A Study on Facilitating Social Signaling Using Paired Devices with Visual and Haptic Cues

March　２０１９

Nunez Morales Eleuda Rosa

A Study on Facilitating Social Signaling Using Paired Devices with Visual and Haptic Cues

Graduate School of Systems and Information Engineering
University of Tsukuba

March ２０１９

Nunez Morales Eleuda Rosa

Abstract

There are several benefits that humans obtain from social interaction as it was found to be related to individuals well-being. A high degree of these positive effects comes from being able to recognize and convey the intention to communicate. Social signals are observable behavioral cues used during the exchange through social interaction, and they convey the intention to communicate. This research proposes to mediate human-human communication using paired devices that deliver visual and haptic cues. These cues are used with the role of social signals in scenarios where the available social information is limited. This approach was evaluated on two different communication scenarios with populations that might benefit from the characteristics of the proposed devices. One application is for remote communication scenarios, based on the assumption that the physical embodiment and touch-based cues can complement the affective social signaling of individuals trying to communicate from different places. Another application involves individuals with Autism Spectrum Disorder, using the proposed devices to mediate and describe the exchange in turn-taking interventions, using the cues from the devices as a guide. The methodology of this research involved: 1) Select target users that, based on literature review, could benefit from this approach, 2) Design and implement solutions involving paired devices for facilitating social signals, 3) Performance evaluation: to explore the capabilities of the devices for sensing and conveying cues, and 4) User study: to understand the role of the cues and their effect on the user. At the end of this study it is discussed if the proposed visual and haptic cues delivered by paired devices worked as social signals, and what was the effect these cues have on human communication. Based on this, different parameters that condition the effect of this approach are described. This new interaction setting comes together with different questions, especially considering that 1) "the behavior of the user" is different from "the behavior detected by the sensors" and 2) the decoded message representing a human behavior is made only with simple visual and haptic cues.

Abstract

他者との社会的相互作用の量と質は，個人の感じる幸福に密接に関連していることが明らかになっている．これらの効用の大部分は，人がコミュニケーションに関する意図を認識し伝達できることに起因しており，社会的相互作用の場面において意図の伝達のために使用される観察可能な非言語行動は「ソーシャル・シグナル行動」と呼ばれる．本研究では，環境や障害などにより利用可能なソーシャル・シグナルが限られている場面において，ペア型デバイスによる視触覚性の情報を用いることで複数人間のソーシャル・シグナル行動を仲介・促進することを提案する．ここでは，提案手法が有効であると考えられる 2 つの社会的相互作用の場面を想定した．一つは，触覚を用いるインタラクションが困難な遠隔コミュニケーションの場面であり，実体のあるデバイスを用いた視触覚情報に基づく相互作用を実現することで感情や情動に関するソーシャル・シグナル行動を補完できるという仮説に基づき，ペア型デバイスの設計・開発と評価を行った．もう一つは，他者の意図の認識に困難が生じる自閉症スペクトラム障害を持つ人々を対象とし，デバイスにより提供される視覚触覚情報を手がかりとしてターンテイキング行動を促進するための介入を行った． それぞれの事例研究は次のような統一的な方法論に沿って遂行された．1) 文献調査に基づく対象ユーザの選択，2) ソーシャル・シグナル行動を促進するためのアプローチとそれを実現するペア型デバイスの設計，3) デバイスによるソーシャル・シグナル行動の認識と伝達に関する性能評価，および4) それがユーザに与える影響を理解するユーザ研究である． 最後に，これらの事例研究に基づき，ペア型デバイスにより提供された視触覚性情報がソーシャル・シグナルとして機能したかどうか，またこれらの手がかりが人のコミュニケーションにどのような影響を及ぼしたかについて議論する．それにより，提案手法によるソーシャル・シグナル行動の促進効果に作用する様々なパラメータに関する示唆が得られた．さらに，行動に対する人の認知とセンサーによる計測結果との齟齬や，視触覚以外のモダリティを用いたソーシャル・シグナル行動の促進など，新たな発展的課題や解決のアプローチについての知見も得られた．

