OUKA
The University of Osaka
Institutional Knowledge Archive

| | |
|---|---|
| Title | Studies on Cooperation Between Tele/Para-Operatable Semi-Autonomous Robots and Their Operators |
| Author(s) | 窪田, 智徳 |
| Citation | 大阪大学, 2022, 博士論文 |
| Version Type | VoR |
| URL | https://doi.org/10.18910/88100 |
| rights | |
| Note | |

*The University of Osaka Institutional Knowledge Archive : OUKA*

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

The University of Osaka

# Studies on Cooperation Between Tele/Para-Operatable Semi-Autonomous Robots and Their Operators

TOMONORI KUBOTA

MARCH 2022

# Studies on Cooperation Between Tele/Para-Operatable Semi-Autonomous Robots and Their Operators

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

BY

TOMONORI KUBOTA

MARCH 2022

# Abstract

Among the social robots that can interact with humans, operated robots, in which humans control the robot, could alleviate the constraints of time and place at work, and they can bring out more than the capabilities of a single person. In particular, semi-autonomous robots, in which the autonomy of the robots and operation by humans work together, can be of practical use because both functions complement each other during the execution of tasks.

In conventional research on human-robot interaction, the effect of the robot on the interaction partner has been verified in many cases. However, in cooperation between a semi-autonomous robot and its operator, the effect of the robot on the operator through operation and interaction with the robot has not been sufficiently verified. Furthermore, although the tele-operation type has been mainly studied previously, I consider that the relationship between the robot and its operator can be used in the para-operation form, where the robot and its operator exist in the same space. In para-operatable robots, the operator can treat the robot as an avatar and perform tasks, or directly interact with the robot; thus, the robot can play a role that is more than just an operating object for the operator.

By classifying the cooperation between semi-autonomous robots and their operators into two forms, tele-operation and para-operation, this research shows that although both are the same operatable robots, they have different effects. Additionally, the research discusses the further development of the operatable robots. I believe that this research can clarify the effects of the robot on the operators that have not been considered in the practical use of semi-autonomous robots.

This research consisted of three studies. The first aims to demonstrate a method for building an autonomous dialog function that is essential for semi-autonomous robots through a field experiment. In the proposed method, a tele-operator first operates a robot in the field where it is applied, and then the autonomous dialog function is gradually improved using the operation logs. I selected the Internet live service as a place to apply an android robot and conducted a field experiment to improve the autonomous dialog function for the android that interacts with viewers using the proposed method. Thereafter, the conversation between the android with the system and viewers was evaluated as the android's responses were natural. From the results, this study demonstrates the ways of building autonomous dialog ability, which is a basic function of semi-autonomous robots.

In the second study, focusing on the cooperation between tele-operatable semi-autonomous robots and their operators, the influence of the autonomous behavior of the robot on the operator was clarified through a laboratory experiment. I hypothesized that the attitude of tele-operators may align with that of the robots through tele-operation. The experimental results indicated that the attitude of participants who operated the bodily behavior of a robot that autonomously expressed an attitude aligned with that of the robot than those who did not. Through this study, I revealed the influence of the autonomous behavior of the robot on the tele-operator.

In the third study, focusing on cooperation between para-operatable semi-autonomous robots and their operators, I conducted a field experiment to verify the real-world feasibility of the para-operation form and its influence on the operator through interaction with the robot. I implemented a para-operatable semi-autonomous android and applied it to a boutique. Through an 11-day field experiment, during which the salespersons operated the android, it was observed that the android could contribute to the salespersons. The experimental results indicated the real-world applicability of the para-operation form in this study. Furthermore, it has been demonstrated that an operator of a para-operatable semi-autonomous android can accept the contents of a dialog with the android to be operated, and they can relate to the robot they operate as if it were another person.

Through these studies, by organizing the cooperation between semi-autonomous robots and their operators into tele- and para-operation, I demonstrated that each operation form induces different effects on operators. In cooperation between semi-autonomous robots and their operators, this research clarified the unclear effect on the operators. I believe that this research will contribute to the future applications of semi-autonomous robots. Moreover, I discuss the possibility of interaction between a semi-autonomous robot and its operator, taking advantage of the machine-like aspect of the robot owing to its operatability.

# Table of Contents


| | | |
|---|---|---|
| Chapter 1 | Introduction | 1 |
| 1.1 | Social robots | 1 |
| 1.2 | Cooperation between semi-autonomous robots and their operators | 3 |
| 1.3 | Research purpose | 5 |
| Chapter 2 | Related Works | 9 |
| 2.1 | Human-robot interaction | 9 |
| 2.2 | Semi-autonomous robot | 10 |
| Chapter 3 | Development of Autonomous Dialog Function | 15 |
| 3.1 | Background | 15 |
| 3.2 | Experiment 1: Collecting a chat dialog corpus | 21 |
| 3.3 | Experiment 2: Implementation of autonomous dialog system | 27 |
| 3.4 | Discussion | 36 |
| 3.5 | Conclusion | 40 |
| Chapter 4 | Cooperation in Tele-Operation | 41 |
| 4.1 | Background | 41 |
| 4.2 | Method | 44 |
| 4.3 | Result | 55 |
| 4.4 | Discussion | 57 |
| 4.5 | Conclusion | 62 |
| Chapter 5 | Cooperation in Para-Operation | 63 |
| 5.1 | Background | 63 |
| 5.2 | Para-operated android robot system | 66 |
| 5.3 | Field experiment | 74 |

5.4 Result . . . . . 77
5.5 Discussion . . . . . 83
5.6 Conclusion . . . . . 92

Chapter 6 Conclusion 95
6.1 Summary . . . . . 95
6.2 Future directions . . . . . 98

Appendix 103
A Dialog script used in the experiment in chapter 4 . . . . . 103
B The list of ten paintings used in the experiment in chapter 4 . . 105
C Questionnaires used in the dummy task in chapter 4 . . . . . 106

Acknowledgments 109

Bibliography 111

Publications 123

# Chapter 1

# Introduction

## 1.1 Social robots

Research on robots that can participate in human communication is being conducted actively because such robots could provide communication support for humans, compensate for labor shortage in communication tasks, or allow humans who are far away to communicate with each other through tele-operated robots [1, 2]. Such robots can be categorized into autonomous and tele-operated robots. An autonomous robot can perform certain tasks by itself and can be used in various scenarios, such as shopping malls [3], hotels [4], offices [5], schools [6], and museums [7]. Previous research has demonstrated that a human interlocutor can anthropomorphize such robots [8]; therefore, it is expected that humans will recognize these robots as human-like agents. A tele-operated robot is controlled by a tele-operator to perform tasks; the tele-operator can transmit their presence to a remote space through the robot [9, 10]. A tele-operated robot can be used in various domains, such as education [11], childcare [12], customer service [13, 14], and communication support [15, 16]. In the case of tele-operated robots, the tele-operator recognizes a robot as a thing (tool) rather than an agent with its mind. Thus, based on the above scenario, it can be inferred that there are two sides to social robots: they are either recognized as agents or things.

In this study, I considered a scenario in which a person tele-operated some parts of the behavior of an autonomous robot. In this setting, the robot is operated by a tele-operator while retaining its autonomous functions; thus, the robot is classified as a semi-autonomous robot [17, 18]. Under this scenario, the autonomy of the robot

and human tele-operation cooperate simultaneously to accomplish a task. For example, the autonomous behavior of the robot controls the bodily movements and the human-performed tele-operation is responsible for the verbal behavior of the robot. This approach, in which multiple intentions of different agents, such as autonomous systems and humans, are collectively employed to operate a single robot is called collaborative control [19].

Studies conducted on semi-autonomous social robots [20–22] have aimed to reduce the burden of tele-operators and realize robotic applications in the face of technical difficulties. Currently, it remains difficult to develop a fully autonomous social robot that can communicate with humans in the real world owing to technical difficulties such as sensing environmental information or understanding language. A semi-autonomous robot that functions with help from a human to compensate for its incomplete autonomous system has been adopted to overcome these difficulties. In the case of tele-operated robots, it is difficult for a human tele-operator to control the entire behavior of the robot because it comprises consciously performing both verbal and various nonverbal tasks, including subtle expressions that are often produced unconsciously in human-human communication. Therefore, it is expected that a semi-autonomous style that employs the autonomous behavior of the robot along with inputs from tele-operation will yield successful results. Additionally, a previous study proposed a model in which one tele-operator operated multiple semi-autonomous social robots [23].

When humans use semi-autonomous robots as their avatar, the cooperation between human operation and robot autonomy allows humans to participate in social activities by extending their capabilities and presence; such an avatar in which humans can work with the assistance of the autonomous functions of the robot system has been proposed as a cybernetic avatar [24] and is being pursued. For example, in work, having robots with a certain degree of autonomy allows people to work while operating multiple robots, and in everyday dialog situations, the autonomous function can assist people with the content of their responses and how to proceed with the conversation. I think that the goal of semi-autonomous robot research is to establish a technology that enables people to expand their abilities and activities by cooperating with robots, such as cybernetic avatars. This research, which deals with the cooperation between semi-autonomous robots and their operators, will contribute to the realization of a society in which people can expand their abilities and sense of

existence by using such robots.

## 1.2 Cooperation between semi-autonomous robots and their operators

As mentioned above, it is important to realize semi-autonomous robots, in which their autonomous functions and human operation cooperate to achieve their goals. However, in the human-robot cooperation in semi-autonomous robots, it is still unclear how the operator of the robot is influenced by the robot to be operated. While much attention has been paid to the influences of robots on their interlocutors [25,26], the influences on robot operators through operating or interacting with the robots to be operated have not been sufficiently examined.

Therefore, in this research, I focus on the relationship between a semi-autonomous robot and its operator, i.e., how the two cooperate with each other. Through this research, I believe that it will be possible to clarify effects and influences that have not been considered in the practical use of semi-autonomous robots. This clarification will lead to more effective cooperation between the robot and its operator, such as maintaining the operator's motivation to continue using the robot, by considering the effect of the robot on the operator as well as on its interlocutor. Alternatively, it may also reveal potential risks to the operator of the robot through operating that were not previously recognized Therefore, I believe that focusing on the cooperation between semi-autonomous robots and their operators is essential for the practical application of such robots in society.

Additionally, when focusing on the cooperation between semi-autonomous robots and their operators, there are two possible forms of robot operation: tele-operation and para-operation forms (Fig. 1.1). tele-operation is a form of operation from a remote place considered conventionally. Para-operation is an unprecedented form in which a robot and its operator are in the same space. Few social robotics studies have dealt with para-operation form in the past, and it is a new form of application of operatable robots. Shimaya et al. [15] adopted an operation form in which a robot and its operator are placed in the same space, and the presence of the operator is revealed to their interactive partner. In the para-operation form proposed in this study, however, the operator does not reveal to their interlocutor that they are operating

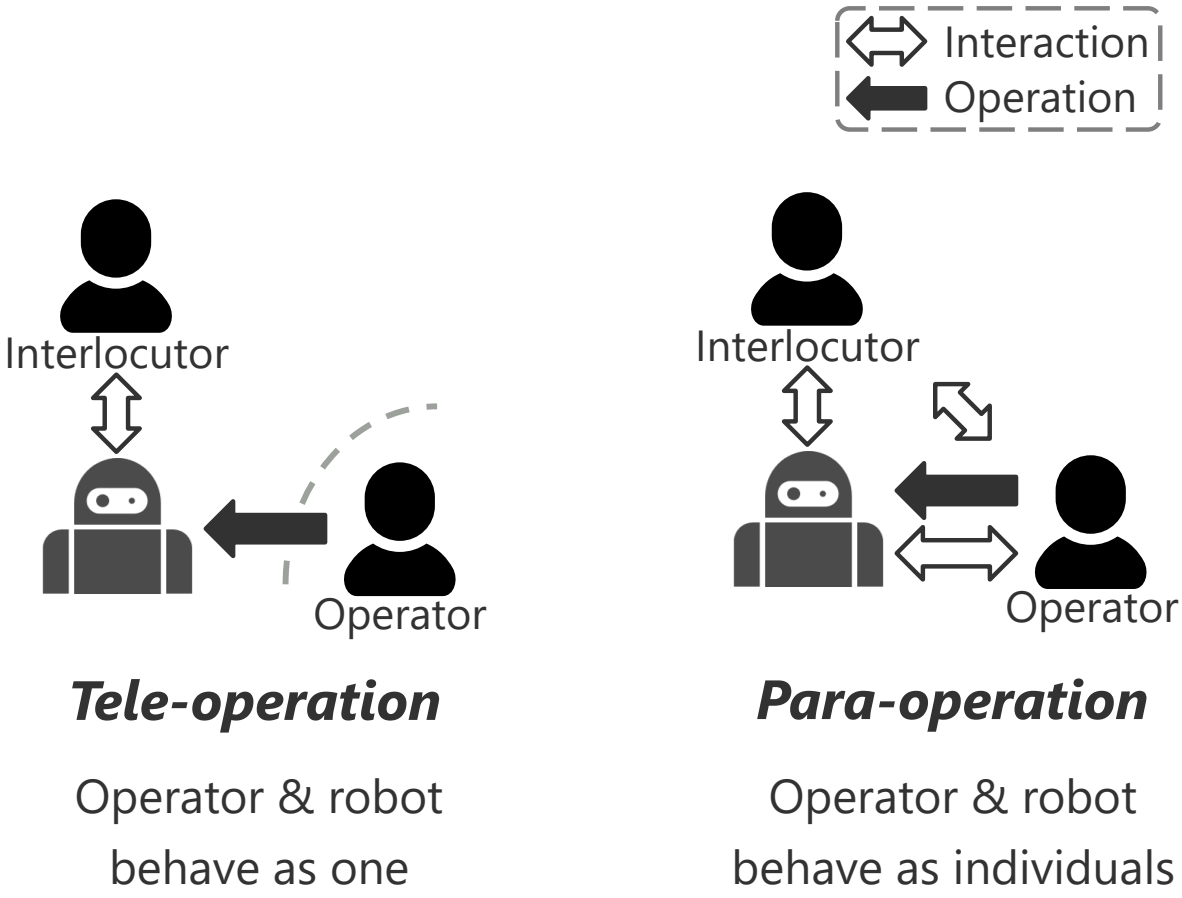



Fig. 1.1 Conceptual diagrams of tele-operation (left) and para-operation (right) style.

the robot. That is, the para-operation form will allow the operator to approach their interlocutors in pairs with the robot, for example, by operating the robot in secret and proceeding with the conversation as they intend. For example, a salesperson working in a store may want to recommend a product to a customer, but may not be able to do so due to concerns about being considered calculating. In such a situation, by using a para-operatable robot, salespeople can let the robot speak on their behalf to say what they want to say but cannot. In conventional tele-operation, a robot and its operator are one and the same, but in para-operation, there is a possibility that they can play different roles and work as two entities. Therefore, by focusing on the coordination between a robot and its operator, operated robots have been operated remotely, but I can now consider para-operation, and different cooperation can be achieved for each of them.

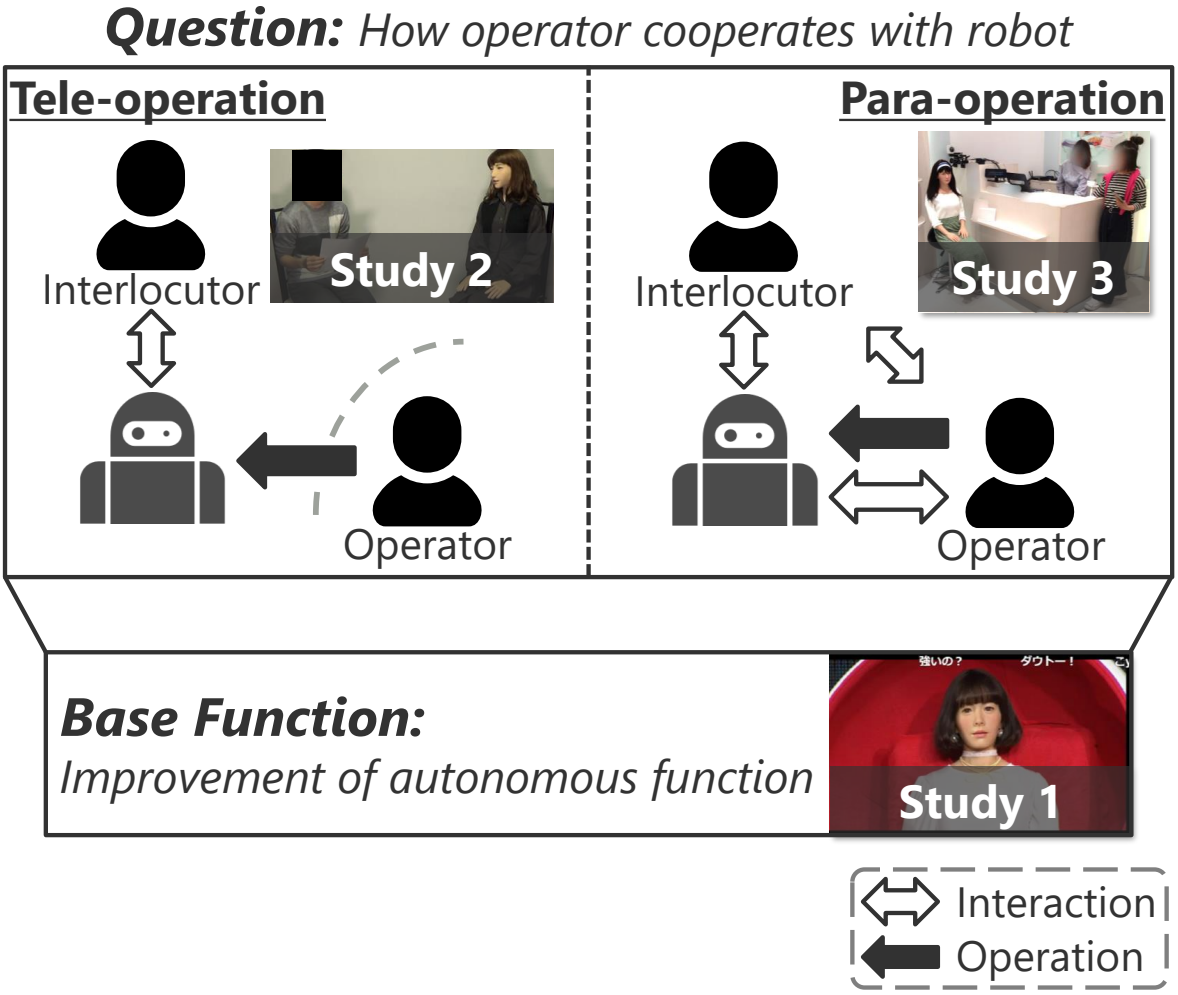



Fig. 1.2 Research map. Study 1 demonstrated how to construct an autonomous dialog function, which is necessary to realize semi-autonomous robots. Then, the cooperation between the robot and its operator was verified in two forms: tele-operation (Study 2) and para-operation (Study 3).

## 1.3 Research purpose

The purpose of this research is, by classifying cooperation between semi-autonomous robots and their operators into two forms: tele-operation and para-operation, to show that each form of operation has a different effect on the operator, and to discuss the further development of semi-autonomous robots. To achieve this, the following three

studies were conducted.

Study 1 Development of underlying autonomous dialog function for semi-autonomous robots

Study 2 Cooperation in tele-operation

Study 3 Cooperation in para-operation

Figure 1.2 shows a research map. In Study 1, I first demonstrated how to construct an autonomous dialog function, which is necessary to realize semi-autonomous robots. Then, the cooperation between the robot and its operator was verified in two forms: tele-operation (Study 2) and para-operation (Study 3).

In Study 1, I demonstrated through field experiments how to build autonomous dialog functions essential for semi-autonomous robots. Semi-autonomous robots are required to have a certain level of dialog capability. However, it is still difficult to generate appropriate responses in situations where robots are applied. In this study, I proposed a method in which a tele-operator first operates the robot in an actual application field, and then the autonomous dialog function is gradually improved using the operation logs. I selected an Internet live service as a place to apply an android robot and conducted a field experiment to improve the autonomous dialog function of the android that interacts with viewers. Through this experiment, I evaluated the acquisition of autonomous dialog function by the proposed method and the practicality of the method.

In Study 2, I conducted a laboratory experiment to clarify the influence on the operator from the autonomous behavior of the robot in the cooperation between the tele-operatable semi-autonomous robot and its operator. In the experiment, participants were asked to operate a part of the body of an android robot that was autonomously conversing with people and expressing a certain preferential attitude. A behavioral index was used to evaluate whether the participants' preferential attitude aligned with the one expressed by the android through an operation.

In Study 3, in the cooperation between a para-operatable semi-autonomous robot and its operator, I conducted a field experiment to verify the real-world feasibility of a para-operatable semi-autonomous robot and its effects on the operator through interaction with the robot to be operated. In the experiment, I selected a boutique as a place where robots are applied and implemented a system that allows salespeople to para-operate an android robot without being noticed by customers. Through the

actual operation of the android in the store, I evaluated the feasibility of a para-operatable semi-autonomous robot in the real world and how the operator interacts with the robot.

In this study, an android robot [27] with a human-like appearance was used. First, the reason why I used a robot with a physical body instead of an on-screen agent without a physical body is that it is said that a robot can adapt to real-world tasks better than an on-screen agent [28, 29]. In tele-operation, the operator sees the robot on the monitor, and there is no significant difference in Lifelikeness between a collocated robot and an on-screen robot [30]. Furthermore, the reason for using androids instead of mechanical-looking robots is that it has been shown that people feel a higher degree of anthropomorphism and lifelikeness and are more likely to have empathy for a robot with a human-like appearance [31, 32]. It has also been shown that people prefer robots with a human-like appearance rather than those with a mechanical appearance in terms of cooperation [33]. Therefore, in human-agent cooperation, I consider that an android robot with a physical body and human-like appearance can effectively cooperate with the operator in the real world, and the operator can feel a high degree of autonomy on it. For these reasons, I used androids in this research because I believe that they have the potential to facilitate the relationship between semi-autonomous android robots and their operators.

Chapter 2 introduces related works of this research. In Chapters 3, 4, and 5, I describe Study 1, 2, and 3, respectively. In Chapter 6, I conclude this research and discuss future directions of cooperation between tele/para-operatable semi-autonomous robots and their operators.

# Chapter 2

# Related Works

## 2.1 Human-robot interaction

Research on robots that interact with people, such as social robots in this study, has been conducted in the research field of human-robot interaction (HRI) [1, 2]. To realize robots that can relate to people harmonically, various influences of robots on people have been clarified in order to establish how the relationship between robots and people should be designed: for example, the influence of the robot's appearance [34, 35], voice [36], nonverbal behavior [37], the distance between robots and people [38], social touch [39, 40]. These robots are a type of agent. An agent is defined as "an autonomous system that interacts with the outside world" or "a system that appears to be such an autonomous system" [41]. Therefore, HRI is included in the area of human-agent interaction, which deals with the relationship between humans and agents, and research is being conducted comprehensively on artifacts that interact with humans, including not only robots but also comparisons with agents that do not have physical bodies, such as on-screen agents [28, 29].

In order for social robots to interact with people, it is important to research to realize their autonomous functions. Autonomous robots need to have the ability to communicate not only verbally [42, 43] but also non-verbally [44, 45] so that they can interact naturally with people in society. In particular, research on the realization of autonomous dialog capabilities is closely related to research on dialog systems. There are various implementation methods for dialog systems, depending on the form of dialogs, such as task-oriented dialog or task-less dialog such as chatting [46]. In recent years, there has been an increase in the number of applications of autonomous robots

in society, but the realization of a dialog robot that can perform tasks in a fully autonomous manner in society still presents technical difficulties and is still under research.

On the other hand, tele-operated robots, which are different from autonomous robots, are also being researched and applied [9]. Tele-operated robots aim to realize a system that allows a human to operate a robot in a remote location to convey their presence to the remote place and to interact with people there while feeling a sense of presence as if they were in the remote location. Various forms of tele-operated robots have been proposed, and research has been conducted on their appearance effects on interlocutors [10, 47], operating systems with a high sense of presence for the operator [48], and adaptation of the operator to the robot's body through operation [49]. I think that this research, which deals with operatable robots, is also one of the HRI studies, including the interaction with the interlocutors and the realization of interfaces for the operators to operate the robots appropriately.

## 2.2 Semi-autonomous robot

Semi-autonomous robots have also been proposed and researched, which fall between the autonomous robots and operated robots mentioned above [17]. Such robots have shown potential for practical use in which the autonomy of the robots and tele-operation by humans work together to accomplish difficult interactions. In tele-operated robots described above, the non-verbal behavior of the robots was sometimes realized autonomously [50]. Therefore, the application of fully operated robots, in which a human operator completely operates all the actions of the robot, is not realistic, and even the so-called operated robots are usually equipped with partial autonomous functions. This research deals with robots whose operators are aware of the autonomous behavior of the robot and can recognize a certain degree of agency in the robot.

### 2.2.1 Autonomization

In semi-autonomous robots, it is as important to realize appropriate autonomous functions as it is in autonomous robot research. However, it is not easy to realize a system with autonomy that is appropriate for the situation in which the robot is applied. Here, a method has been proposed in which the robot is first operated by a

human, and the operation log is used to realize the autonomous functions of the robot required for the application [51]. Yet, the practicality of a method to improve the autonomous dialog function of a robot by using its operation log in the application field has not been demonstrated. In this research, I demonstrate a method to improve the autonomous dialog function for semi-autonomous robots through the operation in the field.

### 2.2.2 Cooperation

Research has been conducted to achieve effective cooperation between semi-autonomous robots and humans: for example, research on the blending of autonomous functions and operations [20, 52, 53], and interfaces for operating robots [54]. Alternatively, there is research into semi-autonomous operation in which the system displays possible options to the operator [14]. It is important to study the appropriate mixing of autonomous functions and operations, as well as robot operation methods related to interfaces, but these are not covered in this research.

In a semi-autonomous robot, the operator and the autonomous functions share control of the robot's behavior. In some cases, the human is responsible for the robot's interactive functions, in other cases, for sensor functions, and in other cases, for bodily behavior such as locomotion. Based on the human model proposed by Norman [55], Shiomi et al. [17] summarize the functions of a semi-autonomous robot into three layers: Reactive layer, which controls unconscious actions in response to sensor information; Behavioral layer, which controls routine actions that humans can perform without any particular awareness; and Reflective layer, which controls conscious actions that humans perform while thinking. This study dealt with the case where the operator operates the behaviors corresponding to the reflective layer in this classification, such as speech and facial expression in a dialog. Additionally, in Study 2, I compared the effects that the operator receives through operations corresponding to the reflective layer and reactive layer in terms of which behaviors of the robot the operator is in charge of.

In semi-autonomous robots, a robot and its operator cooperate to accomplish the task, but there is another form of human-robot cooperation: when an autonomous robot and a human work together [56,57]. For example, a human and a robot can work together to perform a task, such as carrying something, while communicating with each other [58]. Such collaboration of robots and humans is categorized as human-

robot collaboration (HRC) research [59], especially in HRI. This research deals with the cooperation between an operatable semi-autonomous robot and its operator, and I think that this research is also included in the scope of HRC research, where humans and robots work together. However, few studies focus on the relationship between an operated robot and its operator. I believe that this research will contribute to the understanding of the interaction/collaboration between robots and humans in a form that has not been considered previously.

In research on semi-autonomous robots, the effect of reducing the operator's burden is implicitly considered, but few studies focus on the psychological influence that the operator receives from the robot through cooperation. Several studies have concentrated on the tele-operator of semi-autonomous robots [52, 53, 60]. Nishio et al. [61] revealed that the feelings of the tele-operator changed based on the autonomously displayed facial expressions of the tele-operated android robot. In this study, I focused on whether the attitude of the tele-operator changes according to that of the semi-autonomous robot autonomously expressed in words. In Study 2, I investigated whether the attitude of the tele-operator aligns with that expressed by the semi-autonomous robot when they control a part of the robot's body and when they and the robot act together.

### 2.2.3 Para-operation style

Although tele-operation is the mainstream of conventional operated robots, this research proposes a form of para-operation. Previously, Shimaya [62] used an operation form in which the operator operates a robot that exists in the same space. In a series of studies by Shimaya et al., the robot's interlocutor interacts with the robot knowing that the operator is operating it. By having a robot-mediated dialog, the goal is to lower the barrier for the interlocutor to directly interact with the operator, and to promote dialog between the interlocutor and the operator [15, 63, 64]. The form of para-operation proposed in this research allows the operator to approach the interlocutor by extending their presence by not revealing the operator is operating the robot. Therefore, although para-operation and operation forms in studies by Shimaya et al. are similar in terms of the positioning of the three parties participating in the dialog, the usage of the robot and its purpose is different. There are no examples of the application of operated robots in the form of para-operation, and in this research, I examine the possibility of its practical use and discuss the further

development of the para-operation form.

# Chapter 3

# Development of Autonomous Dialog Function

## 3.1 Background

### 3.1.1 Objective of this study

In this chapter, I describe a study in which a method for constructing an autonomous dialog function, which is essential for semi-autonomous robots, was proposed and demonstrated through a field experiment[*1]. Semi-autonomous robots require a certain level of dialog ability, but it is still difficult to generate appropriate responses in situations where robots are applied. In this study, I propose a method in which a tele-operator first operates the robot in an actual application field, and then gradually improves the autonomous dialog function using the operation logs. Using this method, an autonomous dialog system with a certain level of autonomous dialog ability was implemented in a place where the robots are applied, and the practicality of the proposed method was evaluated through a field experiment.

There is a growing demand for robots that provide services and entertainment, which can communicate with people through natural dialog. While robots have

[*1] This chapter is a translation based on a previously published paper [65]. The copyright of the material that appeared in the paper [65] is retained by the Japanese Society for Artificial Intelligence (JSAI). The material is published here with the agreement of the authors and the JSAI. Please be complied with the Copyright Law of Japan if any users wish to reproduce, make derivative work, distribute or make available to the public any part or whole thereof.

various appearances, an android robot, in particular, has an appearance similar to humans and is therefore expected to be deployed in situations where a human-like presence can be used [27]. Nevertheless, adopting an autonomous dialog function for an android is more challenging than for other agents and media. For example, it has been reported that if the robot's actual function is inferior to the function expected by humans, people tend to develop a poor impression of that robot [66]. Therefore, it is important that an android has the ability to behave and provide appropriate responses like humans.