Abstract

Existen varios beneficios que los humanos obtienen de la interacción social, ya que se ha determinado que esta relacionada con el bienestar de las personas. Un alto grado de estos efectos positivos proviene de poder reconocer y transmitir la intención de comunicarse. Las señales sociales son señales de comportamiento observables que se utilizan durante el intercambio a través de la interacción social y transmiten la intención de comunicarse. Esta investigación propone mediar en la comunicación entre humano-humano utilizando dispositivos emparejados que brindan señales visuales y hápticas. Estas señales se utilizan con la función de señales sociales en escenarios donde la información social disponible es limitada. Este enfoque se evaluó en dos escenarios de comunicación diferentes con poblaciones que pueden beneficiarse de las características de los dispositivos propuestos. Una de las aplicaciones es para escenarios de comunicación remota, basada en el supuesto de que la representación física y las señales táctiles pueden complementar la señalización social afectiva de las personas que intentan comunicarse desde diferentes lugares. Otra aplicación involucra a personas con trastorno del espectro autista, y utiliza los dispositivos propuestos para mediar y describir el intercambio en intervenciones para entrenar la habilidad de tomar turnos, utilizando las indicaciones de los dispositivos como guía. La metodología de esta investigación involucró: 1) Seleccionar usuarios objetivo que, según la revisión bibliográfica, podrían beneficiarse de este enfoque, 2) Diseñar e implementar soluciones que involucren dispositivos emparejados para facilitar las señales sociales, 3) Evaluación del rendimiento: para explorar las capacidades de los dispositivos para detectar y transmitir señales, y 4) Estudio del usuario: para comprender el papel de las señales y su efecto en el usuario. Al final de este estudio se discute si las señales visuales y hápticas representadas en dispositivos emparejados funcionaron como señales sociales, así como su efecto en la comunicación humana. En base a esto, se describen diferentes parámetros que condicionan el efecto de las señales visuales y hápticas en dispositivos emparejados. Esta nueva configuración de interacción trae consigo diferentes preguntas, especialmente considerando que 1) "El comportamiento del usuario" es diferente de "el comportamiento detectado por los sensores de los dispositivos" y 2) El mensaje decodificado que representa el comportamiento humano se hace solo con señales visuales y hápticas.

Acknowledgements

First and foremost I want to thank my academic advisor Kenji Suzuki. I am grateful for all those meetings, for your time and effort, and for making my Ph.D. experience productive and stimulating. "Sensei", your enthusiasm was contagious and motivational for me, even during tough times. I could not have imagined having a better mentor for my Ph.D. study, and I will be forever grateful to you.

Besides my advisor, I would like to thank the rest of my thesis committee: Prof. Hideaki Kuzuoka, Prof. Yasushi Nakauchi, Prof. Hirotaka Osawa, and Prof. Hiroya Igarashi, for serving as my committee members, and for their insightful comments and suggestions.

In particular, I am grateful to Prof. Hiromi Mochiyama, your inspiring lectures helped me to shape my research and to find my passion for haptic sensing. To Prof. Soichiro Matsuda, working together was a great and rich experience from me. Thank you for teaching me about the world of research in psychology. My gratitude is also for Prof. Masakazu Hirokawa for always believing in me, for the stimulating discussions and your endless patience. To all the participants who took time to join my experiments. Without you, this would not have been possible.

My special thanks are extended to the staff and members of AI lab, for your support and smiles during my years in this laboratory.

Contents

List of Figures

List of Tables

A mi padre, Mario Luis Núñez Montiel, quien desde niña me inculcó amor por nuestro universo y curiosidad por entender lo que nos rodea.
Siempre seguiré tus pasos.

A mi madre, Rosalba Morales de Núñez, quien desde mis primeras pruebas en la escuela primaria, hasta el día de la defensa de esta tesis recitaste "que Dios te de claridad, entendimiento, inteligencia, sabiduría y mucha calma para que todo salga bien". Tus palabras siempre están conmigo, gracias por siempre llevarme de la mano.

A mi hermana y mejor amiga Andreina Núñez Morales, quien es un ejemplo de perseverancia y sensibilidad.
Siempre estas ahí para recordarme sonreír.

A toda mi familia regada por el mundo que aun desde la distancia siempre me brindaron su amor y apoyo.

Este trabajo es por y para ustedes.

Chapter 1

Introduction

According to Aristotle, as human beings, we are ”social animals," and therefore, we are not self-sufficient, but we depend upon one another. As human beings, we have a profound necessity to connect with others and gain acceptance in social groups. It is through social interaction that we can build social relationships. Different studies support the idea that psychological well-being is related to the quality and quantity of people’s relationship [1] [2]. Moreover, feeling connected to others promotes physical health and well-being as well as emotional well-being [3]. However, a high degree of these positive effects comes from being able to coordinate during interaction with others, and accurately recognize and convey the intention to communicate. Among the different aspects of social interaction, those behavioral cues that help us to perceive and convey intention are related to the social signaling theory in communication.

Social signals are part of every human interaction. It is only by social signals that people can understand and predict others’ behavior. Our social signal capabilities are potent, and they viscerally impact not only on the person we are interacting with but also our psychology [4]. Social signals are independent of the language and content of the message. For example, when we watch a TV program in a different language, and we observe the characters engaged in a discussion, we can understand that there is a conflict even though we are not able to understand their words. The characters’ facial expressions and body language give us enough information to understand that they are arguing. More fundamentally, social signals are a combination of observable behavioral cues used to exchange of intention to communicate (such as speech utterances, body movement, gestures, manipulation of objects) [5] [6]. These behavioral cues work as “guide” or “indications” during the exchange in social interaction. A different definition describes social signals as “observable behaviors that produce, intentionally or not, tangible changes in others, whether this means to modify their inner state (e.g., to stimulate the emotions they experience), to modify their observable behavior (e.g., to make them laugh in response to a joke) or to change their beliefs about the social setting (e.g., to make them aware of conflict or disagreement).” [7]. This thesis will work with the following definition of social signals: “observable behavioral cues that are used during the exchange in social interaction and that produce, intentionally or not, tangible changes in others." However, what happens when we are not aware of or physically unable to perceive social signals? Also, how does social interaction work when the amount/kinds of social signals are limited to use?