An important component in the research on dialog systems is the development of a dialog system that can be applied to androids. However, to elicit problems and requirements from the real world, it is also useful to experimentally deploy semi-autonomous androids equipped with dialog systems that can be realized with the current technology in the real society, so that people can interact with these androids and reveal insights on these problems and requirements. By observing and investigating the interaction between people and an android in the real world, I can get ideas about, for example, what kinds of topical conversations would an android be required to participate in, or what kinds of services can be provided by such an android.

Nevertheless, with current technology, it is difficult to realize a fully autonomous android equipped with a dialog system capable of partaking in dialog equivalent to human beings in all real-world situations. This is due to reasons such as the difficulty of speech recognition in noisy environments, such as in typical commercial facilities, and the challenges involved in discourse generation that is needed to respond appropriately during a dialog. A dialog breakdown caused by speech recognition errors or a discourse generation failure will decrease the android's human likeness and will betray humans' expectations for an android.

This study proposed a method for gradually improving an autonomous dialog function for an android using operation logs by first operating the android with a teleoperator in a situation where it is actually applied. In this study, the Niconico Live[*2] (referred to as NicoLive) Internet live service was chosen as the situation of applying the android. In the NicoLive live streaming service, a broadcaster can deliver live broadcasts to users, and broadcasters can talk to users who write comments, which are displayed on the video stream (Fig. 3.1). Using NicoLive as the interface between the android and the users circumvents some of the technical difficulties out-

[*2] http://live.nicovideo.jp/




Fig. 3.1 Android U, the android that interacts with users on Niconico Live. The white character strings (Japanese) on the screen are comments from users. In the actual video stream, these character strings flow from the right to the left of the screen.

lined above, and in particular speech recognition errors are eliminated because users communicate to the broadcaster via textual messages.

Using video streaming as an interface to users makes it easier to enable an android to communicate with the users without significantly impairing its humanlike presence. In a live broadcast where multiple users post textual comments, the dialog system can maintain a level of consistency of the conversation by selecting comments to which it can correctly respond from those that are shown on the video stream. Furthermore, since such a broadcast would involve dialog with multiple users, it is not necessary for the android to always strictly follow the topic of the dialog. For the sake of simplicity, the scenario of use for the dialog system presented in this section is considered to be that in which "android interacts with users on NicoLive". A dialog system that can talk to users on NicoLive could be developed without dealing with the technical difficulties of deploying an autonomous android and collecting a huge number of examples of dialogs that would be needed for constructing the dialog system. In this

section, I present a chat-oriented dialog system that enables an autonomous android to have a chat with users in the NicoLive broadcast scenario.

The current study demonstrates the method to develop a chat-oriented dialog system that enables an android to chat with users over NicoLive. I present the implementation and evaluation of the proposed chat-oriented dialog system developed using the proposed method. The rest of this chapter is organized as follows. In section 3.2, I describe the experiment performed on NicoLive to collect a chat dialog corpus required for developing the autonomous dialog system. In section 3.3, I describe the experiment performed to evaluate the autonomous dialog between the user and the android installed the developed dialog system in, and the results of this experiment. In section 3.4, I reflect on the findings and discuss their implications, followed by the conclusion in section 3.5.

### 3.1.2 Development of autonomous dialog system

Dialog systems are typically classified into two types: task-oriented dialog systems and non-task-oriented dialog systems that chat casually with users [46]. This study deals with the latter. A task-oriented dialog system conducts a dialog with a specific goal. For example, one dialog system guides travelers while arousing interest in them [67]. This system uses information in its knowledge base to provide recommendations and information to users and answer their questions. The current study develops a chat-oriented dialog system over NicoLive.

A dialog system needs to be able to chat informally with users. For example, research has found that, even in the case of a task-oriented dialog system, users of a dialog system sometimes like to have a casual chat that is not directly related to tasks that the system performs [68]. Further, it has been reported that a robot that guides customers in a shopping mall influenced their purchasing behavior by making friends with them, and through not only introducing stores and goods but also including informal conversation aspects such as self-disclosure of the robot itself [69]. Therefore, considering the importance of informal chat dialogs, as a first step, I developed a chat-oriented dialog system that can partake in an informal chat dialog on NicoLive.

Chat-oriented dialog systems typically require an initial configuring phase, where they take as input many rules or a corpus of data, in order to be able to respond to users' utterances. The systems ELIZA [70] and A.L.I.C.E. [71] use algorithms based on pattern matching for the configuration phase. A difficulty with such a method is

that it requires a lot of time and effort to create a large number of rules. For example, A.L.I.C.E. works with about 41,000 rules created by hand. More recently, a method using a neural network to generate responses to users' utterances has also been studied [72]. This study used about 62 million sentences to train a neural network that enables chat dialog. Therefore to adopt this method requires the preparation of an enormous corpus of dialog data.

In this study, I decided to adopt a retrieval-based method that employs dialog pairs. Utilizing a chat dialog corpus of dialog examples, composed of pairs consisting of an utterance from a user and a reply to the utterance, this method finds a past utterance in the corpus that is most similar to a user's utterance input to the dialog system and returns a reply sentence linked with the past utterance as a response. To deploy such a method, therefore, requires a chat dialog corpus.

A previous study constructed a chat-oriented dialog system based on a retrieval-based method, using collections of pairs of utterances and replies from Twitter as a chat dialog corpus [73]. However, since the chat dialog corpus gathered from an SNS consists of dialogs between unspecified users, it is difficult to maintain the consistency of the personality of the dialog system, when it outputs its response purely based on the corpus. In the case of an android, to ensure human-like behavior, it is necessary to collect examples of dialogs that would confer a consistent personality to the android. Therefore, this study adopted a methodology that assumes a consistent personality for the android while collecting a chat dialog corpus.

In order to collect the chat dialog corpus, I first delivered a live broadcast on NicoLive with a human operator tele-operating the android, behind the scenes, like the Wizard of Oz (WOZ) method [74]. On NicoLive because a distributor can show video with audio, users can see the android interacting with them, and it is possible to efficiently collect comments that users actually send to the android. Moreover, since the android can talk with many and unspecified users, diverse utterances can be collected. Furthermore, since a user talks to the distributor through a comment, users' utterances can be gathered in text.

A study by Sugiyama et al. [75] has examined the problem of collecting a dialog corpus for a dialog system that has a consistent personality. In the study, multiple questioners created a series of questions for the dialog system to answer, and only one answerer-provided answers to them in order to confer the dialog system a consistent personality. In Sugiyama et al.'s study, the questioners have presented an image of

the character's appearance. In the method using NicoLive proposed in this study, it is possible in addition to present the actions and voice of the android. Thus, I might expect to gather user utterances such as "Smile for me" or "Tell me a tongue twister". As such, there is no need to use an android to acquire user utterances such as these. Even when the action and voice are presented with a virtual CG agent or the like, there is some possibility that utterances such as the above can be gathered. User utterances that can be expected to be gathered more exclusively in the case of an android are what arise from the fact that an android has a physical body, such as "I want to see you". In this study, the tele-operator interacts with users via an android to collect a chat dialog corpus with consistent reply sentences by prescribing the personality of the android.

In this study, it is considered that dialog in NicoLive is a multi-party dialog. A previous study on a multi-party dialog system using a robot estimated whether sound input to a robot is responsive by using the information on users' movement or waist orientation [76]. However, the scenario over NicoLive is different from the face-to-face multiparty dialog. For example, users in NicoLive cannot make speech utterances. Particularly, in this study, since users in NicoLive comment anonymously, the problem of detecting which user to respond to cannot be dealt with. In NicoLive dialogs, it is possible to decide when and whom to respond by using the information in the context of the dialog, but in this study, the method for responding to users is to select comments to which the dialog system can provide correct responses.

There has also been previous work on a dialog system that has adopted NicoLive as its platform. Radiobots [77] is a system that conducts a real-time dialog with users while reading the news as a radio personality. It adopts NicoLive for achieving real-time dialog between a radio personality and a user but does not use the potential benefits of applying the dialog system to NicoLive as described in this study. The chat-oriented dialog system presented here is based on the following premises for using NicoLive: (1) speech recognition is unnecessary, (2) it is possible to maintain a degree of consistency of the conversation by selecting comments to which the dialog system can correctly respond, and (3) it is possible to use both video and audio.

This study to achieve an autonomous dialog system can be divided into two broad stages. In the first stage, I collected a chat dialog corpus through broadcasting NicoLive programs in which a tele-operator operated an android behind the scenes. In the second stage, I developed the dialog system, performed an experiment of the

android's autonomic dialog with users on NicoLive, and evaluated the dialog system to show the effectiveness of the proposed autonomization method.

## 3.2 Experiment 1: Collecting a chat dialog corpus

In this section, I describe the experiment to collect a chat dialog corpus by distributing live programs on NicoLive and having a chat with users via a tele-operated android by adopting the WOZ method. Additionally, I incorporated a reply recommendation system that autonomously presents a recommendation for responses to users' comments to the tele-operator, that is during this period, the system becomes a semi-autonomous form. And as the chat dialog corpus used by the reply recommendation system grows, I also verify whether the number of times the tele-operator adopts the recommended response increases.

In this study, I used Android U (Fig. 3.1), which is a female type android. She has 19 joints in her head and upper body that can be operated by actuators, and therefore she can express actions and facial expressions that are perceived as humanlike by observers. Additionally, the actuators are pneumatic, so that operational sounds do not interfere with the dialog.

### 3.2.1 Tele-operation system

To gather a chat dialog corpus, I implemented a system that enables tele-operating the Android U to talk to users on NicoLive. With this system, the tele-operator can operate Android U with a GUI (Fig. 3.2). As an example of how to use the GUI, it is possible to make Android U utter a reply to a user's comment with a synthetic speech by selecting a comment to reply to, inputting a reply sentence for that comment, and sending it. Also, it is possible to make her express accompanying actions such as a worried face along with her speech. Simultaneously, a dialog triple consisting of the selected comment, the reply sentence, and the accompanying action is saved in the database, so that it is possible to gather a chat dialog corpus while distributing a NicoLive program.

Here, I describe each part of the GUI for tele-operating Android U shown in Fig. 3.2. In the comment display section, comments posted by users during a NicoLive program are displayed in real-time, and the tele-operator can select a comment to reply to from the displayed comments. In the reply input section, it is possible

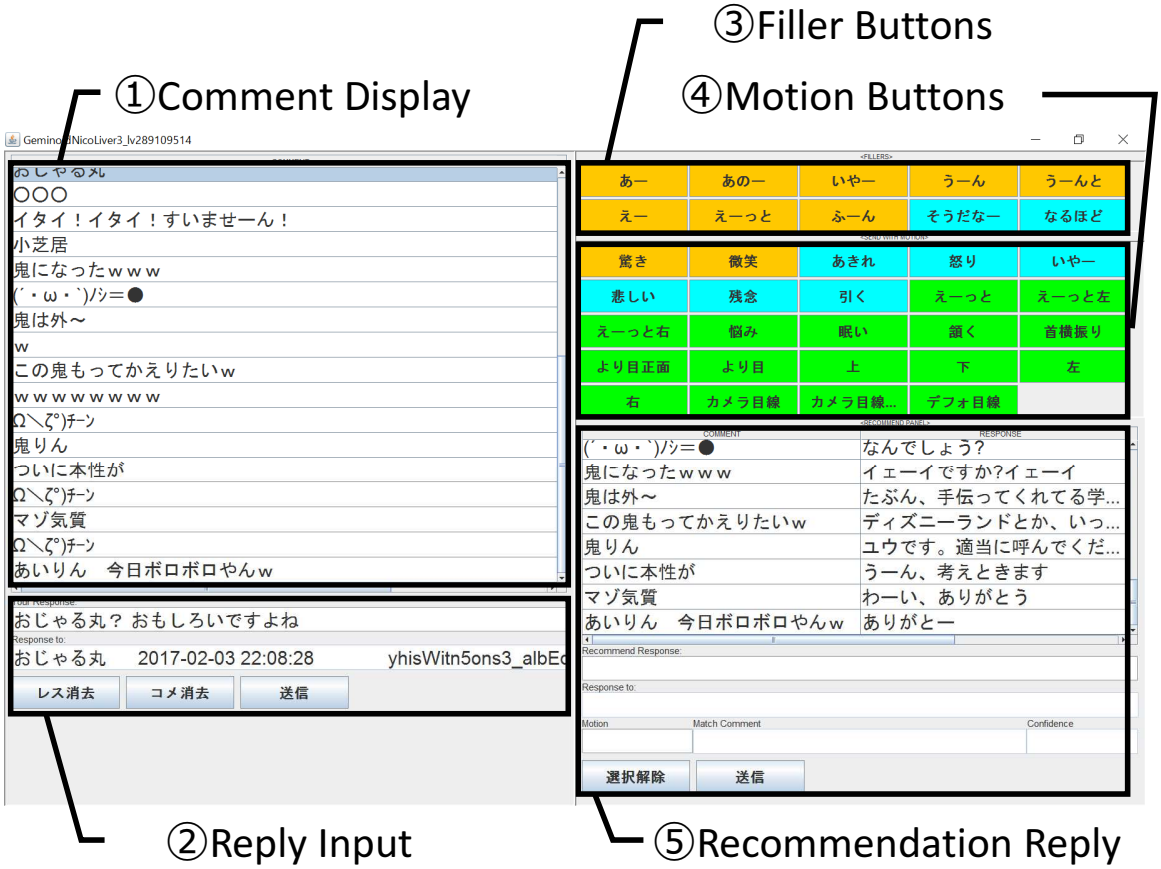



Fig. 3.2 GUI for tele-operating Android U.

to input a reply sentence for the selected comment and to make Android U utter a reply sentence by pressing the send button. Also, the tele-operator can input a sentence that s/he would like to make Android U utter, without selecting the comment to respond to. With a filler button, it is possible to make Android U utter a filler ("hmm", "uh", etc.) associated with each button. With an action button, it is possible to make Android U express action (such as "smile" or "looking at the camera") associated with each of these buttons. Also, by pressing an action button after inputting the reply sentence at the reply input section, it is possible to make Android U utter the sentence and express the action simultaneously. In the recommendation reply display part, it shows reply sentences for each comment recommended by the reply recommendation system. The tele-operator can select the recommended reply sentence and press the send button so that the reply sentence is

uttered by Android U.

Next, I describe the reply recommendation system. When receiving a comment posted during a NicoLive program, this system outputs the response and the accompanying action that the system infers as most appropriate for Android U. For inferring the recommendation reply, the system uses the chat dialog corpus collected in this experiment. As a process flow of this system, a sentence similar to the received comment (input sentence) is searched for in the corpus, and the system finds a matching sentence with the highest similarity. Next, the reply sentence and accompanying action associated with the matching sentence are output as a recommendation reply to the input sentence.

I adopted the Word Mover's Distance (WMD) [78] algorithm to retrieve similar sentences for the reply recommendation system. WMD calculates a similarity value between two sentences using the distributed word representation (i.e. Word2Vec [79]). Therefore, even if two sentences do not contain common words, but their meaning is similar, the similarity value obtained can be high. This system retrieves the previous comment with the maximum similarity value calculated based on WMD as the matching sentence. Additionally, MeCab [80,81] was used for morphological analysis, and mecab-ipadic-NEologd [82,83], which is a dictionary containing neologisms, was used as MeCab's dictionary. Distributed word representation necessary for WMD was acquired from Japanese Wikipedia's full-text data [*3].

### 3.2.2 Experimental method

Using the tele-operation system and reply recommendation system described in the previous section, I distributed NicoLive programs that Android U performed in at Niconico Channel "U's room" [*4]. A screenshot from a program is shown in Fig. 3.1 and a scene of the room used for the live streaming is shown in Fig 3.3. The tele-operator of Android U was restricted to a single person and made to act as Android U. By fixing the tele-operator to one person and prescribing the personality of Android U, her personality could be maintained consistently throughout a program. For the character of Android U in a given program, I prescribed nine items that could be supposed in advance, such as that she is 22 years old, unable to eat food because

[*3] https://dumps.wikimedia.org/jawiki/
[*4] http://ch.nicovideo.jp/geminoidu




Fig. 3.3 Scene of the room used for live streaming program of the android.

she is an android, and living in a university. Along with the characters, I regulated her personality by considering that she behaves gently and uses polite language, and asked the tele-operator to act supposing this personality. If a character not specified in advance on Android U appeared newly in a distributed program, the tele-operator acted keeping both the new character and the prescribed characters. The content of the program by Android U was not specified, and it took the form of chatting freely with users who came to see the program. I did not clarify the existence of the tele-operator for Android U to the users during the program to prevent comments directed toward the tele-operator rather than for Android U.

The programs were distributed to the public users of NicoLive, and everyone who

had an account of niconico [*5] could watch them. However, when distributing the Android U NicoLive programs, people concerned with the program, including the author, did not promote the program on NicoLive or other SNSs. From this, it can be inferred that NicoLive programs were viewed by my laboratory members who knew the program in advance, and users of niconico who found it unintentionally. Since users who watched the program in this experiment are anonymous, it is impossible to clarify the attributes of the users, but it is considered that the attributes were not biased.

As dialog samples between Android U and users were collected to build a chat dialog corpus, dialog pairs generated by the tele-operator during programs were always included in the corpus. After each program, the tele-operator made a response to a comment that was not responded to during the program, and I also included the pair created from this response into the chat dialog corpus. However, I asked the tele-operator not to reply to comments with only symbols such as emoticons or comments that the tele-operator could not understand, and not to reply to sentences outside the program.

After I collected about 600 dialog pairs in a chat dialog corpus, I used the reply recommendation system when distributing a program, even as I continued to grow the corpus. In order to precisely measure how many comments were responded to by a recommendation reply, the tele-operator had to always adopt a recommendation reply if it was appropriate as a response.

### 3.2.3 Results and discussion

In this experiment, I distributed programs 20 times (about 25 hours in the total delivery time) and collected a chat dialog corpus with 4,460 dialog pairs. Since I did not specify the content of the programs, it can be said that the content was different each time due to users' comments. There are standard utterances from users in the programs, such as questions asking for the age of Android U, and comments referring to her hairstyle. However, these utterances are not repeated, and therefore it is considered that new kinds of user utterances were obtained from every program.

Fig. 3.4 shows the transition of the proportion of the number of recommendation replies that the tele-operator adopted to the total number of comments and the total

[*5] http://www.nicovideo.jp/

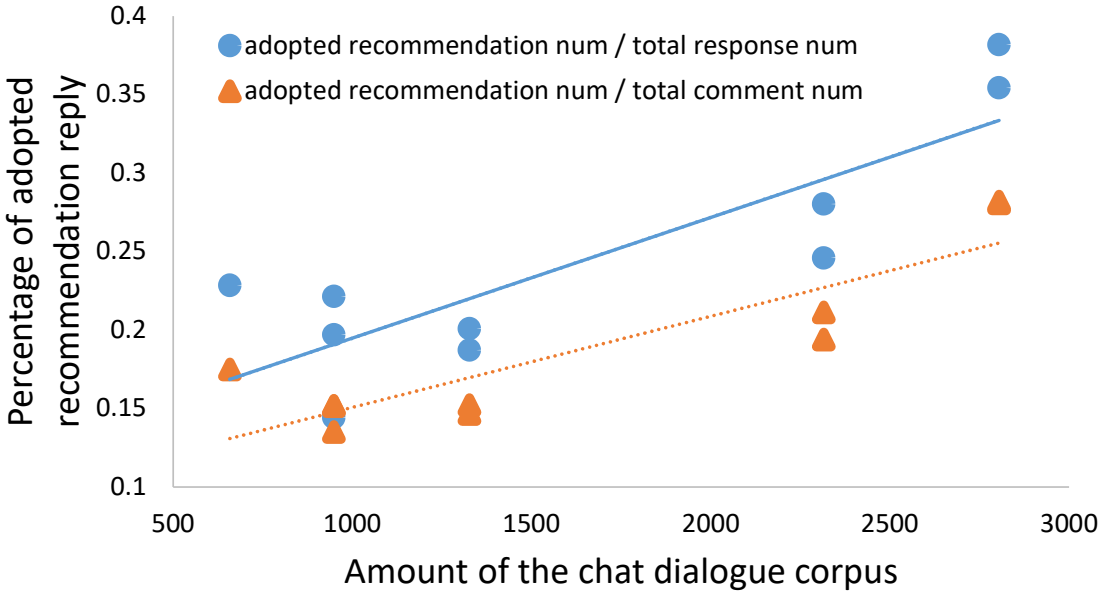



Fig. 3.4 Relationship between the amount of the chat dialog corpus and adopted recommendation reply ratio. The horizontal axis represents the number of dialog pairs included in the corpus.

number of responses. The total number of comments is the number of comments posted by users during programs. The total number of responses is the number of comments responded to by the tele-operator, including the responses using the recommended reply.

I calculated Pearson's correlation coefficient between the size of the chat dialog corpus and the number of adopted recommendation replies. The result is shown in Fig. 3.4. The correlation between the amount of the chat dialog corpus and the number of adopted recommendation replies / the total number of comments was positive and the correlation coefficient was 0.89 ($p < .01$). Also, the correlation between the amount of the chat dialog corpus and the number of adopted recommendation replies / the total number of responses was positive, with a coefficient of 0.85 ($p < .01$).

These results indicate that the collection of the chat dialog corpus through this experiment was successful. Further, Fig. 3.4 shows that when increasing the size of the chat dialog corpus used by the reply recommendation system, the proportion of the number of adopted recommendation replies to the total number of comments and

the total number of responses increased.

User utterances such as "Wave hands." and "Look here and smile to me!" were also included in the chat dialog corpus. I can consider that these utterances were acquired because Android U's appearance and actions were presented to the users. I also obtained user comments such as "Your voice is cute." or "Can you tell us a tongue twister?", and it is considered that these comments were acquired because the voice of Android U was presented to the users. Moreover, user comments such as "I want to go to see you." or "Can I go to see you this time?" were gathered. It is considered that these comments were acquired because of the android having a physical body, and thus if a computer agent was used, I would not obtain these comments.

## 3.3 Experiment 2: Implementation of autonomous dialog system

In this section, I describe the implementation of the chat dialog system "NicoLiver4U" that enables Android U to autonomously have a chat with users on NicoLive, using the chat dialog corpus collected in the experiment described in the previous section. Following this, I describe the results and discussion of the experiment conducted to evaluate the capability of NicoLiver4U based on a questionnaire, by showing subjects the movie of actual dialogs on NicoLive between users and Android U in which NicoLiver4U was installed.

### 3.3.1 dialog system "NicoLiver4U"

NicoLiver4U generates a reply to a comment by adopting a retrieval-based method using WMD to search for a similar sentence from the chat dialog corpus collected in the prior experiment. This retrieval method is identical to that used by the recommendation system referred to in the previous section. Also, NicoLiver4U uses the same morphological analysis method and distributed word representation for calculating WMD as the reply recommendation system. Similarly, for sentence retrieval, NicoLiver4U finds a past comment whose WMD similarity value is the maximum for the input sentence to NicoLiver4U, but only when the WMD similarity value exceeds a certain threshold is the reply sentence linked to the found past comment output as

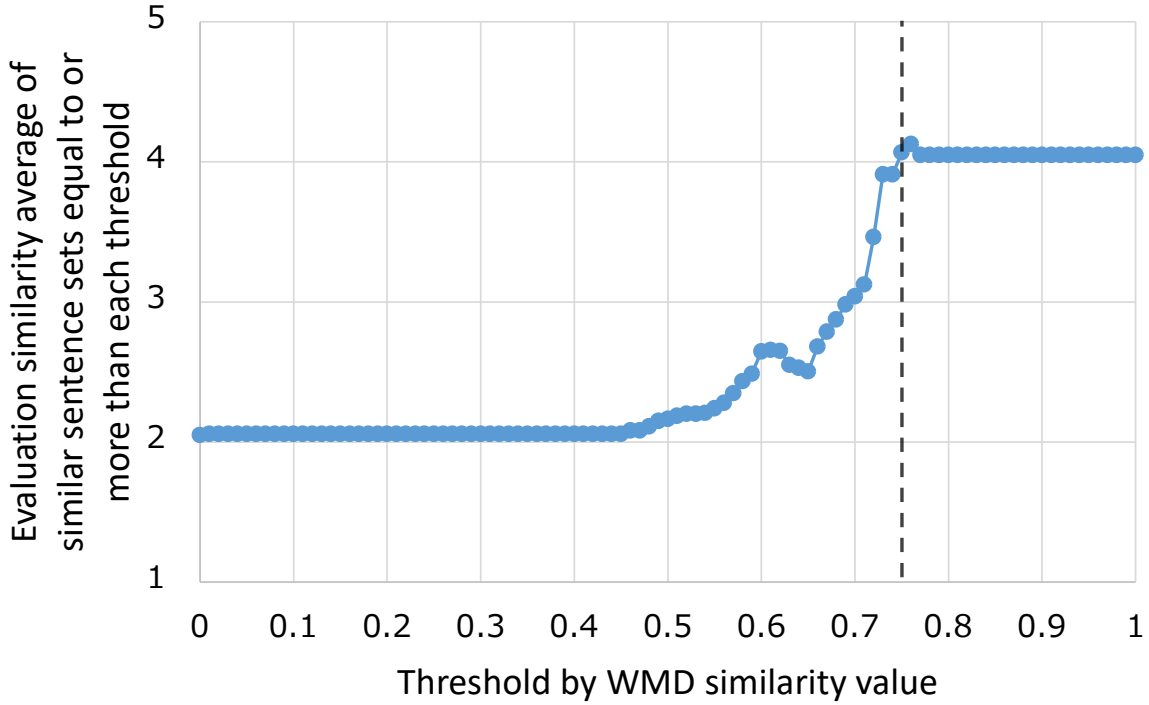



Fig. 3.5 Result of the preliminary experiment to decide the threshold for the similar sentence retrieval.

a reply. This is because a reply sentence linked to a past comment with low similarity to the input sentence is likely to be inappropriate as a reply to the input sentence. The WMD similarity value ranges from 0 to 1, and the closer the value is to 1, the more similar are the sentences [*6].

I conducted a preliminary experiment to determine the above-mentioned threshold value. First, 50 user comments were extracted as test data sentences from the chat dialog corpus. I then obtained the top five sentences with the highest similarity in the remaining corpus for each of the 50 sentences in the test data. Next, I conducted the following steps.

[*6] As a property of WMD, I mention that WMD similarity value taking the value of 1 doesn't mean complete match of two sentences. Since WMD calculates the similarity between two sentences using the distributed word representation, for example, if the same distributed word expression is obtained for word A and word B depending on how to obtain the representation, the WMD similarity value of "I like A." and "I like B." is 1.

1. Three evaluators evaluated the similarities of each set of similar sentences using a five-point Likert scale
2. For each set of similar sentences, I calculated the evaluation similarity value of each similar sentence set by averaging the similarities provided by the three evaluators
3. For each set of similar sentences, I calculated the average of the evaluation similarity value of each similar sentence set exceeding a threshold, while changing the threshold from 0 to 1 in steps of 0.01

The result of the preliminary experiment is shown in the graph of Fig. 3.5. As shown in Fig. 3.5, the average of the evaluation similarity remained at 2 and 4 before and after the thresholds of 0 and 1, but at the above-mentioned part of (1) and (2), there were similar sentence sets with evaluation values of 1 or 5. In this study, it was considered that the possibility of obtaining an appropriate response to an input sentence could be increased by sufficiently securing similarities of similar sentence sets. Therefore, I arrived at the threshold value of 0.75, which is the WMD similarity value threshold at which the average of the evaluation similarity exceeded 4. Thus it can be regarded that 0.75 was a reasonable value to secure the similarity of the similar sentence set, that is, to secure the semantic consistency of the response.

NicoLiver4U does not respond to the input sentence when there is no previous comment with a similarity higher than the threshold value. This leverages the feature of NicoLive that "it is possible to maintain a degree of consistency of the conversation by selecting comments to which the dialog system can correctly respond."

However, in the case where only a comment that cannot be responded to has been posted, Android U will fall silent. In order to reduce the unnaturalness caused by the silence, as a dialog strategy, it utters monologs in such cases. Examples of monologs are "What people are watching?" or "I get nervous".

Additionally, NicoLiver4U stores its history of replies, and when it is about to utter a reply sentence for the second time, Android U adds content to the original response to express that she is aware that this has been mentioned before, such as "I told you a while ago." However, when speaking a reply sentence three or more times, such addition is not performed.

Table 3.1 User comment, detected similar comment in the corpus and reply from the android during the autonomous dialog using NicoLiver4U. All the following sentences were translated from Japanese.

| User Comment | Detected Similar Comment in the Corpus | Reply from the Android |
|---|---|---|
| Wow | Wooow! | Surprised? |
| Hello | Hello | Hello |
| Cute | It's cute!! | Thank you |
| Today is warm. | Today is cold. | It's cold. The outside seems to be very cold. |
| Great | It's great | Thank you. Yes. |
| How old are you? | Hey U, how old are you? | I'm 22. |
| 22 years old!! | ··· and became 22 years old. | Although I'm 22 since I was born. |
| You're young. | Young! | Thanks. Glad to hear that. In fact, I was made as a 22 years old robot. |
| Beautiful | You're beautiful. | Really? Come on, I'm flattered. |
| Hey | Hey | What? |
| Amazing | Amazing! | Is it? |
| Cute | Cute | Am I? I'm blushing. |
| You're scary. | Scary! | Am I scary? Grrr |
| You're not scary. | Scary! | Am I scary? Grrr. Maybe I said it before.* |
| Grrr | Grrr | Grrr? |
| Grrr | Grrr | Grrr? Maybe I said it before.* |
| How many digits of Pi do you know? | How many digits of Pi do you know? | 3.1415926535. I think that remembering a large number of digits is a waste of memory. |
| Funny lol | Funny | Good, good! Very glad to hear that. |

* "Maybe I said it before." was an appended reply sentence mentioned in section 3.3.1.




Fig. 3.6 A scene of the movie of Android U's autonomic dialog shown to subjects. The white character strings (Japanese) on the figure are comments from users. In the actual video, these character strings flow from the right to the left of the screen.

### 3.3.2 Experimental method

I conducted a dialog with a duration of three minutes on NicoLive between users and Android U installed with NicoLiver4U. Table 3.1 shows this dialog in detail: users' comments that NicoLiver4U responded to, the detected similar comment in the corpus corresponding to each user comment, and the reply from Android U during the autonomous dialog using NicoLiver4U. This dialog happened during the NicoLive program "Press Conference on Androidol "U" Raising" which was broadcast live on February 15, 2017. This program was open to the public and could be watched by anyone who had an account in niconico. Also, in the above dialog, I used the entire chat dialog corpus collected by the previous experiment.

In order to evaluate the dialog on NicoLive between users and Android U with NicoLiver4U, I conducted the experiment by showing the part of the movie that cut

out only the scene of the three-minute dialog from the above program to subjects and asking them with the questionnaire on the content of dialog [*7]. The participants in the experiment consisted of 14 university students in their twenties (11 male and 3 female). Fig. 3.6 shows a scene of the movie that shows the state of the dialog used as an experimental stimulus. The movie contained all the responses of Android U shown in Table 3.1.

However, the movie used as an experimental stimulus included performers and stage devices other than Android U. Therefore, when conducting the experiment to evaluate NicoLiver4U's autonomous dialog between users and Android U, I taught subjects to focus on the conversation. While it cannot be said that the effects of performers and stage devices on the study participants were eliminated, it is considered that these effects were minimized. A reason for using this movie clip as an experimental stimulus was to evaluate NicoLiver4U in a situation close to where it is actually adopted. Here, a situation where NicoLiver4U is adopted is where many and unspecified users are engaged in dialog with Android U on NicoLive. For this reason, I used the above-mentioned movie in which Android U could autonomously talk to as many unspecified users as possible.

In the questionnaire, I showed responses to comments made by NicoLiver4U and I sought responses to 4 items on a five-point Likert scale: "whether each response was natural (naturalness of reply)", "whether the whole conversation was natural as a conversation on NicoLive (overall naturalness of conversation)", "whether intervals or gaps between utterances were natural (naturalness of intervals)", and "whether users seemed to enjoy the conversation considering the entire conversation (users' enjoyment)". Additionally, I asked in an interview whether there was any part in the clip that the participant felt was unnatural in the conversation, and where the part was.

The "naturalness of intervals" mentioned above, which refers to gaps or intervals of silence between utterances, might be different in the context of the NicoLive conversation from "intervals" in general face-to-face conversations. For this reason, the experimenter specifically instructed the participants that "intervals" means "intervals between a speech of the android and the next utterance of the android".

Next, regarding "users' enjoyment", the questionnaire specified that the dialog

[*7] The video used as an experimental stimulus can be seen on `https://www.youtube.com/watch?v=g68yHtHacAg`

exchanged not only words but also nonverbal information such as where, when, or how it was said. In other words, the purpose of this question was not to evaluate whether users seemed to enjoy the conversation itself but to evaluate it by including nonverbal elements such as circumstances. Because NicoLiver4U is meant to be operated on NicoLive, this question was included to confirm in particular whether NicoLiver4U being deployed in the system enables the android to autonomously have a chat with users on NicoLive.

In total, Android U responded to 21 comments during the autonomous dialog with users, and in these, there are 18 patterns of replies excluding duplication. Therefore, on the question on the naturalness of her reply, subjects evaluated it with the text of these 18 pairs of users' comments and Android U's responses. Specifically, I showed subjects pairs of users' utterances and Android U's response presented in Table 3.1.

### 3.3.3 Results and discussion

The results of the evaluation experiment based on the questionnaire are shown in Fig. 3.7. For each evaluation value, the larger the value is, the more favorable the impression is. Additionally, since I performed the evaluation using a five-point scale, I conducted a one-sample t-test with a chance level of 3. The average value of the "naturalness of reply" item is 3.98 ($p < .01$) and its standard deviation is 0.73; the "overall naturalness of conversation" item is 3.29 ($n.s.$) and its standard deviation is 0.88; the "naturalness of intervals" item is 3.64 ($p < .01$) and its standard deviation is 0.48; and the "users' enjoyment" item is 4.07 ($p < .01$) and its standard deviation is 0.70. For the "naturalness of reply" item, the evaluation value was calculated as the average of 14 subjects' evaluated values for each response. 8 subjects answered that there was a scene where they felt some awkwardness in the conversation. Specifically, this was a scene where the same reply was repeated, and Android U uttered "Yeah, it's cold." in response to a user's comment "It's warm."

From the result of the "naturalness of reply" item, the response by NicoLiver4U to each comment was evaluated as being significantly natural. It is considered that this result was because NicoLiver4U had appropriate replies to users' utterances in the corpus and the threshold value for determining whether to output the response was set at a reasonable value, and therefore this result indicates that the threshold value was effective.

The result of the "overall naturalness of conversation" item suggests that, when

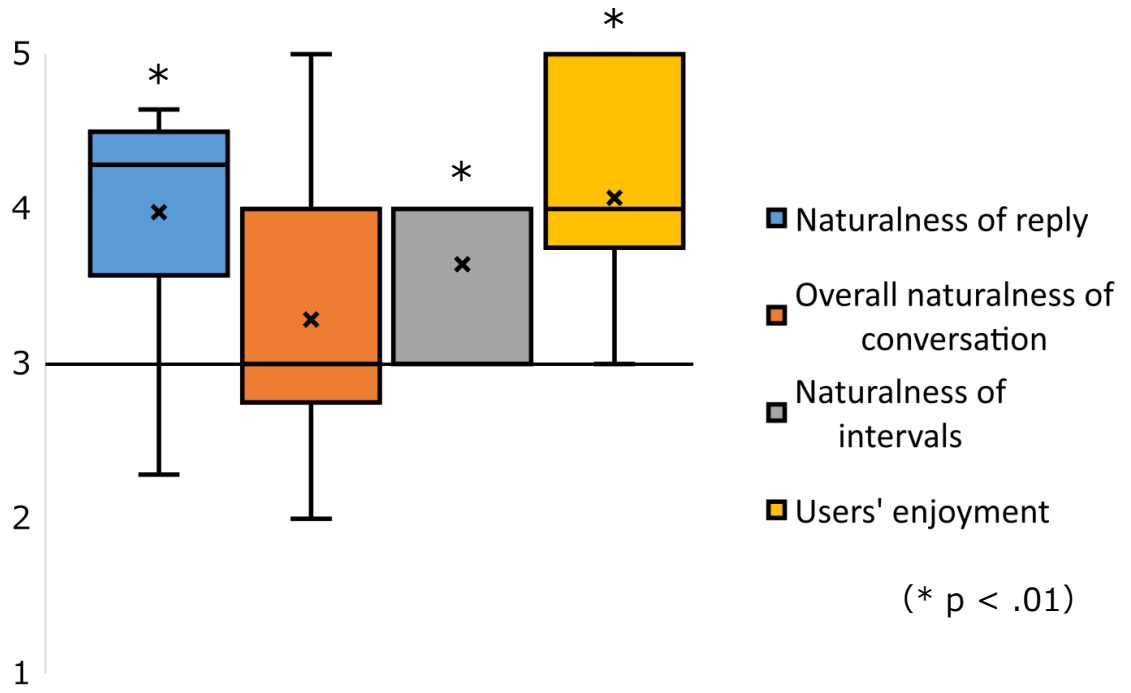



Fig. 3.7 Result of subjective evaluation of the dialog by participants who observed the dialog between Android U and users (Chance level comparison considering 3 as "Neither").

considering the whole of the conversation on NicoLive, the conversation was not sufficiently natural. The number of subjects who answered that there was a scene in which they felt some awkwardness in the conversation exceeds half, and therefore that the scene mentioned as an unnatural response by Android U might affect this result. To address the awkwardness of scenes such as this, where the same reply was consecutively uttered, it might be useful not to generate the same reply sentence twice consecutively, or not to generate a response if it counts as a repetition within a short time, by remembering the output history.

From the results of the "naturalness of intervals" item, intervals in the conversation between users and Android U with NicoLiver4U were evaluated to be significantly natural. NicoLiver4U would not utter anything for up to 5 seconds when only comments that could not be responded to were posted. Therefore, there was a possibility that an interval caused by this silence might result in an unnaturalness in the conversation. This result suggests that I was able to incorporate the proposed feature of

the NicoLive dialog that "it is possible to maintain a degree of consistency of conversation by selecting the comment to which the dialog system can correctly respond", without causing any unnaturalness of the conversation. However, since it cannot be said whether "the 5 seconds" used in NicoLiver4U was the optimum gap, it could be useful to experiment and find the optimum number of seconds for the gap so that it does not make the android's silence seem unnatural.

Based on the result of the "users' enjoyment" item, users who had a chat with Android U seemed to enjoy it significantly. While it is important to know whether a dialog system can generate natural responses as an evaluation index of these systems, it is also important that users who talk to the dialog system are satisfied. Therefore, it cannot be said that NicoLiver4U generates natural responses based on the result of "overall naturalness of the conversation", but it can be said that it entertains users.

NicoLiver4U used the chat dialog corpus containing 4,460 dialog pairs collected in the experiment mentioned in 3.2. Bessho et al. [73], in their study of chatbots conversing with users on a console, examined the transition of the naturalness of the chatbot when increasing the number of dialog pairs, which were collected from Twitter. They found, for example, that the naturalness of the chatbot's responses was evaluated lower than the chance level of the five-point scale when the chatbot used 10,000 dialog pairs. On the other hand, NicoLiver4U's responses were evaluated significantly higher than the chance level, notwithstanding that it used only 4,460 dialog pairs. This result seems to be an outcome of the NicoLive feature that "it is possible to maintain a degree of consistency of conversation by selecting comments, which the dialog system can correctly respond to."

Finally, the results revealed insights into areas for improvement in the method for similar sentence retrieval. For example, NicoLiver4U's output response repeated a past comment in the corpus "It is cold." because the similarity value between the past comment and input sentence "It is warm." was calculated highest. This reply was cited by participants as being awkward in the questionnaire. This is because there was no past comment, which was similar to "It is warm.", and the WMD algorithm used the distributed word representation acquired by Word2Vec. Word2Vec defines a word vector depending on the context in which the word is used. Words such as "warm" and "cold" often have similar contexts of use, and therefore their word vectors are similar. For this reason, WMD might compute a high similarity value for two sentences that have contrary meanings. In future work, I intend to address this

problem in order to make NicoLiver4U's response more natural, by improving the accuracy of similar sentence retrieval.

In this field experiment, since one robot talked with many users, I consider that the android was able to achieve a natural dialog without strictly considering the context of the dialog. In such a situation, it is easy for a robot to take the lead in the dialog, as it can change the replying user each time. However, in a retrieval-based dialog algorithm, if the context is not taken into account in a one-to-one dialog, the system will probably generate a response that ignores the context. Therefore, in a one-to-many situation such as in this experiment, the system that does not consider context might be able to achieve a natural dialog, but this will be discussed in detail in the future.

## 3.4 Discussion

In this section, I provide an overall discussion of this study. Based on the subjective evaluation results mentioned in section 3.3, NicoLiver4U was evaluated significantly higher than the chance level in the naturalness of its responses and users' enjoyment of its dialog. This confirms that the chat-oriented dialog system developed in this study was able to generate natural responses to a certain extent and entertain users. Therefore, this study demonstrated the applicability of the method to develop a chat-oriented dialog system that has the ability to talk to users on NicoLive naturally, by circumventing the difficulties of deploying an autonomous android in the real world with the NicoLive interface and leveraging the chat dialog corpus, which has a relatively modest number of dialog pairs. From this, the dialog system implemented using the proposed method has acquired a certain degree of autonomy. Of course, there is a limitation that this study has only been able to confirm the effectiveness of the method in the NicoLive situation. However, I believe that the method of autonomization using operation logs is an idea that can be applied regardless of the place of application or the form of tele or para operation style. Therefore, I consider that this study has clarified a method for realizing an autonomous dialog function important for semi-autonomous dialog robots that can respond in a way that is appropriate for the application field.

The overall conversation between NicoLiver4U and users on NicoLive was evaluated as not sufficiently natural. This result suggests that in order for the overall

naturalness of the conversation to improve, it is important to improve the human likeness of the android's nonverbal behavior, such as emotional expressions and body movements during utterances. It has been shown that the expression of natural behavior by an android leads to a positive evaluation in dialogs with a human [84], and therefore in this system, it is necessary to consider the expression of nonverbal information in further detail.

A relevant point to consider is that the autonomous dialog using NicoLiver4U was evaluated by third-party subjects, and not by users who participated in the dialog. It could be useful to have the users evaluate the system, but this is challenging on NicoLive because the users are anonymous. Here, a conversation in NicoLive can be regarded as a one-to-many conversation, between the distributor on the one hand, and the users on the other. It is conceivable that users frequently change roles between being the speaker and being bystanders in the conversation. Thus, a user on NicoLive also becomes a "third party" who looks on at the conversation at various times, and therefore the evaluation by third parties in this study is a valid evaluation of NicoLiver4U. However, there may be issues and insights that can only be clarified from the direct evaluation of the users who participated in the dialog, such as whether the response from the android to their comment matches with their beliefs. If there is a platform similar to NicoLive in which I can identify users who participate in the dialog, I could perform such a direct evaluation. The evaluation of NicoLiver4U by a user who participated in the dialog is an issue to be addressed in the future.

The largely successful achievement of the objectives for NicoLiver4U paves the way for further possibilities for the application of a chat-oriented dialog system on NicoLive. For example, interweaving such a dialog system capable of an informal chat with an information provisioning system can result in an effective information-provisioning medium. Nevertheless, for an android that can be an information provision medium based on applying the system developed in this study, there is the restriction that the android can only work with NicoLive. However, since an android has a physical body, it can appear simultaneously, or alternatively, between NicoLive and the real world. One can thus expect an android that might establish a mutual relationship with a user not only face-to-face but also through a NicoLive dialog. Further, in a face-to-face dialog, an android can also interact with users through a touch display [85], and therefore it is possible to interact with people using such a dialog system even in a face-to-face scenario. There is thus a certain value in applying the system constructed

in this study in the realization of an android that can be an information-providing medium in NicoLive.

The total number of users who watched the NicoLive distribution where the android performed an autonomous dialog with users described in section 3.3 is 6,468 people and the number of subscribers following the Niconico channel "U's room" was 584 people (as of June 6, 2017). NicoLive is easily accessible, and the system I developed in this study can be shown to many people.

### 3.4.1 Applications in other situations

The NicoLiver4U dialog system constructed in this study depends on the feature of NicoLive that "users can comment at any time", but it can be possible to apply the system to situations in which an unspecified audience communicate with one person, such as in classes or lectures. However, there are technical problems such as the difficulty of speech recognition to be overcome, and therefore it is not easy to apply the system constructed in this study as it is. Nevertheless, the findings obtained through this study can be valuable for real-world operations of such systems.

Other live streaming media have features similar to NicoLive, such as YouTube Live [*8] and LINE LIVE [*9]. Thus, even if NicoLive stops or changes its service, it is possible to apply the system to an alternative medium. As the forms of communication between people become more variegated and the live streaming media described above are used by a higher number of people, it is meaningful to explore, research and develop useful services on popular media.

All dialogs included in the chat dialog corpus collected in this study are gathered via NicoLive. Therefore, using the chat dialog corpus under circumstances other than NicoLive is likely to be difficult because other circumstances might require different vocabulary and verbalism. In the chat dialog corpus, although there exist some user utterances that can occur in many general situations, such as "How old are you?" or "What is your favorite food?", I have not discussed in detail how many such utterances have been acquired and whether they are likely to occur in situations other than NicoLive. The applicability of the chat dialog corpus collected on NicoLive in other situations is another research subject.

[*8] https://www.youtube.com/live
[*9] https://live.line.me/

There are differences between dialog over NicoLive and face-to-face dialog. dialog on NicoLive is different from face-to-face dialog because the android is not collocated with a user, or a user must communicate with text comments. On the other hand, the common points between these two modes of dialog are the appearance of the android to the user, and the fact that "the android has its physical body" is perceivable in both cases. In the situation where an android interacts with a user on NicoLive since the android is not in the same space as the user, there is a disadvantage that the android's humanlike presence is diminished. However, there is the merit of the users' utterances being possibly influenced by an android having its physical body even on NicoLive dialog, and such utterances were actually confirmed in this study. However, this study did not specifically focus on clarifying the differences between NicoLive dialog and face-to-face dialog, and this must be considered in future work.

### 3.4.2 Limitations

The experimental findings suggest that users were entertained in the conversation with the android using NicoLiver4U. However, this was inferred from responses to the question that "Did users seem to enjoy the conversation, considering the entire conversation?", and this question specified that the conversation included nonverbal elements. Therefore, with this definition of a conversation, it is unclear in this study whether the android could entertain users with only the linguistic elements of the conversation (i.e. the conversation itself). This is a problem that future work can explore.

As described in section 3.3.2, although the movie used as the experiment stimulus for evaluating the dialog between the android and users on NicoLive included performers and stage devices other than the android, I tried eliminating the influence of performers and stage devices on this study participants by instructing them at the time of the evaluation study. As a result, I can consider that the results of evaluations of the dialogs shown in this study have a certain validity.

In this study, I implemented and evaluated a chat-oriented dialog system that enables an android to autonomously chat with users on NicoLive. In the chat between users and the android using the implemented system, the android was evaluated to have entertained users and provided natural replies in the evaluation based on the text of the responses. However, I was unable to obtain evaluations of the system from users who actually had a chat with the android but rather obtained these evaluations

from third-party participants who watched a movie clip of the chat. Therefore, in the experiment conducted in this study, I could not investigate matters that might become obvious only from evaluations provided by the dialog participants themselves, such as whether the response of the android was natural for the participants themselves. I would thus like to obtain evaluations of the dialog between the system and users directly from the users and therefore obtain a more accurate evaluation of the system.

## 3.5 Conclusion

This study has demonstrated, through the field experiment, a method for constructing autonomous dialog function that is essential for semi-autonomous robots. The proposed method is that a tele-operator first operates a robot in the field where it is actually applied, and then the autonomous dialog function is gradually improved using the operation logs. I selected the Internet Live Service (ILS), NicoLive, as the place where the robot was applied, and conducted a field experiment to improve the autonomous dialog function of the robot that interacts with viewers. In this experiment, first, the tele-operator operated the responses of the robot talking with users to collect a dialog corpus of comments and robot responses. After the corpus was acquired to a certain extent, the operating system was changed to a semi-autonomous form, i.e., the tele-operator and the developed autonomous dialog system cooperated in replying to users' comments, and the corpus was further collected. This resulted in the acquisition of the corpus containing 4,460 pairs through about 25 hours of live broadcasting in ILS. After that, to evaluate the performance of the autonomous dialog system, I recorded the conversations between users and the android equipped with the autonomous system. These conversations were then presented to evaluators and the result indicated that the android's responses were natural. From this, the dialog system implemented using the proposed method has acquired a certain degree of autonomy. From the result, this study demonstrated how to build an autonomous dialog function, which is a basic function for semi-autonomous robots.

# Chapter 4

# Cooperation in Tele-Operation

## 4.1 Background

In this chapter, I describe the study on the cooperation between a tele-operatable semi-autonomous robot and its operator, in which the influence on the operator from the robot's autonomous behavior was verified by a laboratory experiment. It is unknown how the attitude represented in the autonomous behavior of the robots affects tele-operators. However, I believe that it is important to know how the operator is psychologically influenced by the robot in order to prepare for real-world applications of semi-autonomous robots. By clarifying this point, I will be able to consider the effect of the robot on the operator through operation, which will lead to the realization of more effective cooperation between them, such as maintaining the operator's motivation to continue using the robot. Alternatively, it may also reveal potential risks that the operator is exposed to through operation, which have not been clarified so far. In the experiment, participants were asked to operate a part of the body of an android robot that was autonomously conversing with a person and expressing a certain preferential attitude. Through this experiment, I evaluated whether the operator's preferential attitude aligns with the one of the robot through operating the robot.

This study considered a scenario wherein operating a semi-autonomous social robot can be considered as "partial role-play" for the tele-operator. Here, role-play implies a process wherein the role-player assumes and performs a specified role [86]. A semi-autonomous robot autonomously determines its role that comprises intention and/or attitude when facing a task. Then, its tele-operator performs the required operation

following the role to support the robot. In role-play, a human plays an assigned role like an actor with their own body. In the case of a semi-autonomous robot, its tele-operator shares the intention or attitude with the robot and plays the role that is autonomously demonstrated by the robot even when the tele-operator only operates a part of the robot's body.

Previous publications on role-play for change in human attitude reported that those experiencing role-play based on a specific role formed an attitude similar to that of the played role [86, 87]. Recent studies have indicated that the attitude of a person who has experienced role-playing as a virtual human in virtual spaces changed according to the role [88–91]. For example, people who experienced acting as superheroes in a virtual space were motivated to perform prosocial behaviors [89, 90]. Thus, it is evident that people change their own attitudes based on the role they play.

From these findings, I hypothesize that even when a human operates a part of another body, the operator behaves like that body; thus, the attitude of an operator changes based on the behavior of the other person. Although scenarios wherein a part of the human body is controlled are rare, they are sufficiently probable for semi-autonomous robots, as assumed in this study.

In this study, I hypothesized that "when manipulating a part of the body of an autonomous robot expressing a certain attitude, the attitude of the tele-operator aligns with that of the robot." Further, I conducted an experiment and analyzed its results to verify the hypothesis. I believe that the findings reported in this study can be used to design effective robot systems for human-robot collaboration and lead the discussion of risks for the tele-operators of semi-autonomous robots. This study suggests the importance of focusing on the influence of interlocutors in face-to-face human-robot interaction and the tele-operators of semi-autonomous robots.

In the experiment, There were two conditions under which a participant operated a robot and one condition under which they did not. I compared the change in attitudes of the participants under all three conditions to check if their attitudes aligned with that of the robot. For the experiment, I used a female-type android robot, ERICA, which has a human-like appearance (Fig. 4.1a). The human-like appearance of an android makes humans feel that it has higher autonomy than non-human-like robots; therefore, it was considered to be the optimal choice. The two conditions created for tele-operating the android included tele-operating the facial expression of the android and tele-operating subtle unconscious gestures by the android's hand. I created

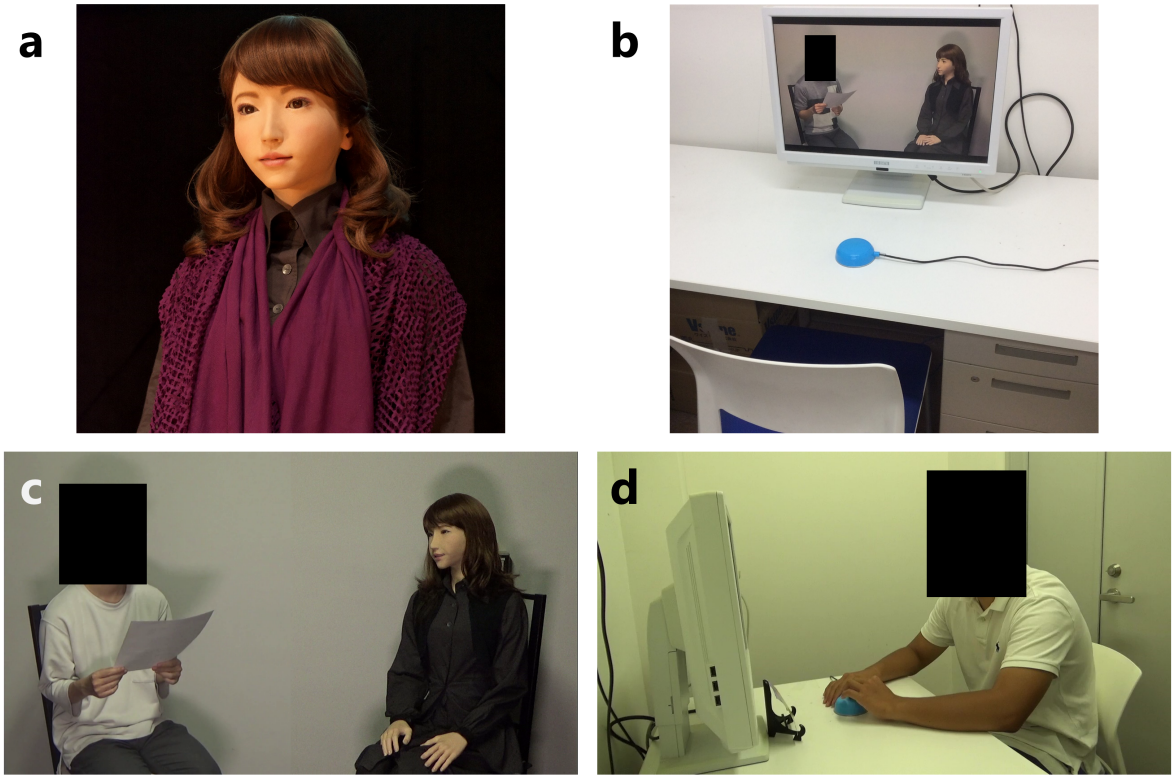



Fig. 4.1 Overview of the experimental setup and scenes: (a) Android robot ERICA used in the experiment. (b) Experiment room. The participants engaged in the experiment in this room. The button in the center of this figure is the one used in the experiment. (c) A scene involving dialog between the android and an experimenter, as watched by the participants on the monitor. The participants were informed that the experimenter was another participant. (d) A scene of the experiment. A participant was engaged in the experiment while watching a scene (like C) on the monitor and pressing the button. A painting printed on the card that the android was recommending to "another participant" was placed in front of the monitor.

two tele-operation conditions because I aimed to study the difference in change of attitude of the participants when tele-operating different parts of the android's body and compare the results of these two conditions. In the former, a participant tele-operated the body part of the android that contributed to the dialog it performed, and in the latter, they tele-operated the part that did not contribute to the dialog.

This study used an android robot in the experiment because it enabled us to prepare variations of operable body parts.

## 4.2 Method

### 4.2.1 Objective

This study performed experiments to verify the hypothesis, i.e., "when manipulating a part of the body of an autonomous robot expressing a certain attitude, the attitude of the tele-operator aligns with that of the robot." Accordingly, I compared whether the attitude of participants aligned with that of the robot under three conditions: 1) SC: a tele-operation condition to make the android smile while observing the dialog over a monitor between the android and the person; 2) HC: a tele-operation condition to perform a subtle hand gesture by the android while observing the dialog; and 3) WC: a no-tele-operation condition besides observing the dialog.

In the experiment, I prepared a scenario wherein the android recommended a painting to a person. A participant was asked under the experimental condition to operate a part of the android's body while watching the scene over a monitor, wherein the android was recommending a painting to a man (Fig. 4.1, b to d). Thus, in this scenario, the participant performed "partial role-play" while operating the part of the android's body that autonomously expressed its favorable attitude toward a particular painting. In fact, the man who listened to the android's recommendation was one of the experimenters because it is needed to equalize the content of dialog between him and the android as much as possible to control each stimulus to the participants. The participants were informed that the man was another participant. The content of dialog between the android and the man was as follows. The android subjectively conveyed to the man how good a certain painting was, including aspects such as the color and composition. First, he was neutral to the recommendations but eventually stated that he liked it after listening to the android's recommendation and accepted the invitation to purchase a postcard of the painting for 216 yen (approximately US$2). The dialog script used here can be found in Appendix.

The following three conditions were set in the experiment: 1) smile condition (SC): a tele-operation condition to make the android express a smile while observing the dialog over a monitor between the android and the man (male: 7, female: 8); 2) hand

Table 4.1 Similarities and differences in the behavior of participants in each condition while observing the dialog between the android and its interlocutor. "Do" implies that the participants in each condition followed the instruction written in the leftmost column. "Do not" implies that the participants in each condition did not follow the instruction written in the leftmost column. All participants pushed the button; however, only those in WC did not tele-operate the android.

| | Smile condition (SC) | Hand condition (HC) | Watch condition (WC) |
|---|---|---|---|
| Push the button | Do | Do | Do |
| tele-operate a part of the android's body | Do | Do | Do not |

condition (HC): a tele-operation condition to cause a subtle gesture of the hand by the android while observing the dialog (male: 7, female: 8); and 3) watch condition (WC): a no-tele-operation condition besides observing the dialog (male: 8, female: 7).

The experiment was conducted based on a between-group design. A total of 45 participants (age: $M = 20.3, SD = 5.1$) were randomly assigned to the three conditions to reduce the gender bias. Seven male and eight female participants each were assigned to the SC and HC, and eight male and seven female participants were in the WC.