A possible approach involves the use of computers to support the social aspects of the interaction. In the computer domain, there is a field of study called Social Signal Processing (SSP) which aims to model, analyze and synthesize social signals, allowing computers to understand SS and facilitate interaction with humans [8]. Studies in SSP mainly focus on non-verbal behaviors exchanged in face-to-face interaction, such as facial expressions, gestures or prosody. The majority of these studies aim to imitate or simulate the way we employ social signals in face-to-face scenarios, with the purpose of achieving a natural human-machine interaction. These approaches commonly require complicated systems and computational resources. However, in the scenario of human-human computer-mediated communication, is it necessary to simulate social signals used in face-to-face communication?

Considering that with the expansion and accessibility to the internet, social networks and mobile technologies, the opportunities for social interaction has mostly increased. For this reason, nowadays a considerable amount of social interaction is mediated by technology. These technologies mediated interaction scenarios sometimes requires the users to interact in new ways, different from face-to-face scenarios. Moreover, as human beings, we can quickly adapt to different interaction scenarios, and in those where social cues used in face-to-face interaction are not available or limited, we can observe new ways of communication that involve simple gestures or cues that convey social information. For example, the like-button used in many social network services, or the emoticons used to illustrate different emotions and feelings. Even if these gestures differ from those we use in face-to-face interaction, they have similar functions. It is observed that the number of cues available is limited compared to face-to-face interaction, but at the same time, new and alternative cues are exchanged. In these new technology-mediated scenarios, the first question that arises is what kind of technology should be used to mediated social signals? As well as, what are the social signals being exchanged and how?

1.1 Purpose of this research

The primary motivation of this research is to improve the quality of human communication, considering its close relationship to individuals well-being. Among the different aspects of social interaction, social signals were selected as they are observable cues and an essential component of human communication. Based on the assumption that it is possible to use computer-generated gestures as social signals, this research proposes a method to facilitate human-human social signaling using cues delivered by paired devices. These cues are meant to have a similar function as the social signals used in face-to-face scenarios. The proposed devices are physically embodied, and they encode the human behavior by touch channel, transfer action to a paired device, and decode it as a representation made by simple visual and haptic cues. With this, this research aims to facilitate social signaling in human-human communication using paired devices with visual and haptic cues.

The proposed approach was evaluated in two different communication scenarios with a population that might benefit from the characteristics of the proposed devices. One application is for remote communication scenarios, based on the assumption that the physical embodiment and touch-based cues will have an effect and improve the affective social signaling of individuals trying to communicate from different places. Another application involves, individuals with Autism Spectrum Disorder as they have difficulties to interpret social cues used in interaction, and a considerable amount of studies has proved the potential benefits of technology for training social skills.

After describing the scenarios and target population, the design and implementation of paired devices with visual and haptic cues are introduced. This includes system overview and description of the method to sense cues from the user, and to convey cues to the user. Then, each device is evaluated in terms of 1) Performance evaluation: to explore the capabilities of the devices for sensing and conveying cues, and 2) User study: to understand the role of the cues and their effect on the user. This new interaction setting comes together with different questions, especially considering that 1) "the behavior of the user" is different from "the behavior detected by the sensors" and 2) the decoded message representing a human behavior is made only with simple visual and haptic cues. Considering this, the following hypotheses will help us to understand if the research purpose is feasible or not.

1.2 Research hypotheses

As explained in the previous section, the purpose of this research is to facilitate social signaling in human-human communication using paired devices with visual and haptic cues. This research will explore this possibility by answering these three hypotheses:

- H1: Social signals can be represented by visual and haptic cues using paired devices for minimal communication. (Referring to the capability of paired devices for sensing/conveying signals)
- H2: Social signaling can be facilitated by the communication through visual and haptic cues. (Referring to the effect of visual and haptic cues on the user)
- H3: The effect of the represented social signals is conditioned by interaction factors: timing/order, interface and feedback. (Referring to the parameters that condition the effect of the signals)

Chapter 2

Literature review

Based on the terminology used to define the approach of this research, this section includes a description of the main features of the devices and the reasoning behind the selection of each feature. Previous works that describe the potential and/or limitations of these particular features are introduced.

2.1 Approach to facilitate Social Signaling

This research proposes an approach for facilitating social signaling using paired devices with visual and haptic cues. This section explores different works that share similar features with the approach of this research.

2.1.1 Minimal communication

Minimal communication is a term introduced on the work of [9] and refers to the transfer of event-type message driven by the user's actions. It was explored employing bidirectional I/O devices in shape of a button that sensed when the user pressed it, and sent a message to a paired button, representing this action of pressing by changes of colored lights. It was found that it is possible to transfer and perceive "intention" by communicating through simple cues. These cues represented the action of sending a message, made by the user or the partner, without any interpretation. One of the potential benefits is that the lack of content of the message emphasizes the context information However, in order for two users to be able to communicate it is necessary that:

- Both persons know the same rule.
- Both persons know the other knows the rule.
- Both persons know what happens on the other side of the line.
- Both persons believe that signals are made intentionally.