A participant was first told that the experiment was conducted at the request of a museum to implement an android working there. For the tele-operation conditions, an experimenter explained to the participant to cooperate with the android and tele-operate the android's bodily behavior while it recommended a painting to the other participant. To tele-operate the android, a push-button was used. The experimenter explained the tele-operation process to the participants; when a participant pressed and held the button down, the android started smiling and retained it in SC (Fig. 4.2a); the hand of the android made subtle open and close gestures under HC (Fig. 4.2b). Furthermore, by telling a participant to press and hold the button only while the android is speaking, I equalized the amount of tele-operation of the android for the SC and HC to the largest extent. In addition, to equalize the behavior of participants between the tele-operation (SC and HC) and no-tele-operation conditions (WC), the experimenter told them in WC to record the section where the android was speaking

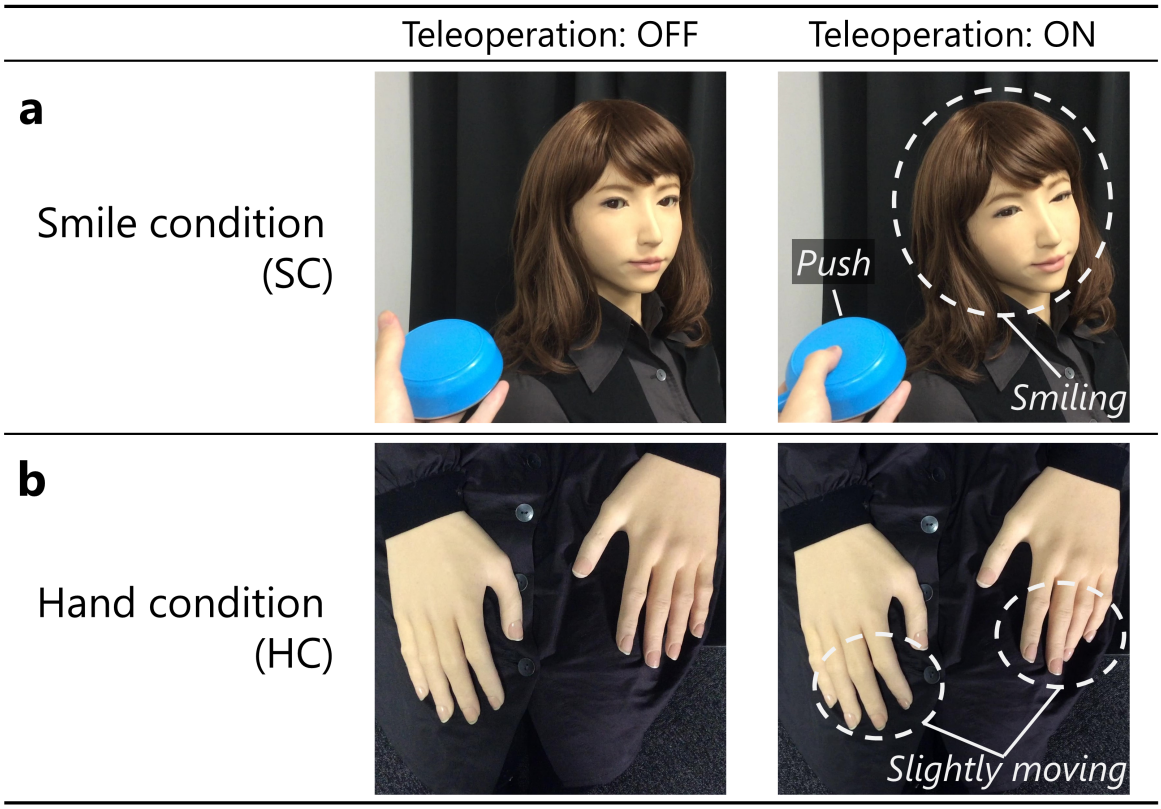



Fig. 4.2 Examples of the behavior of the android tele-operated in the experiment: (a) Smile condition. When the button was pressed and held, the android retained the smile. (b) Hand condition. When the button was pressed and held, the hand of the android kept moving between slightly open and close.

by pressing and holding the button. Thus, the participants in WC pressed and held the button for the same duration as the tele-operating conditions. The similarities and differences in the behavior of participants in each condition while observing the dialog between the android and its interlocutor are summarized in Table 4.1. All participants pushed the button, but only participants in WC did not tele-operate the android.

Further, I controlled the visual stimulus (scene over the monitor when the android talks with the person) provided to the participants during the experiment under all three conditions. For example, when the android was speaking, the subtle motion of the android's hand was automatically executed in SC, the android's smile motion was automatically executed in HC, and both were executed automatically in WC. Thus,



Fig. 4.3 Scene of the dialog between the android and an experimenter, as watched by participants in WC. A small red circle was displayed on the left side of the scene while a participant kept pressing the button.

I ensured that the android movements observed by the participants were the same under all three conditions. Furthermore, in WC, a small red circle was displayed on the left end of the monitor (Fig. 4.3) while the participant pressed and held the button down; thus, the participant was forced to pay attention to the screen in the same manner as the tele-operation conditions.

To test the hypothesis, that is, to assess the change in the attitude of the participants, I applied the method of ranking the paintings considering previous studies [92, 93]. Figure 4.4 illustrates the outline of this experimental flow. In this experiment, participants were asked to rank the 10 given paintings before and after the experiment (see Appendix for the list of 10 paintings). After ranking the 10 paintings, the participants were grouped into one of the three experimental conditions. In all conditions, the android recommended the painting that was ranked 6th by the

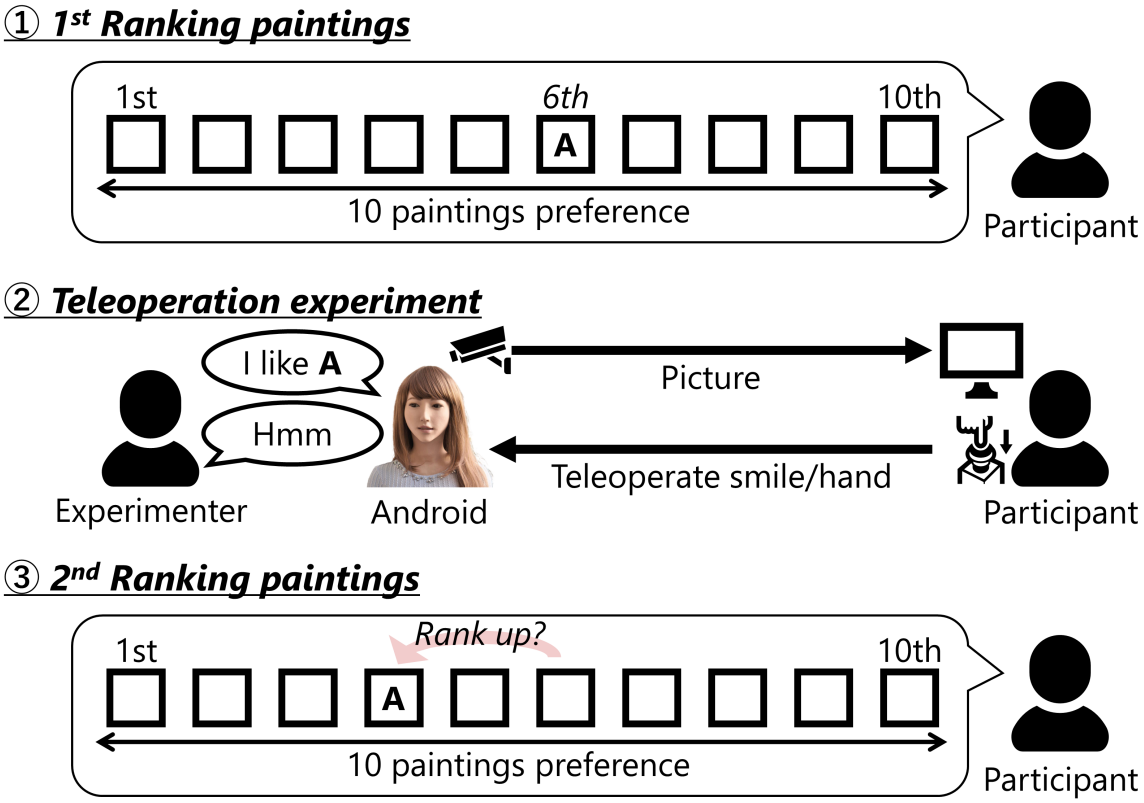



Fig. 4.4 Outline of the experiment: (1) The participants were asked to rank the 10 given paintings. (2) They joined one of the three experimental conditions (SC, HC, and WC). In all conditions, the android recommended the painting that was ranked 6th by the participants to a male experimenter (he was assumed to be another participant by the participants). (3) Fifteen minutes after the experiment, the participants were asked to re-rank the same 10 paintings. If the rank of the painting recommended by the android was higher than that of the first painting, I can argue through this experiment that the attitude of participants is similar to that of the android, thus supporting the hypothesis.

participants to the man. Fifteen minutes after the experiment, the participants were asked to re-rank the same 10 paintings. I assessed the changes in the attitude of each participant by observing the degree of change in the order of the painting that was ranked 6th the first time. A painting that is ranked 6th cannot be considered preferred or not preferred. When its ranking increases during the re-ranking process by even one place, it indicates the change in the attitude of the participants. The

value "6 - the re-ranked 6th painting" indicates the degree of change in the attitude of each participant. If the value is 0, the order of the painting remains unchanged; moreover, if it is positive, the order has increased. If the value becomes positive in SC and HC, and not in WC, I can argue that the attitude of the participants aligned with that of the android based on the results of this experiment, which supports the hypothesis.

From the hypothesis, I aim to confirm the following two points from the experiment: first, if the participants felt that the android behaves autonomously (autonomy of android); second if they felt the sense of controlling the android (sense of agency) in the tele-operation conditions. Only the latter was considered for the participants in the tele-operation conditions (SC and HC). These two points were evaluated by participants on a 9-point Likert scale questionnaire. To gauge the autonomy of android and sense of agency, the following questions were used: "Do you think that the android could think by herself when she spoke?" and "When you were operating the android, did you feel that it was working because you were operating it?"

### 4.2.2 Participants

A total of 45 participants enrolled for the experiment; the gender ratio was almost equal (22 males, 23 females). The age of the participants ranged from 18 to 51 years ($M = 20.3, SD = 5.1$) and they were mostly university students. In addition, the participants were unaware of the topic of the experiment. Before starting the experiment, I explained the precautions that must be observed during the experiment. Then, the participants signed informed consent forms approved by the ethics committee of the Graduate School of Engineering Science, Osaka University, and the methods were performed under the approved guidelines.

### 4.2.3 Materials

This study used a female-type android robot, ERICA (Fig. 4.1a), in the experiment. ERICA has 44 joints on the head, upper body, arms, and both hands; these joints are operated via pneumatic actuators. ERICA can express various facial expressions and motions such as waving and holding hands, thereby making the observers feel a sense of human-likeness. Pneumatic actuators are used in ERICA to ensure that the operation noise does not hinder the dialog. In the experiment, I adopted a

synthesized female voice as ERICA's voice, and ERICA's body movements were generated autonomously by a method based on the work by Sakai et al. [94].

The participants observed the android interacting with a man on the monitor; those who were a part of the tele-operation conditions tele-operated the android using the provided push button. The size of the monitor (IODATA LCD-MF241XWR) was 24.1 in. The push-button had a diameter of approximately 9 cm and could be pressed easily with one hand. The monitor and button are shown in Fig. 4.1b.

### 4.2.4 Experimental Procedure

To prevent participants from knowing the purpose of the experiment, the experiment was conducted under a cover story that this was joint research with a museum. The experiment involved a total of six steps. The scene of the experiment room is shown in Figs. 4.1b and d.

1) Explanation. The experimenter told the participants that the experiment was joint research with a museum to develop an autonomous android that could introduce the paintings to the visitors. Next, as a broad flow of the experiment, the participants were informed that they would be asked to (1) state their painting preference as a general survey of art, (2) participate as a collaborator in an experiment wherein an android recommends a painting, and (3) answer a general survey related to art at the end. For (2), I also told the participants that another participant was participating at the same time as the interlocutor of the android. However, the interlocutor was an experimenter, but this information was not disclosed to the participants.

2) Ranking paintings (1st time). The participants were asked to rank 10 paintings printed on 10 postcard-sized cards as a general survey of art. The participants arranged the 10 cards in the order of their preference. All 10 paintings were Claude Monet's works, and the Japanese name of the work was written on the right side of each card. Prior to ranking, an experimenter told the participants, "Please do not think too deeply; rank the 10 paintings intuitively." During the ranking process, the experimenter stayed out of the room. After the participants completed the ranking, the experimenter recovered the cards and recorded the ranking result in another room.

3) tele-operation experiment. The experimenter told the participants that the purpose of this experiment was to investigate how the person (interlocutor) reacted when they were recommended a painting by the android. The participants were asked

to cooperate in this experiment while watching a scene involving recommendation dialogs between the android and the other participant (interlocutor) on the monitor. For a more realistic experience, the experimenter allowed the participants to see the android in another room. At this time, there was no conversation between the participant and android and the latter performed involuntary natural movements such as blinking and breathing. After returning to the room where the experiment was being conducted, the experimenter turned on the monitor and showed the participants that they could see real-time visuals of the android through the monitor. Next, the experimenter made them have a simple conversation with the android through the monitor to impress the android's autonomy on them. The conversation involved the following: the android first said that her name was "Yu" and asked for the name of the participant. After they provided the name, the android repeated the name correctly and thanked them for participating in the experiment. Then, the experimenter ended the conversation. Further, the experimenter told the participants about the android's preferred painting that it would recommend to the other participant and placed a card of the painting, same as that used in the ranking process, below the monitor screen to ensure that the participants could see the painting. Here, I selected the painting that was ranked 6th by the participants as the painting recommended by the android. In this experiment, considering that there was no bias in the painting preference of participants, I focused on the 6th painting. I selected the 6th painting instead of the 5th because I considered it to have a higher chance of obtaining a better rank in the second round of ranking. Moreover, I did not rank the paintings in a way that yielded a perfect middle rank (e.g., 5th out of 9 paintings) because I presumed that a perfect middle-rank painting might leave a lasting impression on the participants. In addition, the experimenter told the participants that the person who was recommended the painting by the android could buy a postcard of the painting for 216 yen (approximately US$2), and if they bought it, I would consider the recommendation to be successful.

Then, the procedure branched into three conditions, i.e., tele-operating the android's smile (SC) and subtle hand gestures by the android (HC), and not tele-operating but only observing the dialog between the android and another participant (WC). In SC and HC, the participants were informed that each movement could contribute to the android's natural behavior and in WC, the android's utterance section data were required for future system development. Next, the participants

were instructed on how to use the push-button for this experiment: While pushing the button in SC, the android would keep smiling; in HC, the android's hand would keep moving slightly; and in WC, a section where the button was pressed would be recorded. Further, the experimenter asked the participants to keep pushing the button only when the android was speaking. Because the conversational content between the android and the person was almost the same in all conditions, the amount of operation for the button of each participant was equalized maximally via this instruction.

The participants were asked to practice pressing the button several times, and I confirmed that the specified part of the android moved in SC and HC. Further, in WC, a small red circle was displayed on the left edge of the monitor (Fig. 4.3). In SC and HC, the tele-operated movement of the android was shown on the monitor as feedback for pressing the button; therefore, the participants focused on the monitor. In WC, there was no such feedback; thus, I added feedback on the monitor using the red circle to ensure that the participants pay attention to the monitor; moreover, the red circle indicated that the section was being recorded when the button was pushed. The experimenter informed the participants to perform the predetermined button tele-operation when the dialog started; then, the experimenter left the room. The participants pressed the button while watching the dialog between the android and the other participant on the monitor. The flow of dialog was as follows: First, the interlocutor entered the room where the android was present, sat in a chair in front of the android, and they briefly introduced themselves. Then, the android said that she would like to introduce her favorite painting, followed by her reasons for liking it, referring to coloring and composition. The interlocutor initially pretended to not particularly be interested in the painting, but gradually started to show interest after listening to the android's recommendation; eventually, he came to like the painting and purchased a postcard. I prepared 10 different dialog scripts because, in the experiment, the android referred to the painting that was ranked 6th by the participants in the first ranking process. I eliminated the differences between the 10 dialog script patterns by equalizing the structure of each dialog and making the android utterances as identical as possible. The dialog length was approximately 6 min. During the dialog, each participant pressed the button 22 times, for a duration of approximately 3 min (pressing and holding the button down), on average.

The scene of the dialog that the participants viewed on the monitor consisted

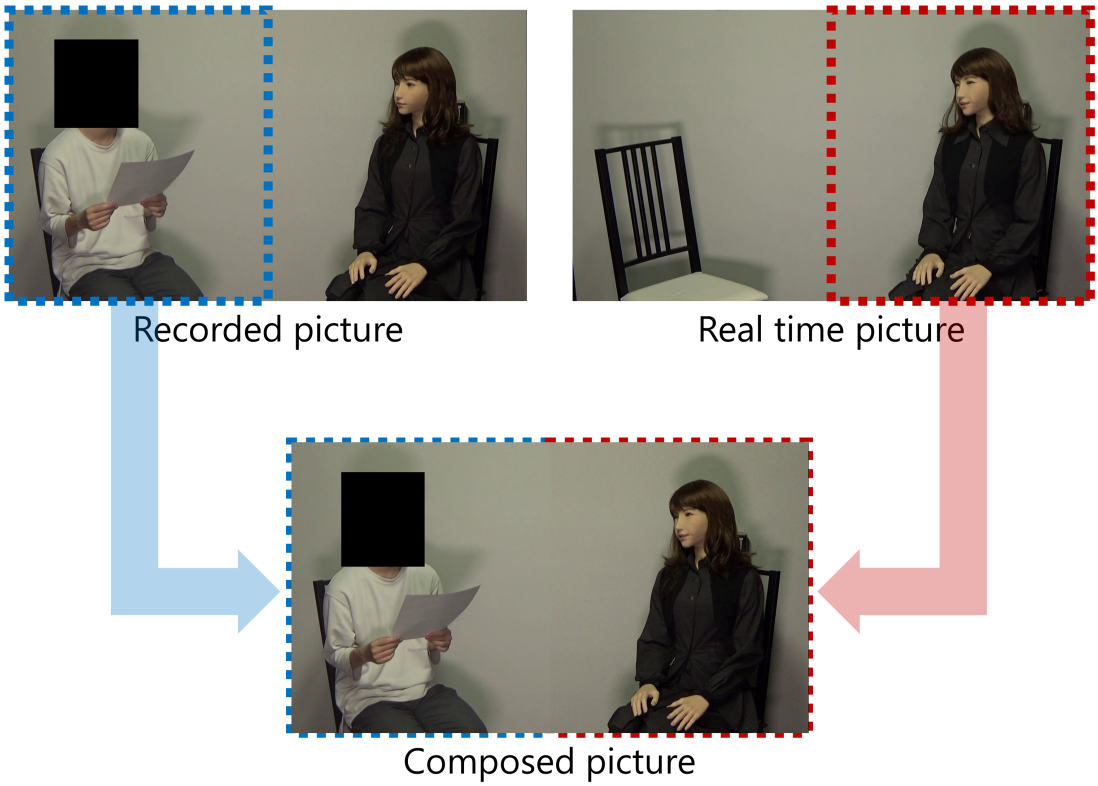
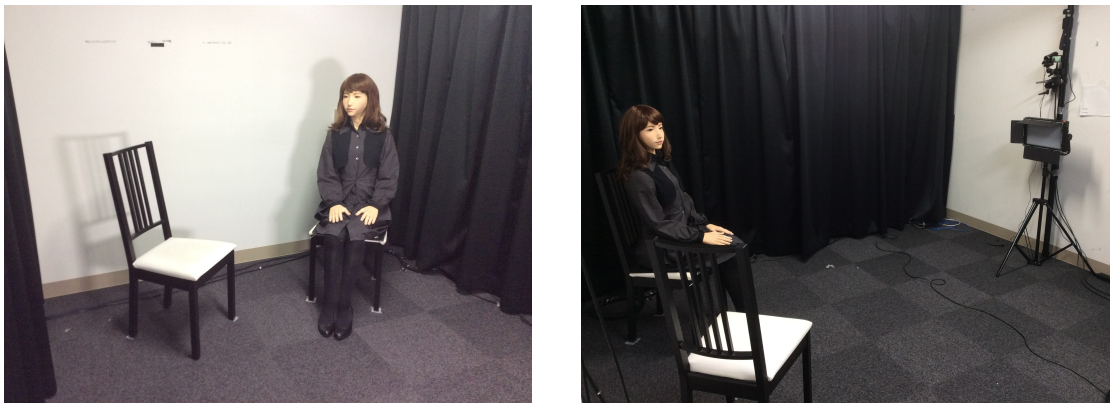



Fig. 4.5 Outline of how to synthesize the real-time and recorded pictures. The right half of the real-time picture and the left half of the recorded picture were combined.

of a real-time picture of the android and a recorded picture of the interlocutor. An overview of how real-time and recorded pictures can be synthesized is shown in Fig. 4.5. The right half of the real-time picture and the left half of the recorded picture were combined. I believed that it was desirable to eliminate the differences in the visual stimuli presented to the participants throughout all experiments. In the experiment, the tele-operation performed by a participant was reflected in the android; therefore, it was necessary to present a real-time video to the participant. I devised dialog scripts and ensured that the comments made by the interlocutor were similar in all 10 patterns. This video was recorded in advance; hence, I used the recorded video in the experiment and presented the same behavior of the interlocuter to all participants. The scene of the room where the video was recorded is shown in Fig 4.6.



(a) View from the side of the camera capturing the android.

(b) View from the android side.

Fig. 4.6 The scene of the room where the video used in the experiment was recorded. Participants also watched the real-time video of the android in this room on a monitor.

4) Dummy task. The participants were asked to work on a 15 min dummy task as a general survey of art. After this, the participants were supposed to re-rank the 10 paintings, and I organized this dummy task to ensure that the impression of the android's preferred painting was not significant. The dummy task comprised closed- and open-ended questions related to art, such as interest in artworks that were not related to the 10 paintings used in the ranking process. The questionnaires used are presented in Appendix.

5) Ranking paintings (2nd time). This was similar to the procedure of ranking the paintings for the first time. The participants were asked to rank the 10 paintings printed on 10 postcard-sized cards. Before ranking, the experimenter informed the participants that "I want the ranking data based on intuition; therefore, do not worry about the results of the first ranking; please rank the paintings according to your current mood." Further, during this ranking process, the experimenter exited the room. After all participants had finished this task, the experimenter recovered the cards and recorded the ranking results in another room.

6) Questionnaire. Finally, the experimenter requested the participants to respond to a questionnaire under the context of hearing their impressions about the experi-

Table 4.2 Statistics of the Wilcoxon rank-sum test with a chance level under each condition for "autonomy of android."

| | M | 95%CI | SD | W | P |
|---|---|---|---|---|---|
| Smile condition | 6.87 | [5.65, 8.08] | 2.12 | 165 | $9.64 \times 10^{-3}$ |
| Hand condition | 6.8 | [5.63, 7.97] | 2.04 | 195 | $3.22 \times 10^{-5}$ |
| Watch condition | 6.93 | [6.14, 7.73] | 1.39 | 202.5 | $8.77 \times 10^{-7}$ |

ment. To determine the "autonomy of android" and "sense of agency," I employed questions that were rated on a 9-point Likert scale.

### 4.2.5 Statistical Analysis

All statistical analyses were performed using R software, version 4.0.3 [95]. To analyze the questionnaires "autonomy of android" and "sense of agency," the Shapiro – Wilk test was performed under each condition to test their normality. When I could not confirm normality, the one-sided Wilcoxon rank-sum test was performed with a chance level (5: neither) under each condition using the R package "exactRankTests," version 0.8-32 [96]. To analyze the results of the ranking of the paintings, under each condition, the Shapiro – Wilk test was conducted to analyze the normality of data, and the Bartlett test to examine the equality of variance in the data. When the normality and equality were not rejected, the two-sided Student's t-test was performed using the Bonferroni correction (multiplying p-value by 3) to test whether the changes in the attitude of the participants differed in all conditions. The effect size (Hedge's g) was also calculated among all conditions using the R package "effsize," version 0.8.1 [97]. This study employed an alpha level of 0.05 for all statistical tests.

## 4.3 Result

The results of the two questions to determine the "autonomy of android" and "sense of agency" are shown in Fig. 4.7. First, the Shapiro – Wilk test was conducted to analyze the normality of questionnaire data. For the response related to "autonomy of android," I could not confirm the normality of data for SC and HC (SC: $W = 0.84, p = 0.014 < .05$; HC: $W = 0.84, p = 0.014 < .05$). Furthermore, for the

Table 4.3 Statistics of the Wilcoxon rank-sum test with a chance level under each condition for "sense of agency."

| | M | 95%CI | SD | W | P |
|---|---|---|---|---|---|
| Smile condition | 8.07 | [7.46, 8.68] | 1.06 | 225 | $6.45 \times 10^{-9}$ |
| Hand condition | 7.53 | [6.98, 8.08] | 0.96 | 225 | $6.45 \times 10^{-9}$ |

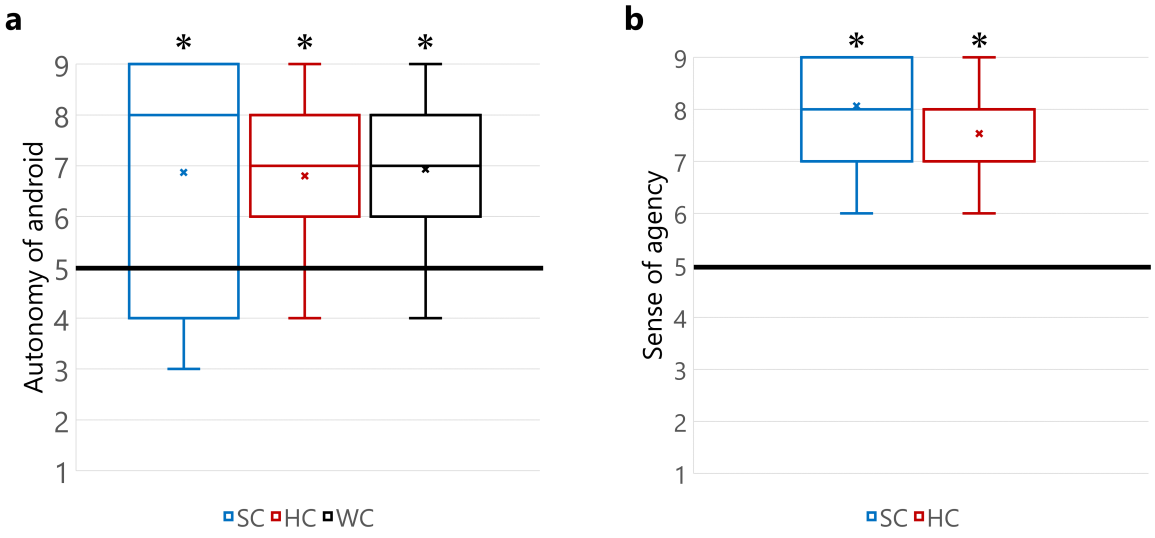



Fig. 4.7 Results of the questions about the autonomy of android ("Do you think that the android could think by herself when she spoke?") and sense of agency ("When you were operating the android, did you feel that it was working because you were operating it?"): (a) Autonomy of android. Based on the results of the Wilcoxon rank-sum test with a chance level (5: neither) under each condition, each evaluated value was found to be significantly higher than the chance level (* $p < .01$). The statistics are listed in Table 4.2. (b) Sense of agency. Based on the results of the Wilcoxon rank-sum test with a chance level (5: neither) under each condition, each evaluated value was found to be significantly higher than the chance level (* $p < .01$). The statistics are listed in Table 4.3.

response related to "sense of agency," I could not confirm the normality of data for SC (SC: $W = 0.80, p = 0.004 < .01$). Therefore, for each response, the Wilcoxon rank-sum test was performed with a chance level (5: neither) under each condition. I confirmed that for both "autonomy of android" and "sense of agency," each evaluated value was significantly higher than the chance level. The statistics for "autonomy of android" and "sense of agency" are listed in Tables 4.2 and 4.3, respectively. Thus, I considered that the "autonomy of android" and "sense of agency" were confirmed by the experiment.

Next, the results of changes in the attitude of participants based on the ranking of paintings were studied. Figure 4.8 shows the results via a graph (SC: $M = 1.33, 95\%CI = [0.36, 2.31], SD = 1.76$; HC: $M = 1.53, 95\%CI = [0.62, 2.44], SD = 1.64$; WC: $M = -0.07, 95\%CI = [-0.64, 0.51], SD = 1.03$). First, the Shapiro – Wilk test was conducted under each condition to test the normality of data; the normality in each condition was not rejected (SC: $W = 0.89, p = 0.070$; HC: $W = 0.95, p = 0.544$; WC: $W = 0.93, p = 0.293$). Next, the Bartlett test was performed to examine the equality of variance in the data; the equality was not rejected either ($\chi^2(2) = 3.99, p = 0.136$).

Finally, the Student's t-test (alpha level of 0.05) was performed using the Bonferroni correction (multiplying the p-value by 3) to determine if the amount of change in all conditions was different. I confirmed that the amount was significantly higher in SC than that in WC, and higher in HC than that in WC (between SC and WC: $t(28) = 2.66$, corrected $p = 0.0385 < .05$, Hedge's $g = 0.944$; between HC and WC: $t(28) = 3.19$, corrected $p = 0.0103 < .05$, Hedge's $g = 1.14$; between SC and HC: $t(28) = -0.32$, corrected $p > 1.0$, Hedge's $g = -0.114$).

## 4.4 Discussion

From the results of the ranking of paintings shown in Fig. 4.8, it can be observed that the participants in the tele-operation conditions (SC and HC) significantly changed their preference to the android's choice of painting in comparison to those in the no-tele-operation condition (WC). Consequently, I considered that the attitude of participants in the tele-operation conditions aligned with that autonomously expressed by the android, thereby supporting the hypothesis. In addition, from the results of the responses to "sense of agency" (Fig. 4.4a) and "autonomy of android" (Fig.

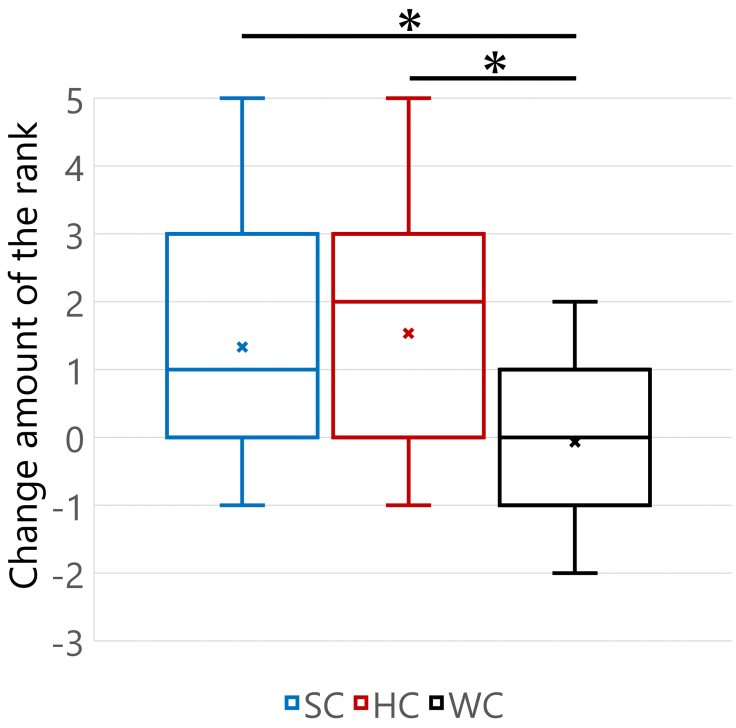



Fig. 4.8 Results of the change in attitude of the participants using the ranking of paintings. The results of Student's t-test (alpha level of 0.05) obtained using the Bonferroni correction (multiplying the p-value by 3) were used to determine if the amount of change in all conditions was different. I confirmed that the amount of change was significantly higher in SC than that in WC, and higher in HC than that in WC (between SC and WC: $t(28) = 2.66$, corrected $p = 0.0385 < .05$, Hedge's $g = 0.944$; between HC and WC: $t(28) = 3.19$, corrected $p = 0.0103 < .05$, Hedge's $g = 1.14$; between SC and HC: $t(28) = -0.32$, corrected $p > 1.0$, Hedge's $g = -0.114$). (* $p < .05$)

4.4b), I concluded that the participants in the tele-operation conditions had a sense of operating a part of the android's body and the android's behavior was sufficient to make participants feel its autonomy. Therefore, I considered that the scenario in which participants tele-operated a part of an autonomous android's body was realized in the experiment.