The proposed approach for facilitating social signaling is based on minimal communication, and its effect will be explored and discussed.

2.1.2 Physically embodied mediators

One of the features of the proposed approach is that the devices have a physical embodiment. An interface with a physical embodiment directly supports touch. Moreover, physically embodied and collocated interfaces have a stronger social presence, and can easily modify the user's environment. Physical presence has been shown to potentially have a positive effect on human-agent interaction, including a stronger perception of social presence. A physically present robot can affect a human's perception of the robot as a social partner, which can potentially lead to a more positive interaction [10].

Previous studies have explored mediators with embodiment that supported touch-based interaction, as well as the effects they had on human-human remote communication. A mobile and stationary hybrid communication system designed for long-distance relationships showed that a hybrid setup (a combination of a mobile phone and hardware) encouraged more frequent message exchange, which increased the feeling of closeness in users [11]. Similarly, a tangible interface was designed to remotely convey a sense of presence by combining mobile communication and touch-based interaction [12]. A similar concept introduced an embodied interface to increase the social awareness in remote communication [13]. These three devices used patterns of colored lights and touch-based interaction to convey a sense of presence to the partner.

Another study emphasized the positive effects of robot embodiment on the social aspect of human-human remote collaboration tasks mediated by a robot [14]. In the field of mediating communication with embodied interfaces, huggable devices are similarly found. A conceptual design intended to establish intimate communication remotely using the idea of transferring ambiguous information by hugging one device and making a paired one react [15]. Moreover, a teddy bear-like robotic device uses the same concept of enhancing traditional communication means through touch-based interaction, allowing users to exchange tangible expressions of affection, including hugs [16]. A different study found that it was possible to enhance physical co-presence in remote communication using a huggable interface [17].

2.1.3 Paired devices as mediator of interaction

In the literature, solutions for mediating human-human communication can involve a single device or paired devices. Studies using a single device commonly make efforts in designing interfaces that can sense human behavioral cues, and deliver back cues from the interface that results natural and acceptable for the user. In this group, one study explored different modalities of socio-feedback for a humanoid robot that worked as a social mediator in a conversation between two persons in collocated scenarios [18]. A different study investigated the social behavior of a humanoid robot mediator for encouraging co-actions among a group of children [19]. Similarly, a humanoid robot with human-like behavior was used to mediate questions among a group of people [20]. Considering that children with ASD have shown interest in robots, robots as a mediator of interaction are popular approaches [21, 22]. These studies made efforts in designing natural behaviors for a positive human-robot interaction. It was pointed out the importance of identifying human behavioral cues, representing an appropriated behavior on the robot, as well as being able to distinguish different users. However, approaches using a single robotic device for mediating human interaction are limited to collocated scenarios. Considering that nowadays a considerable amount of interaction occurs remotely through media, it is necessary to consider alternative solutions to use robotic devices as mediators of remote human-human interaction.

In the literature, the majority of the studies that employ robots to mediate remote human-human interaction exhibited a paired devices configuration. Paired devices are homologous technologies that share similar characteristics and affordances. Paired devices as social mediator commonly support connectivity, allowing two users to exchange messages between them. Studies using paired devices for mediating interaction implemented a more straightforward human-computer interaction and more implications in the relationship and communication between humans. For remote communication applications, it was found different implementations using physically embodied paired devices for increasing the feeling of closeness [11], sense of presence [12] or remote social presence [13]. These three devices identified behavioral cues through touch and delivered cues using patterns of colored lights and vibration. Different studies worked in mediating hugs remotely. A conceptual design intended to establish intimate communication remotely, using the idea of transferring ambiguous information by hugging one device and making the paired one react [15]. Moreover, a teddy bear-like paired devices were designed for communicating remotely using touch-based gestures, including hugs [16]. Paired dog-like robots were used as social mediators, and it was observed positive effects on the social aspect of human-human remote collaboration tasks mediated by the robots [14]. Paired devices have been used in collocated scenarios as well. A system of multiple agents was used to mediate the exchanges in a social game between children with autism and the therapist [23].

From this review, it was observed that a paired devices setup might be beneficial 'as it opens possibilities to explore both, collocated and remote communication scenarios. Supporting connectivity through paired devices is of special importance as by having two interfaces that share same affordances, users can easily imagine how the partner is using the device on the other side of the line.

2.1.4 Visual and haptic cues as social signals

The eyes are the body's most highly developed sensory organs. A large part of the brain function to process vision, and visual cues are a large source of information that helps us to figure how the world is perceived. These characteristics are beneficial for the approach of facilitating social signaling. Devices should increase awareness and help users to perceive each other through the cues. Moreover, physically embodied interfaces that deliver visual cues can easily modify the real-life environment, using all the benefits given by the interface's physical representation. On the other hand, touch plays a fundamental role in social interaction, and it is also a fundamental human need. It intensifies and complements information received through other modalities such as audio or visual. Information delivered through the touch channel could potentially enhance the role of the cues as social signals and intensify the effect on users using touch-based sensory information.