Moreover, I created two tele-operation conditions in the experiment: the partici-

pants tele-operated the android's facial expressions in SC that appeared to contribute to the scenario of the dialog performed by the android, and unconscious hand gestures, which did not appear to do so. This was done to check the difference in the attitude of the participants who tele-operated different parts (smile and hand gestures) of the android. From the results, it can be concluded that the participants in both SC and HC preferred the paintings that were preferred by the android. Therefore, I can consider that the tele-operator's contribution to the scenario is not related to their change in attitude, and there is a possibility that a tele-operator's attitude aligns with that of the android even by tele-operating various parts such as shoulder, chest, or legs other than the parts adopted in the experiment.

The findings of this study can be used to design better robot systems for human-robot collaboration and work regulation for tele-operators in future semi-autonomous robot applications. The findings indicate both a light side and shadow side for tele-operators, suggesting a way for better human-robot collaboration and the risks associated with operating semi-autonomous robots. Subsequently, I discuss the details of both the light and shadow sides.

A semi-autonomous robot system that can reduce the mental burden of its tele-operator can be developed. Cases where the robot's interlocutors behaved impolitely with it have been reported [98,99]. Therefore, problems can occur when a human tele-operator engages with people through a semi-autonomous robot, and their impolite behavior can mentally burden the tele-operator. This problem can be addressed by applying the findings of this study to reduce the burden of tele-operators; for example, when a semi-autonomous robot expresses a positive attitude, the attitude of its tele-operator may also become positive. Even if a semi-autonomous robot is treated impolitely, building a system in which the robot appropriately expresses a positive attitude will minimize the mental burden of its tele-operator.

For the tele-operators of semi-autonomous robots, the study results imply a risk that their attitude may unconsciously align with that autonomously expressed by the robot. For example, the findings of this study can be used by someone who wants to instill a particular thought in a person. However, if a malicious administrator makes a semi-autonomous robot select a particular political attitude, the attitude of its tele-operator may be aligned to that expressed by the robot. Therefore, the tele-operators of semi-autonomous robots should be aware that their attitude may align with that of the robot. In the future, the administrator of a semi-autonomous robot may be

responsible for notifying the tele-operator of this concern in advance and obtaining consent. The findings of this study will lead to an unprecedented discussion of the ethical considerations for the tele-operators of semi-autonomous robots.

In previous research on social robots, the various influences on a robot's interlocutor due to face-to-face dialog with a robot were widely investigated [37, 44, 100, 101]. This study revealed another type of influence in human-robot interaction, i.e., a tele-operator of a semi-autonomous robot was influenced by its autonomous behavior. Therefore, the influence of operating such semi-autonomous robots should be investigated further. Otherwise, the potential effectiveness or risks associated with the use of semi-autonomous robots may be overlooked.

Based on the results of this experiment, I consider that sense of agency on a robot is necessary for the attitude change of its operator. However, although the contribution to the dialog in which the robot participates may not be necessary (from the HC condition), if it is a negative contribution, attitude change may not occur. In other words, I believe that if the content of the operation is contrary to the purpose of the dialog in which the robot is engaged, the operator will break the dialog and will not come to accept the robot's attitude. I would like to clarify this point through further verification.

Even if there is no sense of agency on a robot, a similar change in the operator's attitude may occur through contribution to the dialog situation. For example, it is possible to operate the environment in such a way as to contribute to the robot, such as controlling a slide that supports the robot's dialog without operating the robot itself. This also needs to be verified, but there is a possibility that the operator will accept the attitude of the robot, although the process may be different from the attitude change observed in this experiment.

In this experiment, the participants were asked to operate the robot, but there are other ways to generate sense of agency in the robot. For example, when a person wears an HMD and acts as a CG character in a VR environment, it has been shown that giving vibration feedback to the throat of the wearer at the same time as the character speaks increases the sense of agency toward the character [102]. Therefore, even through passive haptic feedback, which is different from an operation, it may be possible to induce attitudinal changes like this, and there is a possibility of future verification.

In this study, I conducted a laboratory experiment in which the android had a

conversation partner, so I cannot discuss the necessity of a conversation partner for the observed attitude change. However, I can plan an experiment in which the robot is operated even when there is no interaction partner by creating a situation in which the robot shoots a video in which it expresses some kind of attitude by itself. I would like to clarify the necessity of a dialogue partner through this kind of experiment.

In this study, I have proposed a para-operation form (detailed in Chapter 5), and I consider that attitude change like that in this experiment can occur even in para-operation, but I think this should be verified again. In particular, in this experiment, the operation was synchronized with the speech of the robot. In para-operation, an asynchronous operation can occur, such as operating the robot while engaging in another task. Therefore, it is unclear whether the same kind of attitude change will occur even with the asynchronous operation, and this needs to be verified.

Further, I analyzed why the attitude of the tele-operator changed. Previous studies have shown that a person's attitude changes when they act like someone else through theatrical role-play or by controlling a character in a virtual space [89, 90]. A previous study used Bem's self-perception theory [103] to rationalize such changes in attitude [89]. The self-perception theory states that the attitude and emotions of humans are often determined by their actions and the surrounding environment. In other words, when acting as someone else with a specific attitude in a theater or virtual space, the role-player recognizes their behavior and adopts an attitude that aligns with the role being played. The tele-operator of the semi-autonomous android in this study behaved like the android with a specific attitude by only partially controlling its body movements. The change in attitude in this scenario is similar to that in previous research; thus, I consider that the tele-operator who controlled the facial expression and hand gestures of the android accepted the android's action of autonomously expressing its preference for a painting as their action. In the experiment, it is not clear whether the tele-operator perceives the autonomous action of the android as theirs; I aim to validate this illusion through further investigations. I believe that studies that focused on the actions outside a person's body that are perceived as those of the body can help to enhance the fundamental understanding of oneself; the question of how to determine one's attitude is explained by the self-perception theory as "What is my action?"

In this study, the hypothesis was confirmed through the experiment; however, I should be careful about the generalization of the obtained results. For example, I

cannot conclude whether the same result will be observed when adopting a different type of robot for tele-operation or when operating a robot's verbal behavior. Moreover, in the experiment, I could not sufficiently reduce the bias of age and population of participants; thus, different population and age samples can yield different results. However, I believe that the effect of operating such a robot on its tele-operator is an important contribution of this study to focus on and investigate the novel aspect of semi-autonomous robots. To develop and implement practical social robots in society, I believe that further studies are required to advance the discussion and investigation on semi-autonomous robots and their tele-operators.

## 4.5 Conclusion

In this study, I conducted a laboratory experiment to clarify the influence on the operator from the autonomous behavior of the robot in the cooperation between a tele-operatable semi-autonomous robot and its operator. The tele-operators of semi-autonomous robots also play the role given to the robots and may assimilate their autonomous expression. I hypothesized that the attitude of tele-operators may align with that of robots through tele-operation. To verify this, I conducted an experiment with conditions under which a participant operated a part of the body of an android robot that autonomously expressed a preferential attitude toward a painting and a condition under which they did not. Experimental results demonstrated that the preferential attitude of participants who tele-operated the android aligned more with that of the robot in comparison to those who did not tele-operate it, thereby supporting the hypothesis. Through this research, I have shown one of the influences that the operator receives from the autonomous behavior of the robot in the cooperation between a tele-operatable semi-autonomous robot and its operator.

# Chapter 5

# Cooperation in Para-Operation

## 5.1 Background

In this study, in the context of cooperation between a para-operatable semi-autonomous robot and its operator, a field experiment was conducted to verify the real-world feasibility of the para-operation form and its influence on the operator through interaction with the robot[*1]. The form of para-operation is an unprecedented form of operation in which a robot and its operator are in the same space. By placing the robot, its operator, and their interlocutor in the same space, it becomes possible for the operator to talk with the interlocutor in pairs with the robot, for example, to carry on a conversation while operating the robot according to their own intentions. In the conventional tele-operation, an operator and a robot are one and the same, but in the para-operation form, they can each play a role and work as two entities. Therefore, by focusing on the cooperation between operatable robots and their operators, I can consider the possibility of para-operation, which is different from tele-operation. In the experiment described in this chapter, I selected a boutique as a place where a para-operatable semi-autonomous robot can be actually applied and implemented a system that allows salespeople to operate the robot without being noticed by customers. Through the actual application of the robot in the store, I evaluated the feasibility of a para-operatable semi-autonomous robot in the real world and how its operator interacts with the robot.

Research on robots that can interact with humans and their applications in the

[*1] This chapter is a translation based on a previously published paper [104].

real world is being actively conducted. Recently, robots that provide services while interacting with people in stores are in actual use, as exemplified by Pepper [105] from SoftBank Corp. The effectiveness of such robots working in stores has been verified in the real world through various studies [3, 106, 107], such as the study that realized a robot that guides customers [108] and the study that showed that robots can be accepted favorably in shopping malls [109].

Watanabe et al. [85] have demonstrated that robots can achieve higher sales results than human salespersons, showing the practicality of robots. However, I consider that robots working in stores are not yet fully active in the real world. In stores, it is difficult for robots to recognize voices in noisy environments, such as the voices of people passing by or announcements in the building, and it is also difficult for robots to respond appropriately to the casual speech or dialect of customers. As a result, robots are not able to interact well with people and provide satisfactory services to customers and stores. Additionally, there are other problems such as the cost of introducing robots and the difficulty of choosing what tasks to assign to robots. In 2018, an article[*2] was published that investigated the fact that only 15% of the companies that introduced the aforementioned Pepper plan to renew their rental contracts, highlighting the difficulty of operating robots in the real world.

In addition to this, a reason why robots working in stores have not taken root in the real world is that we have not sufficiently considered what kind of role robots should play for salespeople. In other words, it is necessary for robots to become indispensable to salespeople in order for them to play an active role in actual stores. For example, Niemelä et al. [110] investigated the expectations of store managers through interviews, and Shi et al. [111] implemented a robot after investigating what functions store managers expected from the robot and actually operated it. Both of these studies dealt with what store workers expected from robots in their work, but they did not consider what robots could contribute to salespeople in areas not directly related to the work.

I believe that robots that can not only perform tasks independently and provide services to customers but also work together with salespeople and contribute to them in areas that are not directly related to the tasks, are required in stores. In the relationship between two people working in a store, they not only work together, but also have conversations that are not directly related to their work, and gain mental

[*2] https://xtech.nikkei.com/atcl/nxt/column/18/00466/101000001/

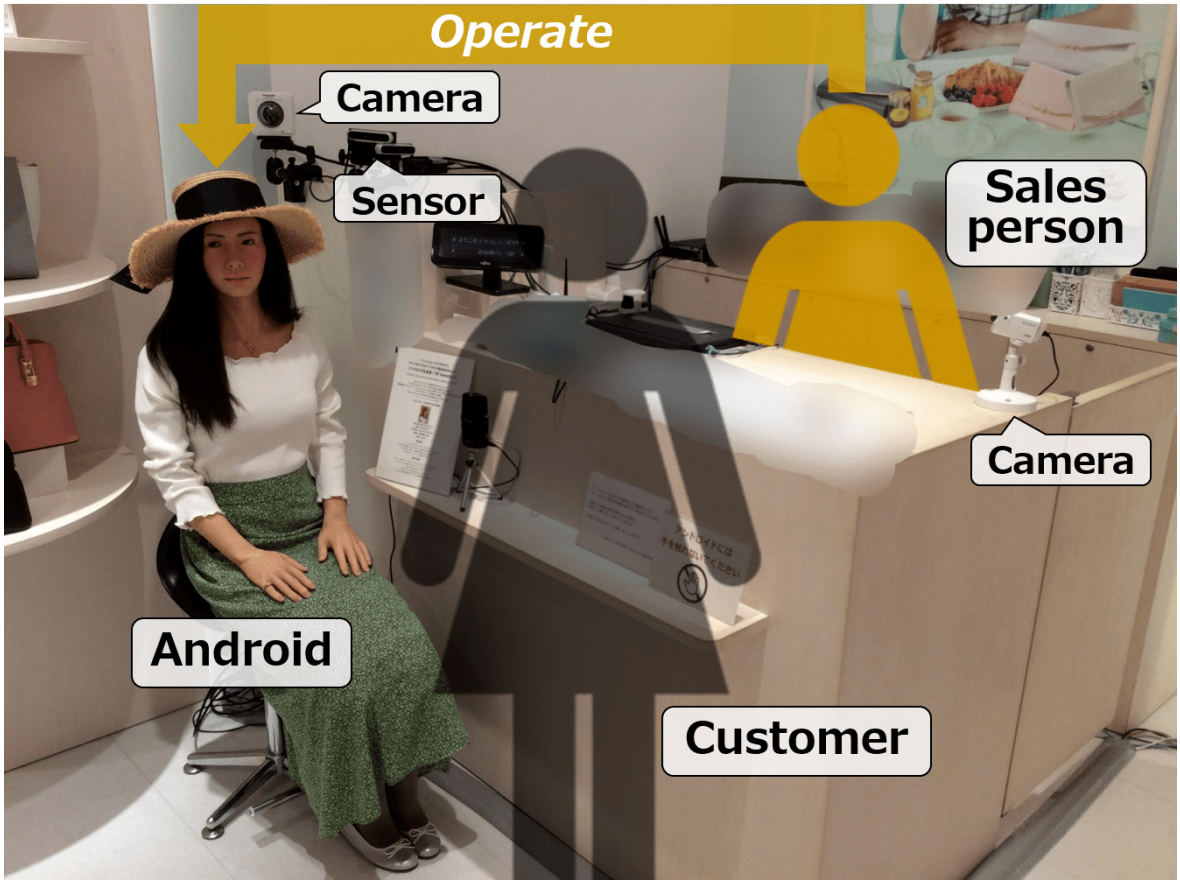



Fig. 5.1 Scene of the field experiment using the para-operatable semi-autonomous android realized in this study at the store. When salespeople operate the android from the side, they can have the android speak.

comfort and encouragement from each other. If robots working in stores are designed so that they can have a relationship with salespeople that goes beyond "working in the same place," they can provide mental support and work support, and such robots will be accepted by store workers and become widespread.

In this study, I design a robot (Fig. 5.1[*3]) with functions that can contribute both to the work and the mental health of salespeople who work with the robot, while avoiding the difficulties of providing services in a real store, where interactive services are important. The purpose of this study is to design, implement, and apply a

[*3] For the figures shown in this chapter, Fig. 5.1, 5.4, 5.5, 5.7 and 5.10 have been mosaicked to protect the information.

para-operatable semi-autonomous android that can cooperate with salespeople in a store, and through this, to verify its feasibility in the real world and the influence on operators through the interaction with the robot. In section 5.2, I describe the robot designed based on the interview results and the robot system implemented based on the design. In section 5.3, I describe a field experiment in which the implemented robot system was operated in a boutique. The results of the experiment were described in section 5.4 and discussed in section 5.5. In section 5.6, I conclude this study.

## 5.2 Para-operated android robot system

### 5.2.1 Design

#### 5.2.1.1 Interview with salespersons

In this study, I first conducted face-to-face interviews with five female salespeople working at a boutique on "What functions would you like a robot to have if it worked with you in the store?" From the opinions obtained in the interviews, the following three functions were extracted as the functions that the salespeople would like in a robot working with them.

Function 1 Work support to provide explanations to customers in place of salespeople

Function 2 Mental support for salespeople

Function 3 Support for building a relationship between salespeople and customers

For the first function, the robot provides necessary explanations to the customer instead of the salesperson. For example, the salesperson must explain the product warranty to the customer at each checkout. A robot can take the place of the salesperson to provide such explanations, reducing the salesperson's workload and improving work efficiency.

The second function is not directly related to business but is required for the salespeople to continue working while maintaining their mental health. According to the salespeople, they wanted a robot that could ease their loneliness when they were alone in the store, or cheer them up when they felt depressed due to failure in their work. As mentioned in the previous section, it is important to consider what the

robot can do for the salesperson outside of work, in addition to the efficiency of work, when considering a robot that works in cooperation with the salesperson. Therefore, in this study, it was decided to equip the robot with a function to encourage the salesperson in order to provide mental support to the salesperson.

Regarding the third point, the salespeople say that in stores such as clothing stores, the salespeople frequently consult with the customers and suggest products suitable for them, and conversation with the customers is essential to provide good service. However, many customers are reluctant to talk to salespeople because they feel nervous about them. Therefore, I decided to install a function that allows a robot to chat with a customer and a salesperson in order to relieve the customer's tension and make it easier for the salesperson and the customer to talk. It is expected that the robot, the customer, and the salesperson will be able to smile while talking, which will activate the consultation about the product between the customer and the salesperson.

In this study, I adopted a system in which the salespeople operate the robot on-site without being noticed by the customers. Specifically, buttons to control the robot were installed in places where customers cannot see them, and the salespeople can make the robot speak by pressing the button associated with the content they want the robot to say. Although it is not easy to realize a fully autonomous dialog robot that can work robustly in the real world due to difficulties in speech recognition and discourse generation, this system avoids such difficulties and enables natural dialog between customers and the robot in a real store, and the salesperson can let the robot say what they want it to say whenever they want. In the interviews, the salespeople agreed that it would be possible to operate the robot by themselves in the store.

In this study, robots that can be operated by a person in the same space are called para-operated robots. In particular, the fact that the operator is operating the robot is not revealed to the conversation partner of the robot. In conventional tele-operated robots, the operator is located in a remote place far away from the robot, and it is rarely seen that the operator exists in the same space with the robot. In the past, Shimaya et al. used such a form of robot, but the operator revealed to the interaction partner that they were operating the robot [15]. The robot implemented in this study has a unique feature that the operator in the same space with the robot can operate it without being noticed by the robot's interactive partner.

In para-operated robots, for example, the usage image of the Function 1 (work

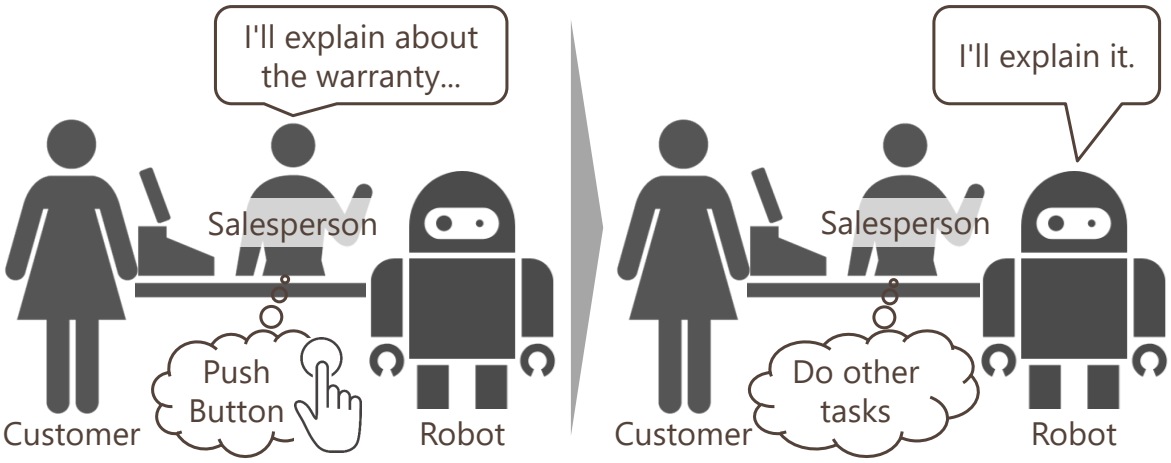



Fig. 5.2 Usage image of the Function 1 (Work support to provide explanations to customers in place of the salesperson).

support to explain to a customer instead of a salesperson) is shown in Fig. 5.2. In other words, the salesperson operates a button from the side when necessary, and the robot speaks. Additionally, for function 2 (mental support for the salesperson), where the robot encourages the salesperson, the salesperson pushes the button to execute the function, and the robot says encouraging words to the salesperson.

This study used Android U (Fig. 5.1), a female-type android robot with a human-like appearance, in consideration of the features of the store. Since the store sells feminine clothing and sundries, it is expected to show customers how to use the Android U by matching it with products and to reduce the loneliness of salespeople when they are alone in the store by its human-like presence. Android U has 19 joints in its head and upper body and can express human-like motions and expressions to the observer. Additionally, pneumatic actuators are used to prevent noise from interfering with interaction.

#### 5.2.1.2 Expected contribution to salespeople

From the above, the following three contributions to the salesperson can be expected by using a para-operated android implemented in this study.

Workload reduction The android can reduce the workload of the salesperson by providing explanations instead of the salesperson.

Mental support The android's presence and encouraging words can reduce the sales-




Fig. 5.3 Button used to operate android robot (nanoPAD2 made by KORG). Long side 325mm, short side 83mm, thickness 16mm.

person's mental burden.

Relationship building The android can assist in building relationships between customers and salespeople through three-party conversations.

This study aims to implement a para-operated android that can be operated by salespeople themselves without being noticed by customers and to clarify its effectiveness through a field experiment in a boutique, in order to realize an android that can contribute to salespeople both professionally and mentally. In the following sections, I report on the implementation of the para-operated android with the functions described in this section and the results of the field experiment using the android in a store.

### 5.2.2 Implementation

In this section, I describe the implementation of a para-operated android system. With this system, salespeople can select one of the pre-registered speech contents to make the android speak by pressing a button placed in a place hidden from customers.

### 5.2.3 Operation Method

In this system, multiple buttons are used to operate the android. Each button is pre-registered with the speech content of the android, and the android can be operated by pressing the button associated with the speech content according to the situation.

Table 5.1 Scenario explaining product warranty.

| Speaker | Speech content |
|---|---|
| Salesperson | I'll explain about the warranty... |
| Android | Well, shall I explain that? |
| Salesperson | Okay, please. |
| Android | Please leave it to me. The warranty is [...]. How about it? |
| Customer | Yes, I see. |
| Android | Thank you. |

In the field experiments described below, a KORG nanoPAD2[*4] (Fig. 5.3) was used as a button to control the android. The nanoPAD2 can be connected to a computer via USB, and button inputs can be captured as MIDI signals. This system does not depend on nanoPAD2 and is implemented so that other types of buttons can be used. This feature is to be able to add or change buttons as needed according to the needs of users of this system.

#### 5.2.3.1 Speech scenario

In this system, each of the 16 push buttons (trigger pads) on the nanoPAD2 can be used to register a continuous scenario or a single word of speech. For instance, in the one-sequence scenario, we can register the speech of an android, which is assumed to be a conversation as shown in Table 5.1. Taking Table 5.1 as an example, pressing the button registered with this scenario will execute the first android's speech ("Um, the explanation..."), and pressing the same button will execute the second android's speech ("Leave it to me..."). In this way, when registering a scenario to a button, users can proceed with the scenario by pressing the button continuously. For example, if a button is associated with the word "Welcome", pressing the button once will make the android say the word. Additionally, it is possible to define behaviors that are expressed at the same time as the android's speech, such as changing facial expression or gaze direction during speech. The registration of speech contents to the buttons can be defined in a JSON format scenario file, which can be easily added or changed. In the field experiment described below, I used two nanoPAD2s and registered a total

[*4] `https://www.korg.com/jp/products/computergear/nanopad2/`

of 32 different speech patterns to each button.

Here, as mentioned above, the system allows the androids to speak only what they have registered in advance. However, there may be cases where customers do not respond according to the scenario assumed in Table 5.1, or where the androids are freely spoken to by customers in the store. To prepare for such cases, I have registered such utterances as "I'm sorry, I don't understand" or "Please help me, staff" in advance, so that even if the android cannot respond appropriately, the salesperson can naturally follow up by letting the android make such utterances.

Each button can be used not only for speech but also for android behavior and to reset the scenario. By registering actions to a button, for example, you can make the android follow the face of the nearest person or make the android smile. The reset function can also be used to terminate a scenario that is in progress. The specific scenario of what the android would say was decided in consultation with the salespeople before the field experiment described in the next section. Through the consultations, we created scenarios in which the android explained what the salespeople had to explain every checkout (to reduce the salespeople' workload), introduced itself and the salespeople (to help build a relationship between the salespeople and the customers), and so on. Additionally, we prepared buttons for the android to say "Welcome," "Thank you," "Staff, please," and so on, which are expected to be spoken frequently.

I also prepared a button to make the android speak words of encouragement to the salesperson (to reduce the mental burden of the salesperson). When the button is pressed, the android randomly speaks one of the 16 types of encouraging words to the salesperson (e.g. "You always work hard. Thank you for your hard work," "Don't overwork yourself", etc). The number of encouraging words was increased to 23 during the field experiment, based on requests from the salespeople.

Table 5.2: Scenarios for the android used in the field experiment.

| Scenario topic | Content |
|---|---|
| Product warranty description | Description of the product's warranty that is given after purchase. |
| Introduction to the app | Description of the store's app. Used in two ways depending on whether there was a campaign or not. |

Continued on next page

**Table 5.2 – continued from previous page**

| Scenario topic | Content |
|---|---|
| Introduction of the point card | Description of the point card issued by the store. |
| Introduction of the facility's point card | Description of the point card issued by the facility where the store is located. |
| Introduction of the credit card of the facility | Description of the credit card issued by the facility where the store is located. |
| Explanation of how to care for the product | Description of how to take care of the product after purchase. |
| Product recommendations | Scenarios for recommending products. Four different scenarios were used depending on the product. |
| Introduction of campaigns | Description of campaigns being conducted at the store. Used in two ways depending on the contents of the campaign. |
| Introduction of different brand stores | Introduction of another brand store in the same facility. |
| Self-introduction of the android | A scenario in which the android introduces itself to customers. |
| Introduction of the laboratory | A scenario in which the android introduces the laboratory to which it belongs. |
| Introduction of the salespeople | A scenario in which the android introduces a salesperson to customers. Used in four different ways depending on salespersons. |
| Introduction of the song | A scenario in which the android hums a song related to the store and introduces the song. |
| Chat for regular customers | A chatting scenario that includes content for regular customers of the store. |
| Single-word | There were 12 ways to say "Welcome", "Thank you", etc. |
| Behavior only | The behavior of looking at the nearest person from the android and salesperson. |

The above-mentioned pre-designed scenarios were partially modified and added by me in the field, based on the suggestions from the salespeople during the field experiment described below. Throughout the experiment, a total of 36 different speech scenarios were used, except for the pattern of the android's words for encouraging the salesperson. The contents of these speech scenarios are shown in Table 5.2. During the experiment, all 32 buttons described in the previous section were used with scenarios, one-word speeches, and action instructions, choosing which scenario to reflect in the system according to whether there was a campaign going on in the store.

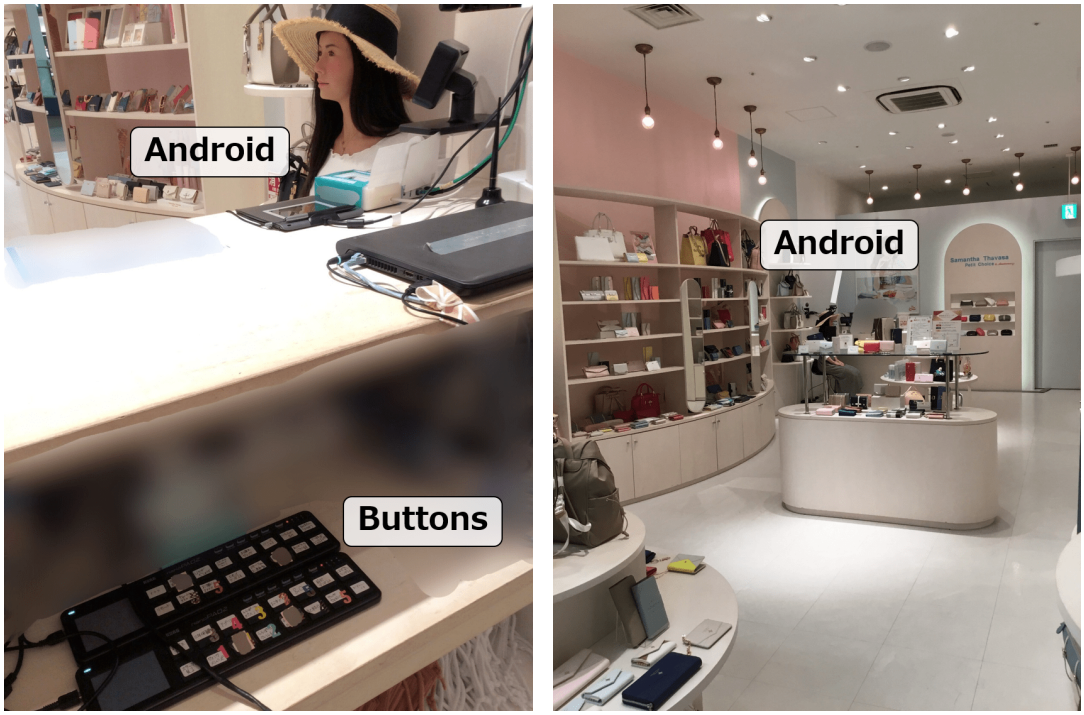
#### 5.2.3.2 Behavior generation

For the voice of the android, I used the voice of the real-time speech synthesis function of AITalk SDK[*5]. By not using prerecorded speech, the system can easily deal with changes in speech content in the field.

For generating android behavior, the software "UltraHuman" developed by Sakai et al. was used to generate natural lip-sync, head, and body movements of the android based on its speech [94]. By using this software, in addition to the above behavior, unconscious behavior such as blinking and subtle body movements of the android are automatically generated, and the facial expressions and gaze of the android can be changed.

Additionally, by using the three-dimensional positional coordinates of the face of the person closest to the android, the android is able to follow the face of the person. In the experiment described below, the android was placed in a location where customers were expected to come closest to the android (in front of the cash register). This allowed the android to follow the faces of customers when customers were in front of the cash register and when the android performed the function of following the face of the person closest to it. The three-dimensional position coordinates of the face of the person closest to the android were obtained from a system using RealSense D415 from Intel[*6] as a sensor. In the field experiment described in the next section, the sensor was installed at the position shown in Fig. 5.1. The timing for the android to follow the human face can be defined in the scenario described above.

[*5] `https://www.ai-j.jp/products/sdk/`
[*6] `https://www.intelrealsense.com/depth-camera-d415/`




(a) View from inside the cash register. The buttons for operating the android are placed out of sight of customers. Each button has a sticker with the name of the scenario or function on it.

(b) View from the store entrance.

Fig. 5.4 Android system installed in the store during the field experiment.

## 5.3 Field experiment

### 5.3.1 Method

In this section, I describe a field experiment in which the android system described in the previous section was installed and operated in a boutique. This experiment was conducted with the cooperation of "Samantha Thavasa Petit Choice & Anniversary

Lalaport EXPOCITY"[*7]. The experiment was conducted for 11 days in June 2019, from 10:00 a.m. (opening time) to 9:00 p.m. (closing time)[*8], with the android installed next to the cash register in the store. Figures 5.1 and 5.4 show the android installed in the store. The buttons for operating the android were placed on the shelf of the cash register so that customers could not see the salesperson operating the buttons (Fig. 5.4(a)). To make it easier to understand what function each button performs, stickers with the name of the scenario or function were attached to each button.

Five female salespeople worked in shifts at the store, and I let them freely operate and use the android system. Before the experiment, I gave the salespeople a lecture on how to use the android system and asked them to practice operating it. After the experiment started, it was found that the salespeople often left the checkout area except for the checkout, and were not able to operate the android. Therefore, I implemented a system to remotely control the android from the backyard of the store. This system is a GUI on a computer that replicates the system used by the salespeople to control the android with buttons, allowing the same operation as the salespeople to be performed from the computer. The system also has a function to make the android speak the contents of the keyboard input. When the salespeople were too far away from the cash register to operate the androids, the experimenters used this system to operate the androids. In particular, the function to help build a relationship between salespeople and customers was not operated by the salespeople themselves during the experiment, but by the experimenters. However, for the function of having the android provide necessary explanations on behalf of the salesperson, the salespeople themselves operated the android during the checkout process, and the experimenters did not operate it remotely.

In this experiment, it was not realized that the function of helping to build a relationship between the salesperson and the customer was operated by the salesperson himself, but it may be possible to realize it in the future by improving the system. In this experiment, the experimenters remotely operated this function and investigated how it would be accepted in a real store so that I could consider the possibility that this function could contribute to the salespeople. Therefore, in this experiment, the experimenters executed this function by remote control.

[*7] `https://mitsui-shopping-park.com/lalaport/expocity/shopguide/703192.html`
[*8] The detailed dates of the experiment are June 3, 4, 6-9, 11 and 13-16, 2019.

This experiment was conducted under the approved guidelines of the ethics committee of the Graduate School of Engineering Science, Osaka University. In addition, this experiment was conducted following the ethical regulations of the company that runs the store where the experiment was conducted and the company that runs the facility where the store is located, and with the approval of the responsible person of each company.

### 5.3.2 Evaluation Method

The evaluation of the android system realized in this study was conducted by analyzing interviews with salespeople, video recordings of customers and salespeople interacting with the androids, records of the authors' observations, and log data of the android system obtained during the experiment. The video recording was done using two cameras as shown in Fig. 5.1. For observation, the experimenters refrained from standing in the store and observed from the backyard of the store through the real-time images of the two cameras. Also, questionnaires and interviews were conducted with customers only when their consent was obtained. In the questionnaire, customers were asked to answer the following three questions on a 9-point Likert scale (1: "Not at all" to 5: "Neither" to 9: "Very much agree").

Q1 I felt that conversations with the android were natural.
Q2 I felt that the android was a nice robot.
Q3 I would like to meet the android again.

In this experiment, salespeople were asked to operate the android using buttons, so it was important to investigate whether the operation was burdened for them. At the end of the field experiment, the salespeople were asked to answer two questions: "The operation of the android was difficult" and "The operation of the android was troublesome" on a 9-point Likert scale (1: "Not at all" to 5: "Neither" to 9: "Very much agree"). As for the video recording of the interaction of the androids with customers and salespeople in the store, the store requested not to record the scene in the store at all times. Therefore, the video of the android during checkout was always recorded, while other scenes were recorded only when human interaction with the android occurred.

During the experiment, the experimenters observed and recorded the salespeople using the android from the backyard of the store. At the end of the experiment,

an annotator watched the recorded videos and analyzed the number of people who interacted with the android and whether the customers smiled while interacting with the android. However, due to a problem with the video recording equipment, it was not possible to record video on one of the days during the experiment. Therefore, the results obtained from the video analysis described below are based on the analysis of only 10 days of video.

## 5.4 Result

This section describes the results of the field experiment described in the previous section. The following figures show the scene during the field experiment, but the actual interactions between the customers and the android cannot be shown here because I did not get permission to show the photos of the customers who visited the store.

As a result of video analysis, the number of customers who interacted with the android during the experiment was 312, which can be determined from the recorded video. Here, "customers who interacted with the android" refers to those who exchanged at least one word with the android. As for the number of customers who did not interact with the android, I could not record all the scenes where there was no interaction between the android and customers because I had to refrain from recording the entire store during the experiment. However, the number of customers who did not interact with the android was 395, as far as I could determine from the video analysis.

### 5.4.1 Workload reduction

I describe the results of verifying the effectiveness of the workload reduction function of the salesperson in the android system implemented in this study. Figure 5.5 shows the interaction between the salesperson, the customer, and the android when the salesperson uses the android's substitute explanation function during checkout. In Fig. 5.5, a salesperson played the role of a customer, not during the actual checkout process, but during the practice of using the android system.

As a result of video analysis, it was found that there were 55 times during the entire experiment when the android was able to explain on behalf of the salesperson during checkout, and 50 times when the salesperson operated the android and

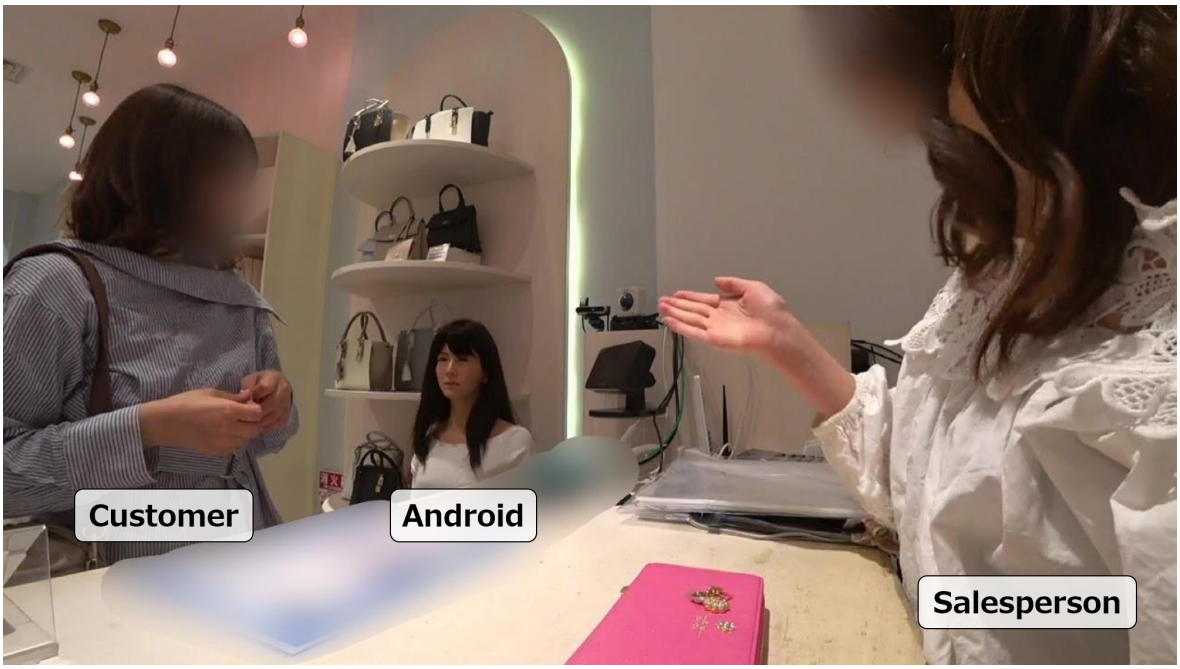



Fig. 5.5 Interaction between a salesperson, a customer, and an android during accounting (image). This picture shows a scene of practice for using the function the android explains instead, so the salesperson is playing a role of a customer.

used the substitute explanation function. Figure 5.6 shows the number of times this function was used each day. The usage rate for each day is shown above each bar in Fig. 5.6, and the overall usage rate of this function during the experiment was 90.9%. In the video analysis, "the scene in which the android was able to explain on behalf of the salesperson during the checkout" was defined as "the scene in which the android actually explained on behalf of the salesperson" or "the scene in which the salesperson explained what the android could explain." Conversely, when neither the android nor the salesperson explained anything, it was not included in the "the scene in which the android was able to explain on behalf of the salesperson during the checkout". This is because the explanations needed vary from customer to customer, and some customers may not need any explanation built into the substitute explanation function. Therefore, such cases are not included in the scene. In the five cases where the salesperson did not use this function despite the android could do so, there were four cases where the salesperson made a quick explanation on her own, and one case where the salesperson forgot about the function of the android.

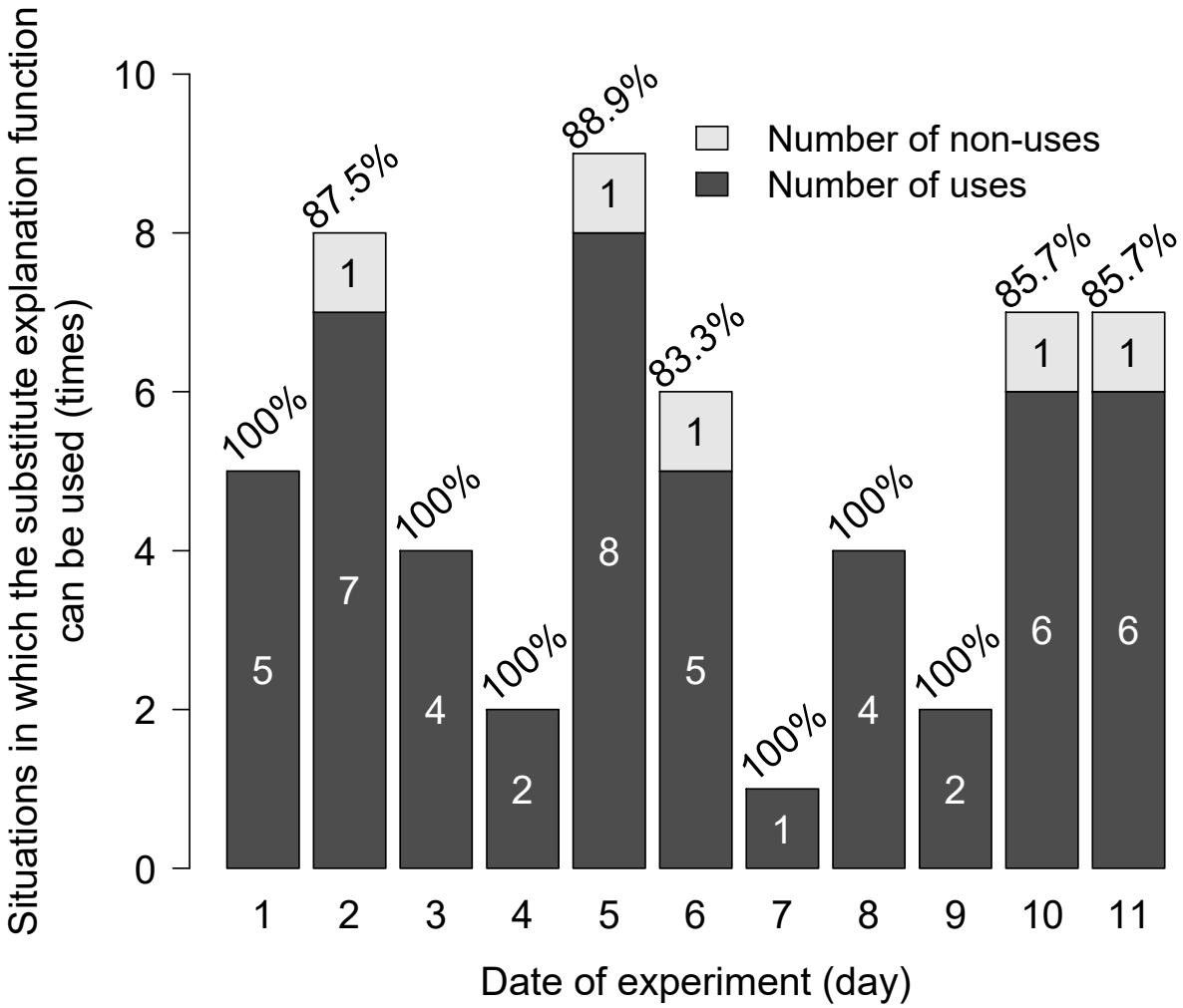



Fig. 5.6 The number of use times (black bars) and non-use times (gray bars) of the proxy explanation function that the android explains on behalf of salespersons. The value in the black/gray bar is the number of use/non-use times per day. The value above each bar is the usage rate of this function per day.

Furthermore, observation and video analysis showed that while the android was explaining on behalf of salespeople, the salespeople were always able to proceed with other tasks such as packaging or paperwork. Interviews with the salespeople revealed that they were able to use their time efficiently because they were able to do other things while the androids were explaining to customers and that the android was able to convey information without mistakes. A salesperson who had just been assigned to the store where the experiment was conducted commented, “I have just been assigned




Fig. 5.7 Scene that salespersons are using the function to have the android say words of encouragement.

to the store and have not yet fully mastered the workflow. I am glad that the android gave me time to think about what to do next. I think I will be able to memorize the explanations by listening to the android's explanations."

### 5.4.2 Mental support

Next, I describe the results of the contribution of the android system in reducing the mental burden of the salespeople (i.e., the presence of the android and its encouraging words may reduce the mental burden of the salespeople). First, from the observations and the video recorded during the experiment, it was observed that the salespeople smiled and said "Let's do our best" when they used the android's encouraging speech function. Figure 5.7 shows a scene of the salespeople using this function. From the interviews with the salespeople, they said that the function of having the android speaks encouraging words to the salespeople made them "feel happy and lighthearted (when they are encouraged by the android)", "I felt as if I was really encouraged even though I was operating the android", and "When I

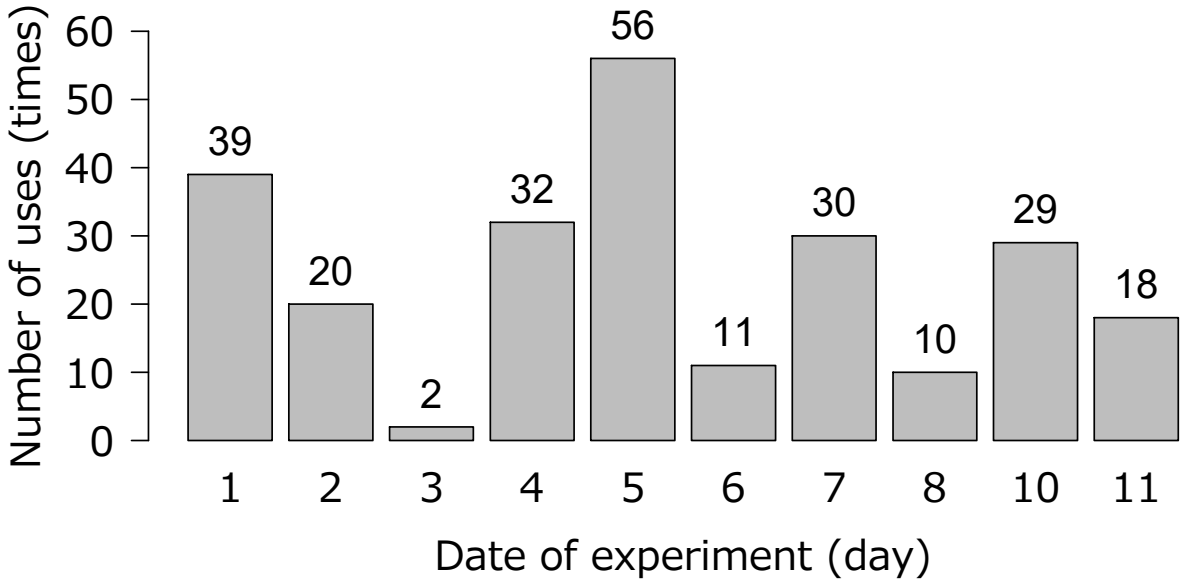



Fig. 5.8 The number of daily use of the function that the android utters encouragement words for salespersons. The value above each bar is the number of uses each day.

want to be encouraged, the android will encourage me until I am satisfied, which is something that is difficult to ask of others." As mentioned in section 5.2.3.1, during the experiment, the salespeople asked for more patterns of the android's words of encouragement, and the experimenter increased the number of patterns from 16 at the beginning of the experiment to 23. As for the presence of the android, during the conversations among the salespeople, some of them said, "When I am alone in the store, I really want U-chan[*9] to be there," and during the interviews, some of them said, "U-chan is always smiling, so I feel relieved when I see U-chan, and I also feel like smiling and trying my best."

Figure 5.8 shows the daily frequency of use of the android's encouraging speech function by the salespeople during the experiment. The average number of times per day was 24.7, with a standard deviation of 15.07. However, this number includes several times when the salesperson used the android for purposes other than getting encouragement from the android. For example, from observations during the experiment, there were cases where the salesperson used this function to practice operating

[*9] U-chan is the name of the android used in the field experiment.

the android. In this experiment, I was not able to accurately record the number of times the android was used for other purposes. However, from the observations during the experiment, the number of times it was used for other purposes was small, so the data shown in Fig. 5.8 can be used as a reference for the number of usages of this function. On the ninth day of the experiment, a TV station was shooting the android, and this function was sometimes used as a demonstration. Therefore, I excluded the data of the number of usages of this function on the 9th day from the results.

### 5.4.3 Relationship building

I describe the results of the contribution of the android system in this study to assisting the building of relationships between salespeople and customers (i.e., the android can assist the building of relationships between salespeople and customers through three-way conversations with customers and salespeople). As described in section 5.3, there were many cases where the salespeople were away from the cash register and could not operate the android except during the checkout process, so the experimenters remotely operated the android from the backyard to perform the functions related to assisting in building relationships between salespeople and customers.

As for the assistance in building relationships between salespeople and customers, interviews with salespeople revealed that four of the five salespeople gave positive evaluations, such as "it is better to have the android in the store because it makes me feel more comfortable talking to customers and easier to start a conversation with them," "I am able to break the ice with customers much faster with U-chan," "Many customers have stiff facial expressions due to nervousness before and after entering the store, but talking with U-chan makes customers smile" and "The android has become a bridge between customers salespeople." The remaining salesperson did not express any opinions on this perspective. As a result of the video analysis, 83 customers talked with the android and a salesperson, and 80 customers smiled during the conversation, which means that about 96% of customers smiled.

#### 5.4.3.1 Operability of the android system

In the interviews with salespeople, three of five salespeople commented on the operability of the android system, saying that "it was difficult to operate the buttons when the cash register was crowded" and "I could not remember the sequence of the

buttons until the end." After the end of the field experiment, the salespeople were asked to answer two questions about the operation of the android, as described in section 5.3. The results showed that the mean and standard deviation of the evaluation of "the operation of the android was difficult" was 1.6 and 0.89, respectively, and of "the operation of the android was troublesome" was 1.8 and 0.45, respectively.

#### 5.4.3.2 Customers' impressions of the android

Figure 5.9 shows the results of the questionnaire survey on customers' impressions of the android. Questionnaires were answered by 28 customers (21 females and 7 males). The questionnaire was administered only to customers who exchanged at least one word with the android and who consented after the checkout. The mean of the ratings for Q1 "I felt that conversations with the android were natural" was 6.11 with a standard deviation of 1.61. The mean of ratings for Q2 "I felt that the android was a nice robot" was 6.46 with a standard deviation of 1.99. The mean of the ratings for Q3 "I would like to meet the android again" was 6.93 with a standard deviation of 2.05. The Mann-Whitney U-test was conducted for each question with the chance level of 5 (neither), and each p-value was $p < .01$ (Q1: $p = 0.0029$, Q2: $p = 0.0013$, Q3: $p = 6.2 \times 10^{-8}$), indicating that each rating was significantly higher than the chance level.

## 5.5 Discussion

### 5.5.1 Discussion of expected contributions

I describe the contribution of the android system implemented in this study in reducing the workload of salespeople (i.e., the android can reduce the workload of salespeople by providing explanations on their behalf). First, the results in Fig. 5.6 show that the salespeople continuously used the function of letting the android explain on their behalf. It was also confirmed that the salespeople were always working on other tasks while the android was explaining. Moreover, considering the positive comments from the interview with the salespeople, I conclude that the android's substitute explanation function contributed to the salespeople from qualitative evaluation.

Additionally, a salesperson who had just been transferred to the store where the experiment was conducted said that she could remember the content of the explana-

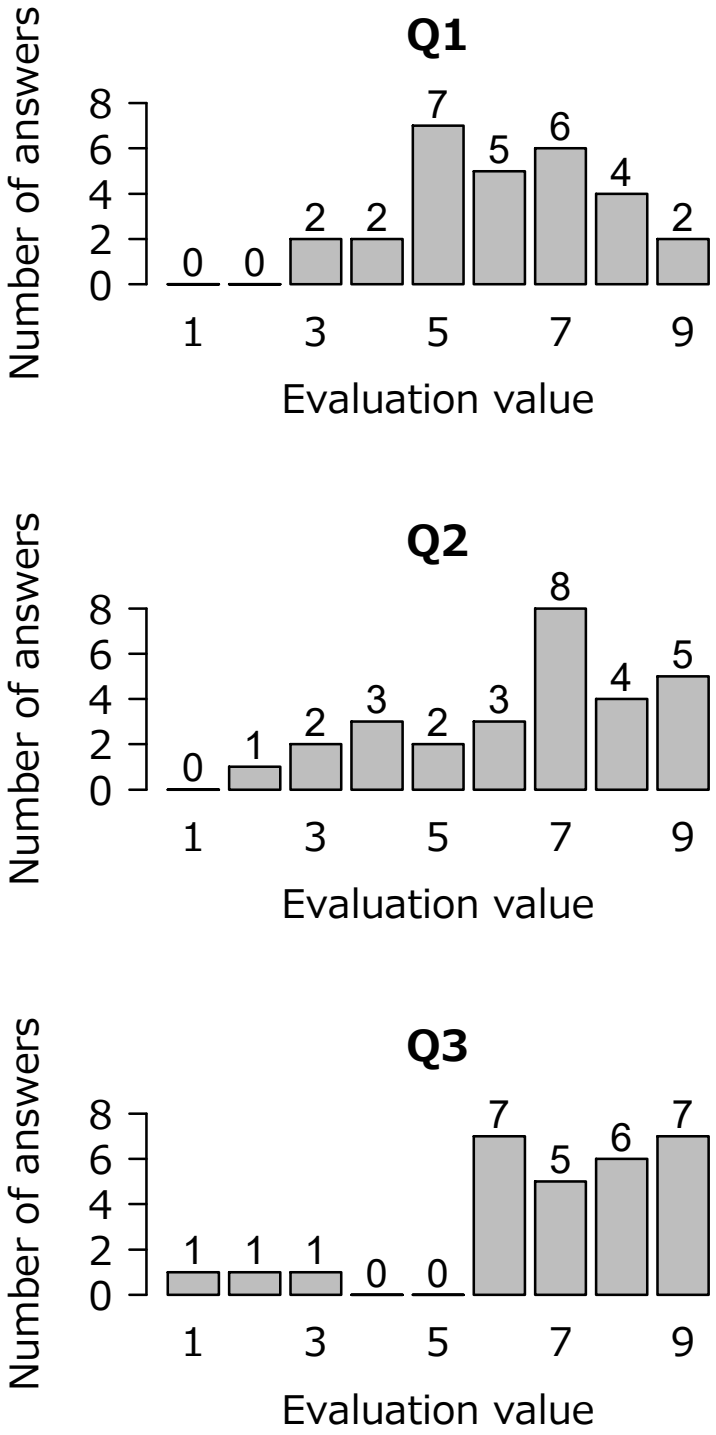



Fig. 5.9 Result of the questionnaire survey on impressions of the android by customers who made payments. The value above each bar is the number of respondents for each evaluation value. From the Mann-Whitney U test with the chance level (5: Neither) for each question, each question was evaluated significantly higher than the chance level.

tion by listening to the android's explanation. This suggests that this system could also be applied to help people learn their jobs.

In the interview with the salesperson, she mentioned the difficulty of operating the buttons on the android. In the system, 32 buttons were used to control the android, and it would have been difficult to remember all the functions of the buttons. Therefore, I think that improving the operation interface is an issue that should be considered in the future in order to create a more user-friendly interface. However, the results of the questionnaire survey on the operation of the android conducted on salespeople in the experiment in this study showed that the operation of the android was not difficult or troublesome. The results of the questionnaire showed that the salespeople did not find the operation of the android system difficult or cumbersome. This suggests that the use of the android system in this experiment was not too much of a burden for the salespeople.

Next, I discuss the contribution of the android system in reducing the mental burden of the salespeople (i.e., the presence of the androids and their words of encouragement can reduce the mental burden of the salespeople). This feature, which allows the salespeople to operate the android and have it speak words of encouragement to them, was favorably accepted by them based on the behavior of them and their requests for more patterns of words of encouragement that I learned from the interviews and observations. Some of the salespeople also said that they felt as if they were really encouraged. From these results, I believe that this function was accepted without any problems. The salespeople who observed the system said that they felt as if they were really encouraged. In particular, some salespeople said that they were able to receive encouragement until they were satisfied when they wanted to be encouraged, suggesting that there was merit in daring to operate an android to speak words of encouragement.

The results of the number of times this function was used during the experiment, obtained by analyzing the log data of the android system, showed that this function was used on average 24.7 times per day, although there were a few times when it was used for other purposes. Considering that the store is open for 11 hours a day, I consider that this function was accepted and used by the salespeople. They also said that they wanted the android to stay with them when they were alone in the store and that they felt relieved when they saw the android smile. Therefore, the presence of the android in the store itself may have contributed to reducing the mental burden

of the salespeople in this experiment. Therefore, I believe that the android system has contributed to the reduction of the mental burden of the salespeople, although only from a qualitative perspective.

I describe the contribution of the android system in this study to assist in building a relationship between salespeople and customers (the android can assist in building a relationship between salespeople and customers by having a three-party conversation with customers and salespeople). From the interviews with salespeople described in the previous section, I found that the presence of the android enhanced the conversation with customers and that the android enabled the salespeople to quickly establish rapport with customers. The results of the video analysis showed that about 96% of the customers who talked with the android, the customer, and the salesperson smiled during the conversation. It is known that smiling increases positive emotions in people [112], and it is likely that smiling eases the tension of customers and makes it easier for them to talk to the salesperson. However, this function was not used by the salespeople themselves, but by remotely operated by the experimenters. This is because the salespeople were away from the cash register except during checkout and could not operate the android. As described above, although the salespeople were not able to operate this function by themselves, the qualitative evaluation indicates that this function contributed to the salespeople. Therefore, there is a possibility that this perspective can be contributed to salespeople in the future even when salespeople can operate this function by themselves. To realize this, it is necessary to improve the operation method and the autonomy of the android, which will be discussed later in this section.

Regarding the customers' evaluation of their impressions of the android after the checkout, the result of Q1, "I felt that the conversation with the androids was natural," indicated that the android was able to communicate with customers naturally. Therefore, I consider that the android realized in this study was able to behave appropriately to customers.

The results of the above impression evaluation questions Q2 "I felt that the android was a desirable robot" and Q3 "I would like to see the android again" indicates that the android was considered favorable by customers. In particular, Q3 was rated highly, suggesting that customers may come back to the store to meet the android. Therefore, the installation of androids with interactive functions in stores may have potential applications to increase the number of customers visiting the stores. How-

ever, these results were obtained only from customers who exchanged at least one word with the android and who agreed to participate in the questionnaire survey after the checkout. Therefore, it is not possible to study the impressions of all customers who interacted with the android. In order to conduct a strict evaluation of the impressions of the android in this study, it is necessary to conduct controlled experiments, which will be the subject of future work.

From the qualitative results, I believe that the android system contributed to the salespeople in terms of reducing their workload, reducing their mental burden, and helping them build relationships with customers. Additionally, the android was able to behave naturally for customers. However, the function of helping to build a relationship between the salesperson and customers was performed by the experimenters by remote operation, and the salespeople could not operate the system by themselves in this experiment. Therefore, although the salespeople were not able to operate the function that assisted in building a relationship between them and customers, I believe that this study has resulted in an android system that is effective for salespeople and can be used effectively in actual stores

In this field experiment, the experimenters remotely operated the android when the salespeople were out of the cash register and could not operate it. This showed that in a store like the one in this experiment if the position of the operation interface of the para-operated android is fixed, there will be situations where no one can operate the android. Therefore, it is recommended to improve the system to be independent of the operator's position by using a small portable interface or a camera that can recognize gestures to operate the android. In the case that the android is chatting with customers, the system could be improved to allow the android to talk with them autonomously. However, due to the difficulties in speech recognition and discourse generation, the android may not be able to respond to customers autonomously and appropriately. Therefore, taking advantage of the characteristics of para-operation, in a system that is not bound by the operator's position as described above, operators may be able to partially assist the android's response by operating it. This would make it easier to realize natural conversations, rather than having the android talk with customers fully autonomously. In the future, I would like to improve the system so that the autonomous function of the android and the operation by an operator can coexist, and clarify its effectiveness.