The role of the proposed devices includes: to encode haptic cues, transfer them to the paired device, and decode this message by representing it as visual and haptic cues on the paired device. These cues are meant to contain the users' intention and have similar roles as those of social signals. In the literature, a common cue identified by physically embodied paired devices is touch [11–13]. The role of body contact is considered to be important for expressing affection. It is possible to convey positive affection just by touches, hugs, and strokes, as well as negative ones such as by hitting or pushing [24] [25]. The field of affective haptics explores the integration of touch in computational systems. Among the different applications, affective haptics is related to the computational aspect of mediating human-human affective touch [26]. Moreover, design strategies for technologies to mediate intimacy and relatedness described the importance of supporting meaningful gestures that convey affection [27]. Among the various touch gestures, hugs are a very important part of human communication, as they can transfer comfort and give an emotional lift [26]. Following this, it was observed different approaches to mediate hugs by paired devices [15, 16].

These gestures are supposed to be represented on the paired device as well. After investigating the available paired devices in the literature, it was observed that simple cues were a common characteristic. This approach is advantageous to reduce the complexity of the hardware, which facilitates the possibility of being adopted by the user. Among the different visual modalities, simple cues with colored lights have been used to endow robots with expressions. On the one hand, dynamic colored lights have been used to convey a

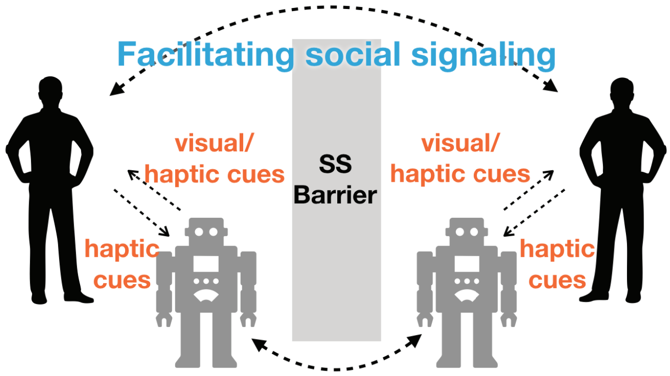


FIGURE 2.1: Model describing how to facilitate social signaling using paired devices with visual and haptic cues.

robot's states and actions [28], and simple expressions using colors were used to express the life duration of a companion robot [29]. Colored lights have been used in combination with sounds and vibrations as a simple and low-cost alternative to express a robot's emotional expressions [30]. Several models relate colors to different emotions [31]. One study proposed a methodology to express a robot's emotions by changing the color of its body [32], and it showed that different emotions were perceived in the agent when it displayed a certain color luminosity at a particular frequency. From this, it was understood that patterns of colored lights and vibrations could, even if limited, convey information to the user about the robot's state. However, in the context of mediating communication, the message or cues from the robot, are used to convey information to users about the actions of the other user. Different colored lights have been used to convey "user's emotion," similarly as traditional emoticons are used [12]. More fundamentally, colored lights have been as a presence indicator, as they are simply linked to the action of sending the message by touching the interface [11, 13], or by hugging it [33, 34].

2.2 Summary of the proposed approach

The proposed paired devices are designed to support minimal communication, which can be interpreted as "an event happened or not" type of message. The devices sense and encode the user's cues through touch channel. Then, this action is transferred as a message to the paired device. The message is decoded into visual and haptic cues that are used to represent the action of sending the message. Considering the devices are physically embodied, they directly support touch. Thus, touch was selected as the main medium of interaction. And based on the impact of visual cues represented on physically embodied interfaces, these are used as a way to represent the user's actions.

Based on the aforementioned reasons, this research involves the use of paired devices that support touch-based messages, by encoding the user's behavior through touch, and decode it using simple visual and haptic cues. It is assumed that an effect on human-human communication will be observed when two persons interact through paired devices that support these cues with the role of social signals. Figure 2.1 illustrates the proposed model for facilitating social signaling using paired devices with visual and haptic cues.

Chapter 3

Methodology

The methodology of this research involved: 1) Select target users that, based on literature review, could benefit from this approach, 2) Design and implement solutions involving paired devices for facilitating social signals.

3.1 Scenarios of use

The approach for facilitating social signaling using paired devices with visual and haptic cues will be evaluated in two different communication scenarios: remote interaction and collocated interaction. To understand the effect of the proposed approach, target populations that might benefit from it were selected. The first scenario involves mediating remote interaction through paired devices, and it is going to be evaluated with individuals with close relationship trying to communicate from different places. Similar to the role of emoticons in remote communication, the approach involves designing social signals by paired devices. Considering that traditional communication media support the transfer of visual and auditory cues, it might be beneficial for this target user to have additional information through the tangible channel. For collocated paired devices-mediated interaction was evaluated with individuals with Autism Spectrum Disorder (ASD). They need training to acquire skills used in social interaction. Considering their natural interest in visually engaging toys, they might benefit by the approach of social signaling with paired devices.

3.1.1 Social signals and remote communication

The differences between computer-mediated and Face-to-Face communication have been extensively explored and analyzed [35] [36]. Traditional computer-mediated communication conveys affective communication to a certain extent. Facial expressions, body gestures, and voice variations are transmitted with video. Even text can support affective communication via emoticons. The greatest difference between computer-mediated and Face-to-Face communication is indeed the absence of body contact [37]. This decreases affective and social presence. Even when messages conveyed by computer-mediated are full of emotional content, they are reported to be less intense [38]. Therefore, augmenting traditional communication with mediators for physical contact is an effective strategy to transmit positive messages such as hugs, or strokes; as well as negative ones like hitting or pushing [39] [24] [25].

Typically, technologies that support touch-based gestures are designed as wearable devices, tangible objects or robotic devices [40]. Touch-based interaction with these devices has positive effects on users [38]. They provide mental stress relief [41], and even simple objects or interfaces can prompt responses similar to those that other people would elicit [42]. Different interfaces supporting remote touch have been designed to convey presence and improve connectedness over distance [13]. Among them, studies have made effort on proposing different designs of mediators of touch, evaluated the general impressions [12], and proposed design guidelines regarding shape and material [34] Mediators of social touch might have different shapes, but all of them share a common feature: to encode the human intention to communicate by touch, and to decode this message using different combination of cues. It is therefore necessary to understand which functions the interface should support, and how to use the cues to convey the information to the user.

A tangible interface was designed to remotely convey a sense of presence by combining mobile communication and touch-based interaction [12], and the users' perception was evaluated. Among the different tangible gestures, hugs contain a strong affective connotation. In the field of mediating communication with physically embodied interfaces, huggable devices are similarly found. A teddy bear-like robotic device uses the same concept of enhancing traditional communication means through touch-based interaction, allowing users to exchange tangible expressions of affection, including hugs [16]. Hug over a distance combined a huggable interface with a wearable device, allowing users to convey hugs remotely [43]. On the other hand, researchers have made efforts in exploring the design requirements of this type of interfaces. A design of pillows for ambient-based interaction integrated with connectivity was proposed for enhancing social intimacy among people [33]. Similarly, a set of interactive pillows for conveying affection were designed focusing on the appearance and textiles to make them easily adopted by the user [34].

Previous studies have explored mediators with embodiments that support touch-based interaction, as well as the effects they had on human-human remote communication. A mobile and stationary hybrid communication system designed for long-distance relationships showed that a hybrid setup (a combination of a mobile phone and hardware) encouraged more frequent message exchange, which increased the feeling of closeness in users [11]. A similar concept introduced an embodied interface to increase the social awareness in remote communication [13]. These three devices used patterns of colored lights and touch-based interaction to convey a sense of presence to the partner. A different study found that it. However, the interface did not support connectivity and it was used to encapsulate the phone. From this studies it was observed that the interaction with the interface is simple, but the mediated interaction between two persons is complex. Especially considering the meaning of hug and the potential effect on the user, it is necessary to explore the effect of a huggable mediator that support connectivity on the users' communication.

3.1.2 Social signals and individuals with developmental disorders

Play is one of the ways children learn and develop skills to engage and interact with the world around them [44, 45]. However, children diagnosed with Autism Spectrum Disorder (ASD) have shown impairments in play skills [46–48]. They are known to have difficulties in understanding social cues that are necessary to cooperate and interact with others during play, and usually opt to isolate themselves from others. This impairment is reflected in their motor skills, in the way they manipulate toys, and in social and pretend play [49, 50].

Different studies have made efforts to develop information and communications technology (ICT)-based solutions for interventions for children with ASD [51]. Specifically, robotic toys can potentially be used as social mediators, engaging children with ASD in social play styles that involve others [52]. As opposed to the use of only social cues, toys with sensory feedback, like flashing lights or sounds, might elicit more play behaviors [53]. Robotic toys can be defined as tangible devices with embedded sensors and actuators that can perceive certain aspects of the environment, such as the behavior of children, and automatically react to them.

In general, robot-based interventions are carefully designed to elicit, facilitate, or train a particular behavior, by adapting the hardware and interaction rules according to the goal of the intervention. Among the different interventions designed to address social skills, some specifically target the training of turn-taking skills. Turn-taking is a back and forward exchange between two or more individuals, and it is considered a fundamental skill for different aspects of social interaction such as being patient, being able to recognize others' intentions, and collaborating with others [54]. Through turn-taking children are able for example, to play and let others play or to talk and listen others talking in a conversation. Considering that children with ASD has significant impairments in turn-taking behaviors, different approaches explore improving this condition [55–57]. Moreover, studies discussed the possibility of reducing the severity of ASD and impairments in social communication, by training and improving turn-taking behaviors [58, 59]. To achieve turn-taking, individuals have to be able to switch the initiative during interaction spontaneously. In conventional turn-taking interventions, the therapist supports this back and forward structure by responding to the child in a specific way according to the rules of the activity [56]. Robotic toys can be used as tools for the therapist to: (1) Provide cues that engage children in the activity, and (2) Describe children's behavior during the activity.