The number of customers who interacted with the android during this experiment

was 312, and the number of customers who did not interact with the android was 395 as far as I could determine from the video. For the latter number, the exact number could not be determined because it was necessary to refrain from recording the entire store during the experiment. Therefore, it is difficult to discuss how many customers the android was able to interact with in this study.

In this study, I did not consider how many customers the android could interact with but focused on how the android could contribute to the salespeople at appropriate times. Of course, it would be useful for the android to interact with many customers in order to attract them to the store. Therefore, the development of androids that can interact with a large number of customers in a real store and their verification should be a subject to be addressed in the future.

### 5.5.2 Discussion of non-expected contributions

In the following discussion, I will describe other interesting aspects of the android system, based on the results of interviews and observations of the salespeople that were not presented in the previous section.

When the salespeople were asked about their attachment to the android, they replied, "I became attached to U-chan and felt lonely on days when she was not around," "At first I thought U-chan was a robot, but now I feel like she is one of us. I will miss her. I want her to stay," and "I felt energized just by being with her as if she were a staff member." These results suggest that the salespeople became attached to the android after 11 days of working together. In the field experiment, as shown in Fig. 5.10, the salespeople repaired the androids' makeup and styled her hair. Therefore, the salespeople's attachment to the android is not only due to her contribution to reducing the salespeople' workload and mental burden, or to help build relationships between the salespeople and customers, but also because of their overall interaction with the android. Through the field experiment, the android did not become a bother for the salespeople but rather became an accepted companion for them. This suggests that the para-operated android developed in this study may be able to play an active role in stores such as the one where the field experiment was conducted.

I will discuss the possibility that androids could also play the role of companion for customers in a store like the one where I conducted the field experiment. An interview with a salesperson revealed that when two customers, a man, and a woman, came to




Fig. 5.10 Scene that the salesperson are fixing android's make-up.

the store, the salesperson was able to spend more time with the woman because the man kept talking to the android. The store where I conducted the field experiment is a store that sells feminine products. The salespeople said that in such a store, male customers who visited as companions of female customers inevitably became bored in the store and made a salesperson feel as if they wanted to leave the store quickly. Additionally, such an atmosphere makes salespeople impatient, and they sometimes cannot calmly examine the products with female customers. This suggests that in stores where customers' companions may be impatient, androids could contribute to the store by taking care of the companions. In this study, I did not implement the android system with this perspective in mind, but it is a point that should be considered in future applications of androids working in stores.

I discuss the possibility of increasing the amount of time customers spend in a store by having androids talk with customers in relation to having androids act as their companions. Through interviews with salespeople, I found that in stores such as the one where I conducted this field experiment, the presence of customers in the store

itself is important. This is because it is believed that the presence of customers in the store creates an atmosphere that facilitates entry into the store and encourages other customers to enter. Three salespeople pointed out that the presence of androids in the store may have increased the time spent by customers in this field experiment. In this study, I cannot quantitatively show that the android system actually increased the time spent by customers. However, the idea of having conversations with customers in order to increase their time spent in the store is an important point to consider when applying androids to stores such as the one where this field experiment was conducted.

### 5.5.3 Limitations

This study includes interviews with salespeople who cooperated with the experiment and worked with the android for 11 days. In this field experiment, the authors consulted with the salespeople and exchanged opinions on the android system even before the experiment was conducted. Through such discussions with the salespeople, I think I was able to make them aware of the importance of discussing the system honestly in order to create a better system for them. Therefore, I believe that the results of the interviews are reliable because the salespeople did not give unreasonably high or low evaluations to the system implemented in this study.

The effects of installing the android in this experiment are not quantitatively shown, as I have not compared the results with those without the android. In order to clarify the value of the android system in this study, it would be important to compare it with agents other than androids, or with non-agent systems that simply play explanatory videos. Although it was difficult to conduct multiple conditions in the experiment of this study, I would like to clarify the effectiveness of the system proposed in this study by conducting experiments that enable comparison with the above multiple conditions in the future.

In this field experiment, I used an android, but the significance of using it compared to virtual agents or humanoid robots cannot be clarified from the results of this study. In the interviews with the salespeople, they said that their loneliness was relieved by the presence of the android when they were alone in the store, suggesting that the presence of the android has created a sense of company for the salespeople. Additionally, the salespeople's comment that seeing the android's smiling face motivated them to work harder suggests that the android's ability to express human-like

expressions may have been effective. These are some effects of using the android in this study, but clarifying the differences from using other agents is a subject for future research.

### 5.5.4 Further possibilities for para-operated robot

As a prospect of para-operation style, I aim to realize a "buddy robot," a robot that can have a buddy-like relationship with a human. A buddy-like relationship means a relationship in which both can trust each other and work closely together, and relate to each other with a high level of trust and intimacy outside of work. A buddy robot is a semi-autonomous robot that has both the ability to be operated by an operator and the autonomy to interact with humans. One of the situations in which I envision the use of a buddy robot is when the robot, its operator, and the person interacting with it are in the same space, as described in this study. For example, as in the case of the android used in this study, the operator can operate the buddy robot without being noticed by the interlocutor and has the robot assist them in conversation by saying things that they are not comfortable saying. On the other hand, as mentioned by the salesperson who used the android robot in the experiment, it is difficult to operate the robot all the time, so the robot can speak autonomously and can interact with the operator autonomously in some situations. In other words, for the operator, the robot is a "tool" like their own avatar, which can be operated to speak what they want it to say, while being an "other" that can interact with humans to a certain degree of autonomy.

I consider that a robot that can be used like a tool that can be directly operated, while being like an autonomous other to operators, is an effective tool for them. For example, in a relationship between humans, there are situations where you can give instructions to an autonomous human partner and have them do what you want them to do. Direct operation with a buddy robot corresponds to this kind of instruction between humans. In particular, there is no guarantee that the human other will act as instructed, but there is always a guarantee that the robot will act as operated. In other words, in the case of humans, it is necessary to first build up mutual trust, and then trust the other person and give instructions to work together. However, with buddy robots, this process can be skipped and people and robots can work together through a direct operation. In this way, operators of buddy robots can give instructions to robots from the beginning without worrying, and work together with

robots in close cooperation to perform tasks as they wish. Additionally, I expect that through such collaboration, operators will gain a high level of trust and intimacy with the robots outside the work, and the operators and robots may build a buddy-like relationship. In this way, operators can trust and accept the robots' words beyond the work, and they may be able to consult with the robots about difficult things for them to disclose to others.

I think that a buddy robot that is both operatable and autonomous is a robot design that can be accepted by humans as a partner with whom they can work efficiently, as well as a partner they can trust in other situations. In order to realize such a buddy robot, this study shows that it is possible for a salesperson to work with the robot in an actual store and that the robot can contribute to the salesperson both in work and mental aspects. In the future, I will focus on robots that have both operatability and autonomy, and investigate whether the operator's impression of the robot changes depending on the method of operation, and what kind of existence the operator accepts through interaction with the robot when the autonomy of the robot is improved. I would like to confirm the validity of the design of buddy robots.

## 5.6 Conclusion

In this study, I conducted a field experiment to verify the real-world feasibility of the para-operation form in cooperation between a robot and its operator in a para-operated semi-autonomous robot, as well as the influence on the operator through interaction with the robot as the operation target. In the experiment, I selected a boutique as a place where robots are actually applied and implemented a para-operatable semi-autonomous android robot system that allows salespeople to operate it without being noticed by customers. The android can help the duties of the salespersons at a boutique and provide mental support by cooperating with them. Through an 11-day field experiment in a boutique, in which the salespersons operated the android, it was observed that the android could contribute to the salespersons in three points: 1) reducing the workload of salesperson, 2) providing mental support for a salesperson, and 3) assisting salesperson to establish a good relationship with the customer. In particular, in the mental support feature, there was a function where the salespeople could operate the android themselves and have it say words of encouragement to them. The log data show that the salespeople used this function enough times,

and the results of the interviews suggest that the salespeople accepted the encouraging words of the android even though they were operating it themselves. These results demonstrate the real-world applicability of the para-operation form through this study. Furthermore, it has been shown that operators of a para-operatable semi-autonomous android can accept the dialog with it, and can relate to it they are operating as if the android were another person.

# Chapter 6

# Conclusion

## 6.1 Summary

The purpose of this research is, by classifying cooperation between semi-autonomous robots and their operators into two forms: tele-operation and para-operation, to show that each form of operation has a different effect on the operator, and to discuss the further development of semi-autonomous robots. In this section, I summarize three studies conducted to achieve the aim and conclude this research.

In Study 1, I demonstrated through a field experiment a method for building the autonomous dialog function that is essential for semi-autonomous robots. The proposed method is that a tele-operator first operates a robot in the field where it is actually applied, and then the autonomous dialog function is gradually improved using the operation logs. I selected the Internet Live Service (ILS) as a place to apply an android robot and conducted a field experiment to improve the autonomous dialog function of the android that interacts with viewers. In this experiment, first, the tele-operator operated the responses of the android talking with viewers to collect a dialog corpus of comments and android's responses. After the corpus was acquired to a certain extent, the operating system was changed to a semi-autonomous form, i.e., the tele-operator and the developed autonomous dialog system cooperated in replying to users' comments, and the corpus was further collected. This resulted in the acquisition of the corpus containing 4,460 pairs through a total of 20 hours of live broadcasting in ILS. After that, to evaluate the performance of the autonomous dialog system, I recorded the conversations between users and the android equipped with the autonomous system. These conversations were then presented to evaluators

and the result indicated that the android's responses were natural. From this, the dialog system implemented using the proposed method has acquired a certain degree of autonomy. From the result, this study demonstrated how to build an autonomous dialog function, which is a basic function of semi-autonomous robots.

In Study 2, focusing on cooperation between tele-operatable semi-autonomous robots and their operators, the influence on the operator from the autonomous behavior of the robot was clarified by a laboratory experiment. Previous studies reported that when humans play a particular role, their attitudes align with that role. The tele-operators of tele-operatable semi-autonomous robots also play the role given to the robots and may assimilate their autonomous expression. I hypothesized that the attitude of tele-operators may align with that of robots through tele-operation. To verify this, I conducted an experiment with conditions under which a participant operated a part of the body of an android robot that autonomously expressed a preferential attitude toward a painting and a condition under which they did not. Experimental results demonstrated that the preferential attitude of participants who tele-operated the android aligned more with that of the robot in comparison to those who did not tele-operate it, thereby supporting the hypothesis. Through this study, I have revealed the influence that an operator receives from the autonomous behavior of the robot in the tele-operation form in their cooperation.

In Study 3, focusing on cooperation between para-operatable semi-autonomous robots and their operators, I conducted a field experiment to verify the real-world feasibility of para-operation form and influences on the operator through interaction with the robot to be operated by themselves. In the experiment, I selected a boutique as a place where an android robot was actually applied and implemented a system that allows salespeople to operate the android without being noticed by customers. I realized a para-operatable semi-autonomous android robot that can help the duties of the salespersons at a boutique and provide mental support by cooperating with them. Through an 11-day field experiment in a boutique, in which the salespersons operated the android, it was observed that the android could contribute to the salespersons in three points: 1) reducing the workload of salesperson, 2) providing mental support for a salesperson, and 3) assisting salesperson to establish a good relationship with the customer. In particular, in the mental support section, there was a function where the salespeople could operate the android by themselves and have the android say words of encouragement to them. The log data show that the

salespeople used this function enough times, and the results of the interviews suggest that they accepted the encouraging words of the android even though they were operating it by themselves. These results demonstrated the real-world applicability of the para-operation form through this study. Furthermore, it has been shown that an operator of a para-operatable semi-autonomous android can accept the contents of a dialog with the android to be operated, and can relate to the robot they operate as if it were another person.

Through these studies, by organizing the cooperation between semi-autonomous robots and their operators into tele- and para-operation, I have shown that each operation form induces different influences on operators. In the case of tele-operation, a robot and its operator behave as one, and the operator changes their attitude to accept what the robot says. While in the case of para-operation, the operator treats the robot as if it were another person, and receives the effects of interacting with the robot while operating it by themselves.

This research has dealt with two forms of operation, tele-operation and para-operation, and I will further consider the differences between the two. The difference between the two is that in tele-operation, the operator can act as if s/he exists in a remote location, while in para-operation, the operator can participate in the dialog between the robot and its interlocutor in the location. In the case of para-operation, an operator can feel the presence of the robot being operated more, suggesting that the robot can be treated like another person. In applications, it is thought that different effects can be obtained by using both forms of operation according to the purpose of the operator (worker).

In conventional Human-robot interaction (HRI) research, the focus has mainly been on the relationship between the robot and its interlocutor. In this research, I focused on the cooperation between a tele/para-operatable semi-autonomous robot and its operator and showed the effects through operation and the possibility of application of the para-operation form as the operator can become an interlocutor of the robot to be operated by themselves. It is important to realize robots in which the autonomy of the robot and human operation cooperate, as proposed in cybernetic avatar [24]. Therefore, I consider that the viewpoint that focuses on the relationship between the operatable robot and its operator will become more important, and this research will be the edge of HRI with such a viewpoint.

## 6.2 Future directions

In this section, I describe the future directions of this research. In this research, I focused on semi-autonomous robots and their operators and dealt with the cooperation between them in both tele- and para-operation forms. I discuss possible future topics related to the cooperation between semi-autonomous robots and their operators.

### 6.2.1 Semi-autonomous robot use in the real world: integration of tele- and para-operation

The realization of a robot capable of providing dialog services in the real world requires a framework that supports both the robot and the worker who acts as its operator. The findings of this research will lead to the realization of a framework for semi-autonomous robot operation that combines both tele-operation and para-operation, which is thought to be capable of providing stable dialog services in the real world.

Specifically, a robot installed in a store that provides a dialog service, for example, is operated by three elements: para-operation by the store clerks, tele-operation from outside the store by professional staff, and the autonomy of the robot itself. First, the clerks work together with the robots by operating them in para-operation as seen in Study 3. This is expected to reduce the work and mental burden of the clerks. However, it is difficult to control all the necessary behaviors of the robot by para-operation. Thus, if the robot cannot be handled by para-operation, a tele-operator will operate the robot to continue the dialog service. This division of labor between para- and tele-operation allows the tele-operator to be in charge of multiple robots as needed, instead of always being in charge of one robot. In addition, as seen in Study 2, by having the robot autonomously adopt a positive attitude, it is possible to improve and maintain the tele-operator's motivation for the task. Furthermore, I believe that the robot's autonomous function can also be improved, as seen in Study 1, by using the operation logs of para- and tele-operation. Through the operation, I believe that the robot will be able to divide its tasks into three layers: tasks that do not require a high level of judgment will be handled by the autonomous function, tasks that require a certain level of judgment or cooperation with the clerk will be

handled by para-operation, and complex tasks will be handled by tele-operation. Through such a division of labor, it is expected to be able to efficiently extract the performance of a single person in both proximity and remoteness.

However, further research is also important for the realization of robots as described above. For example, regarding the improvement of autonomous functions through operation, in addition to Study 1, it should be verified whether it can be applied in situations such as stores. Considering the context, it should be possible to collect dialog logs and implement a dialog system while managing the dialog topics. It is also important to realize an interface that can easily carry the operation of multiple robots in tele-operation.

In addition to Study 2, it is necessary to demonstrate the attitude change of the operator through operation in the real world, and it should be verified whether the same attitude change occurs in para-operation. I consider that attitude change like that in this experiment can also occur in para-operation. However, in this experiment, the operator's operation is synchronized with the robot's speech, and in para-operation, asynchronous operations such as operating the robot while engaging in other tasks occur. Therefore, it is unclear whether the same kind of attitude change will occur even with the asynchronous operation, and this needs to be verified.

For para-operation, more flexible operation methods should be established in addition to Study 3. If there is a method that does not use a button interface but can be operated by gestures, etc., it will be possible to operate the robot regardless of the interface. Also, since it is no longer a mechanical operation, there is a possibility of attributing a high degree of autonomy to the robot.

### 6.2.2 Semi-autonomous robot as humanlike and machinelike agent

In the para-operation situation, it was shown through Study 3 that operators themselves can interact with the robot they are operating as a dialog partner. Therefore, as an advanced study of this research, I consider the question of how an operator interacts with the robot when they have both senses of operating and interacting (Fig. 6.1). Specifically, in a situation where a human cooperates with a robot in a store, etc., the two senses described above would be produced when they are interacting with the robot while performing para-operations by themselves. When a human has both of these senses simultaneously, it may have the effect of making it easier to continue to accept the robot, for example, without blaming it completely

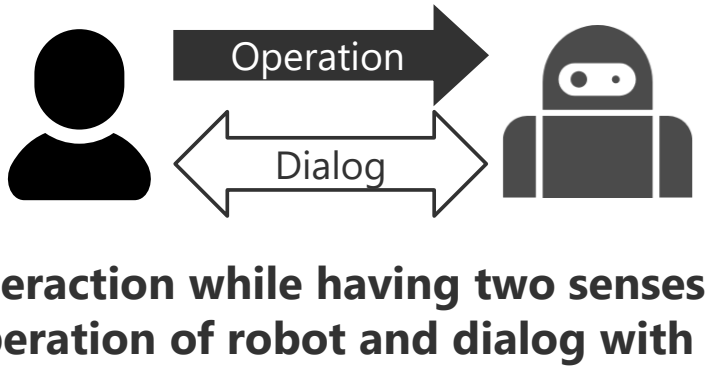



Fig. 6.1 Conceptual image of semi-autonomous robot as humanlike and machinelike agent.

if it makes some mistakes. I think that humans have the nature to hold multiple conflicting thoughts at the same time. Therefore, there is a possibility that people can have the seemingly contradictory senses of operating and interacting with a robot simultaneously. Therefore, I consider we can deepen the research by considering the following hypothesis: "Humans may be more likely to accept a robot if it has both machine-like and human-like aspects."

Conventionally, the autonomy of robots has been emphasized in robots that serve as human interaction partners. However, fully autonomous robots, like humans, may behave in ways unpredictable to humans, which may create a sense of avoidance for them. Therefore, by making the robots operatable, we may be able to control their predictability and make them more acceptable agents. However, if the robot is too machine-like, it may be difficult for it to be considered as an interaction partner, and the effectiveness of the interaction with the robot may decrease. I consider that robots get machine-like aspects through operation and human-like aspects by their human-like appearance and behavior, which can bring out the inevitable human tendency to treat the robots as others. It is necessary to maintain both aspects appropriately by balancing the operatability and behavior of robots. Semi-autonomous robots focused on in this study are not only practical for task performance but also have the potential to be beneficial in the relationship between humans and robots by emphasizing machine-like aspects.

As an extension of this research, I conducted a demonstration in which the operator interacted with a semi-autonomous android equipped with a certain level of

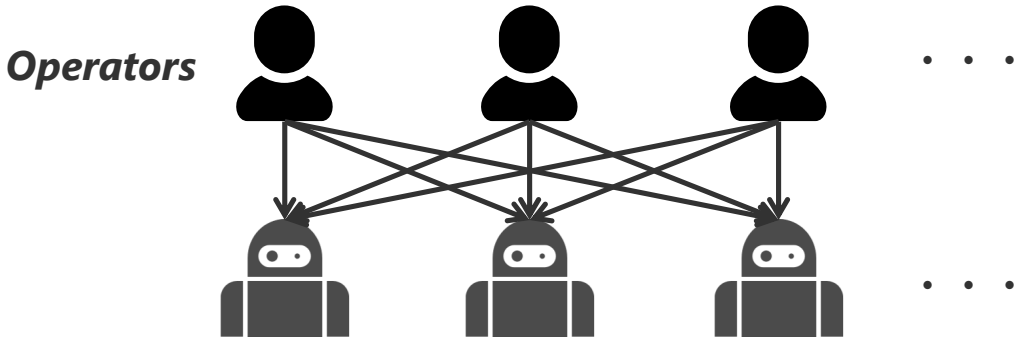



Fig. 6.2 Conceptual image of cooperation between multiple semi-autonomous robots and operators.

autonomous dialog function while performing a task by operation, and an experiment in which the operator interacted with a robot while operating it to verify differences in the sense of dialog depending on the appearance of the robot [113]. For the latter in particular, a robot that simply parrots back human speech with appropriate turn-taking timing was implemented, and an experiment in which participants actually interacted with it was conducted. In the experiment, the appearance of three robots (an android, a humanoid, and a speaker) was compared to investigate the effect of appearance on the sense of dialog. The results showed that participants felt more sense of dialog when they interact with an android or a humanoid than with a speaker, even though they were informed that the robots were only parroting back. This suggests that the robot's appearance is an important factor in fostering the human-like aspect when it is the target of the operation. In the future, while furthering these studies based on the above hypothesis, it is necessary to further investigate how the machine-like and human-like aspects should coexist and be controlled.

### 6.2.3 Cooperation between multiple semi-autonomous robots and operators

In this research, I only considered the relationship between one robot and one operator in terms of how they cooperate. In the future, as the application of semi-autonomous robots in the real world, I believe that this research will advance to the

study of "how the robots and their operators cooperate in complex situations where multiple operators operate multiple robots, both in tele/para-operation form" (Fig. 6.2). Research in this direction is already beginning [23]. As a practical example of application, in a tele-operation situation, it is expected that multiple operators perform tasks efficiently by taking turns to operate multiple robots. In a para-operation situation, it may be possible to provide a dialog service while extending an operator's presence to multiple robots in a store, for example. I think an operator should specialize in either tele-operation only or para-operation only. It may be possible that one operator can grasp both situations by devising an interface, but I guess it is better to specialize in the application.

Organizing the situation in which multiple operators operate multiple robots, can be classified into two categories: One operator and multiple robots and multiple operators and one robot. For example, as for the former, it is necessary to establish a method to realize an efficient operation interface to operate multiple robots alternatively and to know the influence of activities while extending one's presence to multiple robots. For the latter, it is necessary to clarify how to realize more appropriate operations by aggregating the intentions of multiple operators, and how to maintain the consistency of the personality of a single robot despite the existence of multiple operators. By advancing research in these areas, I believe that we can eventually realize real-world applications of systems in which multiple operators cooperate with multiple robots.

I consider how many robots can be operated by a single operator. There are various possible operation methods, but here I consider the case where an operator is responsible for decision making of robots' actions (as in Study 3). In such a case, in order to think how many robots an operator can operate, it is good to consider how many situations involving each robot the operator can grasp simultaneously. In other words, I think that the maximum number of situations that the operator can grasp is simply the maximum number of robots that can be operated. It is difficult to answer specifically how many, because it may be possible to make it easier to grasp multiple situations by devising an interface. It is also possible to realize an interface that allows an operator to recognize different robots as a single situation if their situations are similar. I believe that the number of robots that can be operated simultaneously can be increased to a certain level by developing an interface that reduces the burden on the operator to understand the situation.

# Appendix

## A Dialog script used in the experiment in chapter 4

The dialog script between the android and the experimenter used in the experiment described in chapter 4 is shown in Table A.1. This script has been translated from Japanese to English. There were a total of 10 patterns in the dialog scripts regarding the 10 paintings, and here, I show a version wherein the android recommended "Weeping Willow." The part enclosed in brackets "[ ]" in the script changed depending on the paintings that the android referred to; however, I made the script as similar as possible for all paintings.

Table A.1: Dialog script used in the experiment.

| Speaker | Script |
|---|---|
| Android | Thank you for coming in today. I am Yu, an android. |
| Experimenter | (Bows down to indicate "Thank you.") |
| Android | Well, may I ask your name? |
| Experimenter | I am *** (experimenter's name). |
| Android | ***-san. Nice to meet you. |
| Experimenter | Nice to meet you too. |
| Android | By the way, where did you come from? |
| Experimenter | Well, from Suita. |
| Android | Suita? Then, are you a student at Osaka University? |
| Experimenter | Yes, right. |
| Android | I see. Thank you very much. |
| Experimenter | Ah, no problem. |

Continued on next page

**Table A.1 – continued from previous page**

| Speaker | Script |
|---|---|
| Android | So today, I would like to talk a little about paintings. Are you interested in paintings? |
| Experimenter | Well, I do not usually appreciate paintings. |
| Android | I see. I usually introduce paintings in museums, and I would like to introduce some of my recommended paintings to you. |
| Experimenter | I see. |
| Android | Today, I will introduce the painting you were handed. It is called [“Weeping willow”]. Have you seen this painting before? |
| Experimenter | No. |
| Android | Is that so. This is my favorite painting! Please look at it. How about that? |
| Experimenter | Hmm. Well, it feels normal. |
| Android | I see. I like the colors used in [“Weeping Willow”]. I prefer [a slightly dark image and a calm feeling]. Do you like this type of color scheme? |
| Experimenter | Hmm... Color... I am not sure; I cannot say which. |
| Android | Is that so. How about the expression of light? [Even in dark colors, the effect of light is expressed]. How about that? |
| Experimenter | Ummm... I feel yes and no. |
| Android | I see. ***-san, is there any part of this picture that interests you? |
| Experimenter | Let me see... I do not know how to say it, but is it called balance? I feel the balance is good. |
| Android | Indeed. Certainly, I also think the balance of [“Weeping Willow”] is good. The composition is that [there is a tree in the middle and the space is divided appropriately into the left and the right]. |
| Experimenter | I agree. I feel it is simple. |
| Android | Yes. It is simple, but does it look cool? |
| Experimenter | Yes, I feel the composition is cool, somehow. |
| Android | I agree. I also like the theme of [“Weeping Willow”]. Does not seem like anything, but when it is drawn as a painting, it is so impressive. Do you not feel that way? |

Continued on next page

**Table A.1 – continued from previous page**

| Speaker | Script |
|---|---|
| Experimenter | Oh, I see. Indeed. |
| Android | Right. This painting is interesting because I have seen it many times and discovered many things. |
| Experimenter | Oh yeah? It sounds a little interesting when you say so. |
| Android | That is right. How is it? Do you like my favorite work, [“Weeping Willow”]? |
| Experimenter | Yes. I like it a lot more than when I first saw it. |
| Android | That was good. Today, the postcard of [“Weeping Willow”] is being sold for 216 yen. If you like it, how about it? |
| Experimenter | Really. ...Hmm, then, I would like one. |
| Android | Thank you very much. Please let the staff know after the experiment is over. |
| Experimenter | I see. |
| Android | I am glad that you like the picture. |
| Experimenter | Yeah. |
| Android | Well, our talk is over this time, thank you very much. |
| Experimenter | Thank you very much. |
| Experimenter | (Leaves the room) |

# B The list of ten paintings used in the experiment in chapter 4

I show the list of ten paintings used in the experiment described in chapter 4. All ten paintings were drawn by Claude Monet. Owing to copyright restrictions, I have not shown the images of the paintings.

- Apple trees in blossom (1873)
- Boulevard des Capucines (1873)
- Charing Cross Bridge (1903)
- Landscape: The Parc Monceau (1876)

- The Saint-Lazare Station (1877)
- Marine View with a Sunset (1875)
- The Tuileries (1876)
- Vétheuil (1879)
- Waterloo Bridge (1903)
- Weeping Willow (1918-1919)

# C Questionnaires used in the dummy task in chapter 4

In the dummy task described in chapter 4, the participants answered the following seven mandatory closed-ended questions and five non-mandatory open-ended questions. The participants were asked to work on the dummy task for the full 15 minutes; however, it was not necessary to answer all the open-ended questions.

Closed-ended questions: these were 9-point Likert scale questions (1: Not at all – 9: Very much).

- I am interested in painting.
- I am interested in various kinds of art, not just paintings.
- I often go to museums.
- My friends often go to museums.
- I would like to visit museums as much as possible if I have the chance.
- I prefer contemporary art to classical art.
- I prefer Eastern art to Western art.

Open-ended questions: For the first question given below, I used three paintings that were considered to be different in style from Monet's paintings used in the experiment in order not to influence the participants' preferences in the ranking painting task. Owing to copyright restrictions, I have not shown the images of the three paintings herein.

- Please write freely what you think is good or not good about the following paintings. Also, if you were to recommend this painting to someone else, how would you recommend it?
  - Félix Vallotton, "High Alps, Glaciers and Snowy Summits" (1919)
  - Piet Mondrian, "Gray Tree" (1911)

– Félix Vallotton, “The Visit” (1899)

- What do you think is the most effective way to communicate the appeal of painting and art to college students? You can use bullet points, so feel free to write them.
- Please freely think of a museum or exhibition that you think would make you want to visit and tell us about it.
- If you were to combine robots or artificial intelligence with art, what do you think would be interesting?

# Acknowledgments

I would like to express my deepest gratitude to supervisors of my research during the Ph.D. course, Professor Hiroshi Ishiguro and Associate Professor Kohei Ogawa. Professor Hiroshi Ishiguro gave me valuable insights and a great opportunity to study the latest humanoid robots. Associate Professor Kohei Ogawa spent much time and effort teaching me everything in research, including how to develop, investigate, and write papers. I would like to appreciate Associate Professor Yuichiro Yoshikawa for his careful guidance and discussion on my research. I would also like to thank Associate Professors Hideyuki Takahashi, Midori Ban, Yoshihiro Nakata, Assistant Professors Mahzoon Hamed, and Takahisa Uchida for their valuable suggestions and various supports. I would like to express my appreciation to the members of the dissertation committee, Professors Kensuke Harada and Koh Hosoda, for their time and consideration.

I would like to thank Mr. Takamichi Isowa for collaboration in Study 3, Dr. Kurima Sakai for developing a part of the experimental systems in Study 2, and a colleague for appearing in the experimental video on Study 2. I would also like to thank the companies and staff who were involved in the joint researches described in this dissertation.

I express my thanks to all of the staff of the Intelligent Robotics Laboratory at Osaka University for all support. I also thank my fellow lab mates and all my friends for the meaningful insights and all the fun we have had. Lastly, I want to thank my family for their constant support and encouragement, swearing to what is lost by the words of gratitude.