3.1.2.1 Robotic Toys to Engage Children in Interaction

The use of robotic toys to facilitate turn-taking training has been explored, and positive effects have been observed [21, 22, 60, 61]. Robotic toys can respond to children's behavior and provide repetitive and engaging stimuli to guide children through turn-taking play. These studies share a common aspect, which is having a single robotic toy interacting with the child. The role of the robot is to mediate the interaction between the child and others [21, 22], the robot taking turns with a child to chase each other [60], or the robot as the child's partner in an activity mediated by the therapist [61].

However, one robot can provide contingent feedback to one person at a time. On the other hand, a different approach using more than one robotic devices to mediate turn-taking between the child and the therapist has been proposed. Multiple robotic devices can deliver contingent feedback due to the actions from both, the child and the therapist. Moreover, guiding the participant's attention to different locations/objects can be easily achieved with multiple devices. These characteristics are advantageous in terms of facilitating turn-taking during play activities. A multi-agent platform with simple visual was developed cues in the form of blinking lights in a turn-taking intervention [23, 62]. They assigned meaning to the colors (e.g., blue blocks means water) and simulated the interaction among different blocks (e.g., when a block representing an animal is thirsty, it has to be placed next to the water block. The changes of states are represented by blinking frequency or intensity).

3.1.2.2 Robotic Toys to Describe Children's Behaviors

Considering that the treatment of autism is adapted to each child's needs, it is important then to evaluate the interventions objectively. Traditionally, the evaluation is performed based on observations and post analysis by video coding, with which the therapist identifies those relevant behaviors according to the purpose of the intervention

[55, 56]. This is not only a demanding task, but it can also lead to different conclusions based on the therapist's criteria. In this context, technology can work to provide therapists with quantitative data related to the child's performance during the intervention [63]. This can potentially facilitate the adaptation of the activity to stimulate the child's learning process adequately.

In this field, studies have explored the use of computer vision approaches to automatically collect and describe children's behavior, thereby reducing the human effort required in analyzing the interventions [64]. Computer vision was used to analyze and assess autism, and it was used a robot to engage with children in face to face interactive scenarios [65]. However, solutions using computer vision have limitations such as higher costs, occlusion, and fixed setups. To overcome them, different studies explored the use of wearable devices to describe children's behavior. A combination of three accelerometers has been explored to detect stereotypical movements related to ASD [66]. Wearable devices were used to detect children's smiles as an indicator of enjoyment and positive disposition [67], or to sense the orientation of the facial region, to quantitatively understand children's focus of attention [68]. A framework that combines a robot as engaging element together with wearable sensors and computer vision was used to analyze children performance during the therapy, and based on this personalize the content of the lessons [69].

One of the limitations of the use of wearable devices is the dependence on the child's disposition to wear them. Moreover, wearing something unusual might have an additional effect on their behavior. This is one of the motivations for the approaches using sensors embedded in familiar objects and toys. Robotic toys can follow the child's performance based on the way in which the device is manipulated during play. One study used embedded sensors in rattles to identify early signs of autism [70]. Blocks-like toys were designed to detect developmental delays from patterns of motion while manipulating the toys, time to accomplish a task and accuracy [71]. Toys made with smart textiles were designed to sense and monitor child-toy interaction during playtime [72]. These works identified the benefits of the information extracted from the manipulated toys and proposed alternatives to analyze the sensor's data for describing children's behavior.

3.1.3 Summary and need of this research

In the literature, we found different studies that employed paired devices to mediate human-human communication. Among them, it was investigated those used in applications for mediating remote communication, and for interventions for individuals with ASD.

Applications for remote communication included a variety of designs of paired devices with similar features. These works aim to include touch-based messages in remote communication, considering that traditional media does not commonly support it. It was pointed out not only the benefit of mediating social touch [11] [13] but also the possibility of communicating remotely by relying on simple and abstract cues. These works mostly focused on evaluating the design of the interface in order to point out requirement [16] [33] [34]. Among the different paired devices for remote communication, it was explored those that support hug-based communication. It was observed the necessity of evaluating these technologies in terms of the effect they have on the user. In this context, a method to evaluate the effect of a huggable interface in remote communication pointed out the positive effects of hugging while communicating remotely [17]. However, the interface did not support connectivity; therefore this emphasizes on the necessity of evaluating the effect of paired devices for mediating hug-based communication remotely.

Applications that target developmental disorders such as ASD included interfaces that mediate the exchange of turns. These interfaces are used to provide cues and instruction for children with ASD and thus, potentially facilitate the interaction with others. Most of the approaches involved a single interface, and they were used to engage children in interaction or to describe their behavior in order to follow their progress. Among them, it was found one study that targeted children with ASD and used paired devices in turn-taking intervention [23, 62]. These toys delivered visual cues using colors, and to be able to play with them, the participant needed to understand the abstracted meaning of the color patterns and the relationship between them. The results showed that children were able to understand metaphoric meanings through abstract and simple geometrical objects, and the interactive toys were appealing for the participants. However, it was reported as well that the user group should have the ability to understand metaphoric meaning in order to understand the game. Considering that children with ASD have impairments in understanding metaphors and figurative language [73, 74], the population of children who can participate in the game is limited. Additionally, the turn-taking behavior of participants was investigated on the context of the designed game scenario; hence the effect of sensory feedback policy on turn-taking behavior independent from the play context has not been discussed.