# Bibliography


1) Michael A. Goodrich and Alan C. Schultz. Human-robot interaction: A survey. *Foundations and Trends in Human-Computer Interaction*, Vol. 1, No. 3, pp. 203–275, 2007.
2) Christoph Bartneck, Tony Belpaeme, Friederike Eyssel, Takayuki Kanda, Merel Keijsers, and Selma Šabanović. *Human-robot interaction: an introduction*. Cambridge University Press, 2020.
3) Yingfeng Chen, Feng Wu, Wei Shuai, and Xiaoping Chen. Robots serve humans in public places—KeJia robot as a shopping assistant. *International Journal of Advanced Robotic Systems*, Vol. 14, No. 3, p. 1729881417703569, 2017.
4) Junya Nakanishi, Itaru Kuramoto, Jun Baba, Kohei Ogawa, Yuichiro Yoshikawa, and Hiroshi Ishiguro. Continuous hospitality with social robots at a hotel. *SN Applied Sciences*, Vol. 2, No. 3, pp. 1–13, 2020.
5) Hideki Asoh, Yoichi Motomura, Futoshi Asano, Isao Hara, Satoru Hayamizu, Katsunobu Itou, Takio Kurita, Toshihiro Matsui, Nikos Vlassis, Roland Bunschoten, et al. Jijo-2: An office robot that communicates and learns. *IEEE Intelligent Systems*, Vol. 16, No. 5, pp. 46–55, 2001.
6) Paul Baxter, Emily Ashurst, Robin Read, James Kennedy, and Tony Belpaeme. Robot education peers in a situated primary school study: Personalisation promotes child learning. *PloS one*, Vol. 12, No. 5, p. e0178126, 2017.
7) Takamasa Iio, Satoru Satake, Takayuki Kanda, Kotaro Hayashi, Florent Ferreri, and Norihiro Hagita. Human-like guide robot that proactively explains exhibits. *International Journal of Social Robotics*, Vol. 12, No. 2, pp. 549–566, 2020.
8) Brian R Duffy. Anthropomorphism and the social robot. *Robotics and autonomous systems*, Vol. 42, No. 3-4, pp. 177–190, 2003.
9) Eric Paulos and John Canny. Social tele-embodiment: Understanding presence. *Autonomous Robots*, Vol. 11, No. 1, pp. 87–95, 2001.

10) Shuichi Nishio, Hiroshi Ishiguro, and Norihiro Hagita. Geminoid: Teleoperated android of an existing person. *Humanoid robots: New developments*, Vol. 14, pp. 343–352, 2007.
11) Sang-Seok Yun, Munsang Kim, and Mun-Taek Choi. Easy interface and control of tele-education robots. *International Journal of Social Robotics*, Vol. 5, No. 3, pp. 335–343, 2013.
12) Kasumi Abe, Masahiro Shiomi, Yachao Pei, Tingyi Zhang, Narumitsu Ikeda, and Takayuki Nagai. ChiCaRo: tele-presence robot for interacting with babies and toddlers. *Advanced Robotics*, Vol. 32, No. 4, pp. 176–190, 2018.
13) Jun Baba, Song Sichao, Junya Nakanishi, Itaru Kuramoto, Kohei Ogawa, Yuichiro Yoshikawa, and Hiroshi Ishiguro. Teleoperated robot acting autonomous for better customer satisfaction. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–8, 2020.
14) Masaya Iwasaki, Mizuki Ikeda, Tatsuyuki Kawamura, and Hideyuki Nakanishi. State-transition modeling of human–robot interaction for easy crowdsourced robot control. *Sensors*, Vol. 20, No. 22, p. 6529, 2020.
15) Jiro Shimaya, Yuichiro Yoshikawa, Hirokazu Kumazaki, Yoshio Matsumoto, Masutomo Miyao, and Hiroshi Ishiguro. Communication support via a tele-operated robot for easier talking: case/laboratory study of individuals with/without autism spectrum disorder. *International Journal of Social Robotics*, Vol. 11, No. 1, pp. 171–184, 2019.
16) Kaiko Kuwamura, Shuichi Nishio, and Shinichi Sato. Can we talk through a robot as if face-to-face? long-term fieldwork using teleoperated robot for seniors with alzheimer's disease. *Frontiers in psychology*, Vol. 7, p. 1066, 2016.
17) Masahiro Shiomi, Daisuke Sakamoto, Takayuki Kanda, Carlos Toshinori Ishi, Hiroshi Ishiguro, and Norihiro Hagita. A semi-autonomous communication robot: A field trial at a train station. In *Proceedings of the 3rd ACM/IEEE International Conference on Human-Robot Interaction*, pp. 303–310, 2008.
18) Jun Ki Lee, Walter Dan Stiehl, Robert Lopez Toscano, and Cynthia Breazeal. Semi-autonomous robot avatar as a medium for family communication and education. *Advanced Robotics*, Vol. 23, No. 14, pp. 1925–1949, 2009.
19) Terrence Fong, Charles Thorpe, and Charles Baur. *Collaborative control: A robot-centric model for vehicle teleoperation*, Vol. 1. Carnegie Mellon University, The Robotics Institute Pittsburgh, 2001.

20) Takuma Seno, Kohei Okuoka, Masahiko Osawa, and Michita Imai. Adaptive semi-autonomous agents via episodic control. In *Proceedings of the 6th International Conference on Human-Agent Interaction*, pp. 377–379, 2018.
21) Abolfazl Zaraki, Luke Wood, Ben Robins, and Kerstin Dautenhahn. Development of a semi-autonomous robotic system to assist children with autism in developing visual perspective taking skills. In *Proceedings of the 27th IEEE International Symposium on Robot and Human Interactive Communication*, pp. 969–976, 2018.
22) Katherine M Tsui, Adam Norton, Daniel J Brooks, Eric McCann, Mikhail S Medvedev, Jordan Allspaw, Sompop Suksawat, James M Dalphond, Michael Lunderville, and Holly A Yanco. Iterative design of a semi-autonomous social telepresence robot research platform: a chronology. *Intelligent Service Robotics*, Vol. 7, No. 2, pp. 103–119, 2014.
23) Dylan F Glas, Takayuki Kanda, Hiroshi Ishiguro, and Norihiro Hagita. Teleoperation of multiple social robots. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, Vol. 42, No. 3, pp. 530–544, 2011.
24) Hiroshi Ishiguro. The realisation of an avatar-symbiotic society where everyone can perform active roles without constraint. *Advanced Robotics*, pp. 1–7, 2021.
25) Kazuaki Tanaka, Hideyuki Nakanishi, and Hiroshi Ishiguro. Physical embodiment can produce robot operator's pseudo presence. *Frontiers in ICT*, Vol. 2, p. 8, 2015.
26) Masaya Iwasaki, Jian Zhou, Mizuki Ikeda, Yuki Koike, Yuya Onishi, Tatsuyuki Kawamura, and Hideyuki Nakanishi. "That robot stared back at me!": Demonstrating perceptual ability is key to successful human–robot interactions. *Frontiers in Robotics and AI*, Vol. 6, p. 85, 2019.
27) Hiroshi Ishiguro. Android science. In *Robotics Research*, pp. 118–127. Springer, 2007.
28) Kazuhiko Shinozawa, Futoshi Naya, Junji Yamato, and Kiyoshi Kogure. Differences in effect of robot and screen agent recommendations on human decision-making. *International journal of human-computer studies*, Vol. 62, No. 2, pp. 267–279, 2005.
29) Wilma A Bainbridge, Justin Hart, Elizabeth S Kim, and Brian Scassellati. The effect of presence on human-robot interaction. In *Proceedings of the 17th IEEE International Symposium on Robot and Human Interactive Communication*, pp.

701–706, 2008.

30) Aaron Powers, Sara Kiesler, Susan Fussell, and Cristen Torrey. Comparing a computer agent with a humanoid robot. In *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction*, pp. 145–152, 2007.

31) Kerstin S Haring, David Silvera-Tawil, Tomotaka Takahashi, Katsumi Watanabe, and Mari Velonaki. How people perceive different robot types: A direct comparison of an android, humanoid, and non-biomimetic robot. In *Proceedings of the 8th International Conference on Knowledge and Smart Technology*, pp. 265–270, 2016.

32) Laurel D. Riek, Tal-Chen Rabinowitch, Bhismadev Chakrabarti, and Peter Robinson. How anthropomorphism affects empathy toward robots. In *Proceedings of the 4th ACM/IEEE International Conference on Human-Robot Interaction*, pp. 245–246, 2009.

33) Jennifer Goetz, Sara Kiesler, and Aaron Powers. Matching robot appearance and behavior to tasks to improve human-robot cooperation. In *Proceedings of the 12th IEEE International Workshop on Robot and Human Interactive Communication*, pp. 55–60, 2003.

34) Shuichi Nishio and Hiroshi Ishiguro. Attitude change induced by different appearances of interaction agents. *International Journal of Machine Consciousness*, Vol. Vol.3, No.1, pp. 115–126, 2011.

35) Eileen Roesler, Dietrich Manzey, and Linda Onnasch. A meta-analysis on the effectiveness of anthropomorphism in human-robot interaction. *Science Robotics*, Vol. 6, No. 58, p. eabj5425, 2021.

36) Xiao Dou, Chih-Fu Wu, Kai-Chieh Lin, Senzhong Gan, and Tzu-Min Tseng. Effects of different types of social robot voices on affective evaluations in different application fields. *International Journal of Social Robotics*, Vol. 13, No. 4, 2021.

37) Henny Admoni and Brian Scassellati. Social eye gaze in human-robot interaction: a review. *Journal of Human-Robot Interaction*, Vol. 6, No. 1, pp. 25–63, 2017.

38) Ross Mead and Maja J Matarić. Autonomous human–robot proxemics: socially aware navigation based on interaction potential. *Autonomous Robots*, Vol. 41, No. 5, pp. 1189–1201, 2017.

39) Masahiro Shiomi, Hidenobu Sumioka, and Hiroshi Ishiguro. Survey of social

touch interaction between humans and robots. *Journal of Robotics and Mechatronics*, Vol. 32, No. 1, pp. 128–135, 2020.

40) Masahiro Shiomi, Aya Nakata, Masayuki Kanbara, and Norihiro Hagita. Robot reciprocation of hugs increases both interacting times and self-disclosures. *International Journal of Social Robotics*, Vol. 13, No. 2, pp. 353–361, 2021.
41) Seiji Yamada. 人とロボットの “間” をデザインする (in Japanese). Tokyo Denki University Press, 2007.
42) Dylan F Glas, Takashi Minato, Carlos T Ishi, Tatsuya Kawahara, and Hiroshi Ishiguro. ERICA: The ERATO intelligent conversational android. In *Proceedings of the 25th IEEE International Symposium on Robot and Human Interactive Communication*, pp. 22–29, 2016.
43) Gabriel Skantze. Turn-taking in conversational systems and human-robot interaction: a review. *Computer Speech & Language*, p. 101178, 2020.
44) Shane Saunderson and Goldie Nejat. How robots influence humans: A survey of nonverbal communication in social human–robot interaction. *International Journal of Social Robotics*, Vol. 11, No. 4, pp. 575–608, 2019.
45) Kazuki Sakai, Fabio Dalla Libera, Yuichiro Yoshikawa, and Hiroshi Ishiguro. Generation of bystander robot actions based on analysis of relative probability of human actions. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, Vol. 21, No. 4, pp. 686–696, 2017.
46) Mikio Nakano, Kazunori Komatani, Kotaro Funakoshi, Yukiko Nakano, and Manabu Okumura. *Dialog systems (in Japanese)*. CORONA PUBLISHING CO.,LTD., 2015.
47) Kohei Ogawa, Shuichi Nishio, Kensuke Koda, Giuseppe Balistreri, Tetsuya Watanabe, and Hiroshi Ishiguro. Exploring the natural reaction of young and aged person with telenoid in a real world. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, Vol. 15, No. 5, pp. 592–597, 2011.
48) Susumu Tachi. Telexistence: Enabling humans to be virtually ubiquitous. *IEEE computer graphics and applications*, Vol. 36, No. 1, pp. 8–14, 2016.
49) Maryam Alimardani, Shuichi Nishio, and Hiroshi Ishiguro. Humanlike robot hands controlled by brain activity arouse illusion of ownership in operators. *Scientific reports*, Vol. 3, No. 1, pp. 1–5, 2013.
50) Satoshi Koizumi, Takayuki Kanda, Masahiro Shiomi, Hiroshi Ishiguro, and Norihiro Hagita. Preliminary field trial for teleoperated communication robots.

In *Proceedings of the 15th IEEE International Symposium on Robot and Human Interactive Communication*, pp. 145–150, 2006.

51) Abolfazl Zaraki, Luke Wood, Ben Robins, and Kerstin Dautenhahn. Development of a semi-autonomous robotic system to assist children with autism in developing visual perspective taking skills. In *Proceedings of the 27th IEEE International Symposium on Robot and Human Interactive Communication*, pp. 969–976, 2018.

52) Kohei Okuoka, Yusuke Takimoto, Masahiko Osawa, and Michita Imai. Semi-autonomous telepresence robot for adaptively switching operation using inhibition and disinhibition mechanism. In *Proceedings of the 6th International Conference on Human-Agent Interaction*, pp. 167–175, 2018.

53) Masahiko Osawa, Kohei Okuoka, Yusuke Takimoto, and Michita Imai. Is automation appropriate? semi-autonomous telepresence architecture focusing on voluntary and involuntary movements. *International Journal of Social Robotics*, Vol. 12, No. 5, pp. 1119–1134, 2020.

54) Yuya Ochiai, Kentaro Takemura, Atsutoshi Ikeda, Jun Takamatsu, and Tsukasa Ogasawara. Remote control system for multiple mobile robots using touch panel interface and autonomous mobility. In *Proceedings of the 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 3272–3277, 2014.

55) Norman Donald. *Emotional Design.* Basic Books, 2003.

56) Margaret L Traeger, Sarah Strohkorb Sebo, Malte Jung, Brian Scassellati, and Nicholas A Christakis. Vulnerable robots positively shape human conversational dynamics in a human–robot team. *Proceedings of the National Academy of Sciences*, Vol. 117, No. 12, pp. 6370–6375, 2020.

57) Anca D. Dragan, Shira Bauman, Jodi Forlizzi, and Siddhartha S. Srinivasa. Effects of robot motion on human-robot collaboration. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*, pp. 51–58, 2015.

58) Taemie Kim and Pamela Hinds. Who should i blame? effects of autonomy and transparency on attributions in human-robot interaction. In *Proceedings of the 15th IEEE International Symposium on Robot and Human Interactive Communication*, pp. 80–85, 2006.

59) Arash Ajoudani, Andrea Maria Zanchettin, Serena Ivaldi, Alin Albu-Schäffer,

Kazuhiro Kosuge, and Oussama Khatib. Progress and prospects of the human–robot collaboration. *Autonomous Robots*, Vol. 42, No. 5, pp. 957–975, 2018.

60) Kazuaki Tanaka, Kota Takenouchi, Kohei Ogawa, Yuichiro Yoshikawa, Shuichi Nishio, and Hiroshi Ishiguro. Maintaining the sense of agency in semi-autonomous robot conferencing. *Future Internet*, Vol. 11, No. 7, p. 143, 2019.
61) Shuichi Nishio, Koichi Taura, Hidenobu Sumioka, and Hiroshi Ishiguro. Teleoperated android robot as emotion regulation media. *International Journal of Social Robotics*, Vol. 5, No. 4, pp. 563–573, 2013.
62) Jiro Shimaya. *Using Tele-Operated Robots to Promote Conversation.* PhD thesis, Osaka University, 2020.
63) Jiro Shimaya, Yuichiro Yoshikawa, Oskar Palinko, Kohei Ogawa, Nobuhiro Jinnai, and Hiroshi Ishiguro. Active participation in lectures via a collaboratively controlled robot. *International Journal of Social Robotics*, Vol. 13, No. 4, 2021.
64) Jiro Shimaya, Yuichiro Yoshikawa, Kohei Ogawa, and Hiroshi Ishiguro. Robotic question support system to reduce hesitation for face-to-face questions in lectures. *Journal of Computer Assisted Learning*, Vol. 37, No. 3, pp. 621–631, 2021.
65) Tomonori Kubota, Kohei Ogawa, and Hiroshi Ishiguro. Implementation and evaluation of chat-oriented dialogue system for an android robot in live streaming media in which users can speak at any time (in Japanese). *Transactions of the Japanese Society for Artificial Intelligence*, Vol. 33, No. 1, pp. DSH–G_1–13, 2018.
66) Takanori Komatsu and Seiji Yamada. Adaptation gap hypothesis: How differences between users' expected and perceived agent functions affect their subjective impression. *Journal of Systemics, Cybernetics and Informatics*, Vol. 9, No. 1, pp. 67–74, 2011.
67) Tatsuya Kawahara, Hiroaki Kawashima, Takatsugu Hirayama, and Takashi Matsuyama. "Automated Information Concierge" based on proactive dialog and information retrieval (in Japanese). *IPSJ Magazine*, Vol. 49, No. 8, pp. 912–918, 2008.
68) Ryuhei Nisimura, Yohei Nishihara, Ryosuke Tsurumi, Akinobu Lee, Hiroshi Saruwatari, and Kiyohiro Shikano. Takemaru-kun: Speech-oriented information system for real world research platform. In *Proceedings of First International Workshop on Language Understanding and Agents for Real World Interaction*,

pp. 70–78, 2003.

69) Takayuki Kanda, Masahiro Shiomi, Zenta Miyashita, Hiroshi Ishiguro, and Norihiro Hagita. An affective guide robot in a shopping mall. In *Proceedings of the 4th ACM/IEEE International Conference on Human-Robot Interaction*, pp. 173–180, 2009.

70) Joseph Weizenbaum. ELIZA–a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, Vol. 9, No. 1, pp. 36–45, 1966.

71) Richard S Wallace. The anatomy of ALICE. In *Parsing the Turing Test*, pp. 181–210. Springer, 2009.

72) Oriol Vinyals and Quoc Le. A neural conversational model. *arXiv preprint arXiv:1506.05869*, 2015.

73) Fumihiro Bessho, Tatsuya Harada, and Yasuo Kuniyoshi. Dialog system using real-time crowdsourcing and two-length tweet corpus (in Japanese). *IPSJ SIG Technical Report*, Vol. 2012-SLP-91, No. 13, pp. 1–8, 2012.

74) Norman M Fraser and G Nigel Gilbert. Simulating speech systems. *Computer Speech & Language*, Vol. 5, No. 1, pp. 81–99, 1991.

75) Hiroaki Sugiyama, Toyomi Meguro, Ryuichiro Higashinaka, and Yasuhiro Minami. Large-scale collection and analysis of personal question-answer pairs for conversational agents. In *Proceedings of the 14th International Conference on Intelligent Virtual Agents*, pp. 420–433, 2014.

76) Takaaki Sugiyama, Kotaro Funakoshi, Mikio Nakano, and Kazunori Komatani. Estimating response obligation in multi-party human-robot dialogues. In *Proceedings of the 2015 IEEE-RAS 15th International Conference on Humanoid Robots*, pp. 166–172, 2015.

77) Yasutomo Kimura, Rafal Rzepka, and Keiichi Takamaru. Proposal of Radiobots based spoken dialogue system (in Japanese). *Proceedings of the 22th Annual Conference of JSAI*, Vol. 29, pp. 1–4, 2015.

78) Matt Kusner, Yu Sun, Nicholas Kolkin, and Kilian Weinberger. From word embeddings to document distances. In *Proceedings of the 32nd International Conference on Machine Learning*, Vol. 37, pp. 957–966, 2015.

79) Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeffrey Dean. Distributed representations of words and phrases and their compositionality. In *Proceedings of the 26th International Conference on Neural Information Pro-*

*cessing Systems - Volume 2*, pp. 3111–3119, 2013.

80) Taku Kudo, Kaoru Yamamoto, and Yuji Matsumoto. Applying conditional random fields to Japanese morphological analysis. In *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing*, pp. 230–237, 2004.

81) Taku Kudo. MeCab : Yet another part-of-speech and morphological analyzer. *https://taku910.github.io/mecab/*, 2005.

82) Toshinori Sato, Taiichi Hashimoto, and Manabu Okumura. Implementation of a word segmentation dictionary called mecab-ipadic-NEologd and study on how to use it effectively for information retrieval (in Japanese). In *Proceedings of the Twenty-three Annual Meeting of the Association for Natural Language Processing*, pp. NLP2017–B6–1, 2017.

83) Toshinori Sato. Neologism dictionary based on the language resources on the web for mecab. *https://github.com/neologd/mecab-ipadic-neologd*, 2015.

84) Kurima Sakai, Takashi Minato, Carlos T. Ishi, and Hiroshi Ishiguro. Speech driven trunk motion generating system based on physical constraint. In *Proceedings of the 25th IEEE International Symposium on Robot and Human Interactive Communication*, pp. 232–239, 2016.

85) Miki Watanabe, Kohei Ogawa, and Hiroshi Ishiguro. Can androids be salespeople in the real world? In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, pp. 781–788, 2015.

86) Aryeh G. Kidron. The effectiveness of experiential methods in training and education: The case of role playing. *Academy of Management Review*, Vol. 2, No. 3, pp. 490–495, 1977.

87) Irving Lester Janis and Bert T. King. The influence of role playing on opinion change. *Journal of abnormal psychology*, Vol. 49, No. 2, pp. 211–218, 1954.

88) Nick Yee, Jeremy N Bailenson, and Nicolas Ducheneaut. The proteus effect: Implications of transformed digital self-representation on online and offline behavior. *Communication Research*, Vol. 36, No. 2, pp. 285–312, 2009.

89) Gunwoo Yoon and Patrick T Vargas. Know thy avatar: The unintended effect of virtual-self representation on behavior. *Psychological science*, Vol. 25, No. 4, pp. 1043–1045, 2014.

90) Robin S Rosenberg, Shawnee L Baughman, and Jeremy N Bailenson. Virtual

superheroes: Using superpowers in virtual reality to encourage prosocial behavior. *PloS one*, Vol. 8, No. 1, p. e55003, 2013.

91) Domna Banakou, Sameer Kishore, and Mel Slater. Virtually being einstein results in an improvement in cognitive task performance and a decrease in age bias. *Frontiers in psychology*, Vol. 9, p. 917, 2018.
92) Shinobu Kitayama, Alana Conner Snibbe, Hazel Rose Markus, and Tomoko Suzuki. Is there any "free" choice? self and dissonance in two cultures. *Psychological Science*, Vol. 15, No. 8, pp. 527–533, 2004.
93) Claude M Steele, Steven J Spencer, and Michael Lynch. Self-image resilience and dissonance: The role of affirmational resources. *Journal of personality and social psychology*, Vol. 64, No. 6, p. 885, 1993.
94) Kurima Sakai, Takashi Minato, Carlos T. Ishi, and Hiroshi Ishiguro. Novel speech motion generation by modeling dynamics of human speech production. *Frontiers in Robotics and AI*, Vol. 4, p. 49, 2017.
95) R Core Team. R: A language and environment for statistical computing. *R Foundation for Statistical Computing*, 2020.
96) Torsten Hothorn and Kurt Hornik. exactranktests: Exact distributions for rank and permutation tests. 2021. R package version 0.8-32.
97) Marco Torchiano. effsize: Efficient effect size computation. 2020. R package version 0.8.1.
98) Pericle Salvini, Gaetano Ciaravella, Wonpil Yu, Gabriele Ferri, Alessandro Manzi, Barbara Mazzolai, Cecilia Laschi, Sang-Rok Oh, and Paolo Dario. How safe are service robots in urban environments? bullying a robot. In *Proceedings of the 19th International Symposium in Robot and Human Interactive Communication*, pp. 1–7, 2010.
99) Mark Scheeff, John Pinto, Kris Rahardja, Scott Snibbe, and Robert Tow. Experiences with sparky, a social robot. In *Socially intelligent agents*, pp. 173–180. Springer, 2002.
100) Iolanda Leite, André Pereira, Samuel Mascarenhas, Carlos Martinho, Rui Prada, and Ana Paiva. The influence of empathy in human–robot relations. *International journal of human-computer studies*, Vol. 71, No. 3, pp. 250–260, 2013.
101) Mikey Siegel, Cynthia Breazeal, and Michael I Norton. Persuasive robotics: The influence of robot gender on human behavior. In *Proceedings of the 2009*

*IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 2563–2568, 2009.

102) Domna Banakou and Mel Slater. Body ownership causes illusory self-attribution of speaking and influences subsequent real speaking. *Proceedings of the National Academy of Sciences*, Vol. 111, No. 49, pp. 17678–17683, 2014.

103) Daryl J Bem. Self-perception theory. In *Advances in experimental social psychology*, Vol. 6, pp. 1–62. Elsevier, 1972.

104) Tomonori Kubota, Takamichi Isowa, Kohei Ogawa, and Hiroshi Ishiguro. Development and verication of an onsite-operated android robot working cooperatively with humans in a real store (in Japanese). *The Transactions of Human Interface Society*, Vol. 22, No. 3, pp. 275–290, 2020.

105) SoftBank Corp. Pepper, 2014.

106) Masahiro Shiomi, Kazuhiko Shinozawa, Yoshifumi Nakagawa, Takahiro Miyashita, Toshio Sakamoto, Toshimitsu Terakubo, Hiroshi Ishiguro, and Norihiro Hagita. Recommendation effects of a social robot for advertisement-use context in a shopping mall. *International Journal of Social Robotics*, Vol. 5, No. 2, pp. 251–262, 2013.

107) H-M Gross, H Boehme, Ch Schroeter, Steffen Müller, Alexander König, Erik Einhorn, Ch Martin, Matthias Merten, and Andreas Bley. TOOMAS: Interactive shopping guide robots in everyday use - final implementation and experiences from long-term field trials. In *Proceedings of the 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 2005–2012, 2009.

108) Takayuki Kanda, Masahiro Shiomi, Zenta Miyashita, Hiroshi Ishiguro, and Norihiro Hagita. A communication robot in a shopping mall. *IEEE Transactions on Robotics*, Vol. 26, No. 5, pp. 897–913, 2010.

109) Marketta Niemelä, Anne Arvola, and Iina Aaltonen. Monitoring the acceptance of a social service robot in a shopping mall: First results. In *Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, pp. 225–226, 2017.

110) Marketta Niemelä, Päivi Heikkilä, and Hanna Lammi. A social service robot in a shopping mall: Expectations of the management, retailers and consumers. In *Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, pp. 227–228, 2017.

111) Chao Shi, Satoru Satake, Takayuki Kanda, and Hiroshi Ishiguro. How would store managers employ social robots? In *Proceedings of the 11th ACM/IEEE International Conference on Human-Robot Interaction*, pp. 519–520, 2016.

112) Chris L Kleinke, Thomas R Peterson, and Thomas R Rutledge. Effects of self-generated facial expressions on mood. *Journal of Personality and Social Psychology*, Vol. 74, No. 1, p. 272, 1998.

113) Yuka Arai. Verification of self-counseling effectiveness using a parroting dialog robot (in Japanese). Master's thesis, Osaka University, 2021.

# Publications

## Reviewed Journal Papers

1 窪田 智徳, 小川 浩平, 石黒 浩, “不特定多数のユーザが随時発話可能なライブストリーミングメディアにおけるアンドロイドロボットを用いた雑談対話システムの実現と評価”, 人工知能学会論文誌, 2018, 33 巻, 1 号, p. DSH-G_1-13, https://doi.org/10.1527/tjsai.DSH-G

2 窪田 智徳, 磯和 隆道, 小川 浩平, 石黒 浩, “人と共働する近接操作型アンドロイドロボットの開発と実店舗での検証”, ヒューマンインタフェース学会論文誌, 2020, 22 巻, 3 号, p. 275–290, https://doi.org/10.11184/his.22.3_275

3 Tomonori Kubota, Kohei Ogawa, Yuichiro Yoshikawa, Hiroshi Ishiguro, “Alignment of the attitude of tele-operators with that of a semi-autonomous android” (submitted)

## Non-Reviewed Conference Papers

1 窪田 智徳, 小川 浩平, 石黒 浩, “アンドロイドル U：社会的に大きな存在感の獲得に向けたニコニコ生放送における対話アンドロイドの実現への取り組み”, 人工知能学会全国大会論文集, 2017, JSAI2017, p. 2N23, https://doi.org/10.11517/pjsai.JSAI2017.0_2N23

## Competitive Research Funds

1 日本学術振興会 科学研究費助成事業 特別研究員奨励費, “操作者と対話相手双方への有効性に着目した実用的な半自律対話ロボットの実現”（直接経費: 2,100 千円）, 研究課題/領域番号 20J13662, 2020 年 4 月〜2022 年 3 月

# Miscellanies

## Poster Presentations

1 Tomonori Kubota, Kohei Ogawa, Yuichiro Yoshikawa, Hiroshi Ishiguro, “An Attitude of a tele-operator Manipulating a Part of an Android’s Body Gets Close to an Attitude the Android Autonomously Expresses”, The 1st International Symposium on Symbiotic Intelligent Systems, Osaka, Japan, January 2019

## Demonstrations

1 ニコニコ生放送 “「アンドロイドでもニコ生がしたい」シリーズ（全 24 回）”, 2016 年 12 月〜2017 年 4 月
2 ニコニコ生放送 “アンドロイドル「U」育成に関する記者発表会”, 来場者数 6,536 人, 2017 年 2 月
3 ニコニコ生放送 “【初のアンドロイド MC】選択式トークの恋愛実験２”, 来場者数 20,941 人, 2017 年 3 月
4 ニコニコ生放送 “ロボット MC〜アンドロイドルの「U」なま”, 来場者数 16,386 人, 2017 年 4 月
5 ニコニコ生放送 “アンドロイドル「U」の地下倉庫で朝まで寝姿配信”, 来場者数 40,070 人, 2017 年 4 月
6 池袋 PARCO “インフォメーションカウンターでの接客”, 東京, 2017 年 4 月
7 ニコニコ超会議 2017 “石黒研究室の恋愛実験神社”, 千葉, 2017 年 4 月
8 大阪大学共創 DAY@EXPOCITY 大阪大学とあそぼう “アンドロイド特別展示”, 大阪, 2018 年 11 月
9 新学術領域研究「人間機械共生社会を目指した対話知能システム学」記者発表 “半自律アンドロイドを用いた受付デモ”, 大阪, 2019 年 11 月
10 毎日放送 “ロボロボ”, TV 番組, 2019 年 9 月