3.2 Implementation

This section includes description of the proposed solutions for facilitating social signaling using paired devices with visual and haptic cues.

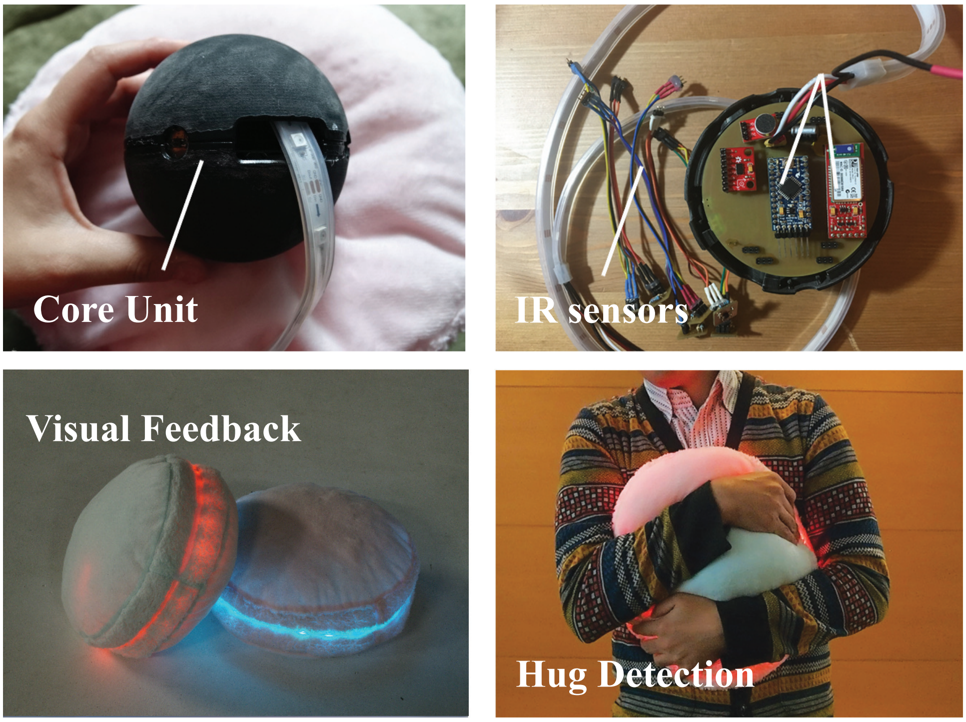


FIGURE 3.1: Top left: Plastic core placed inside the cushion, Top right: Circuit board and sensors, Bottom left: Macaron displaying visual cues, Bottom Right: Macaron detects the user's hugs.

3.2.1 Paired devices for mediating social signals in remote communication: Macaron

Macaron has a simple and round appearance. One of the requirements was to make it look like a product that can be commonly found at home (Figure 3.1 Bottom-left). It is broadly composed of three elements: sensing, feedback, and communication [75]. The sensing part involves the design of a sensor able to distinguish hugs; the feedback is made with LED and vibration patterns, and the communication is managed by a server connected via Bluetooth with Macaron.

The sensor's design was proposed by [76] and it is composed of a plastic core that contains an array of photo reflective sensors used to sense the deformation of a soft interface in which the core is embedded (Figure 3.1 Top-Left). Macaron's core was designed using Solidworks and 3D printed, and it consisted of a plastic sphere with a diameter of 8.5cm. Appendix B contains the schematic of Macaron's case. A cushion with shape of macaron was modified by emptying it and then re-filling it using the proposed granulated filling material [76] necessary for facilitating the detection of the cushion deformation. For this implementation, six photo reflective sensors were embedded and distributed around the plastic sphere. The circuit is contained inside: a board with an Arduino mini pro 5V, a SparkFun Bluetooth module, a vibration motor, a socket for the LED strap

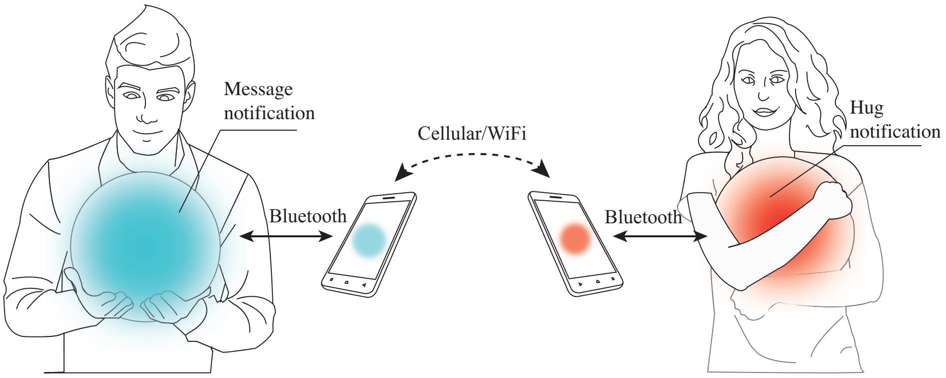


FIGURE 3.2: Illustration of Macaron's concept: two persons living on different places can communicate affective messages mediated by Macaron.

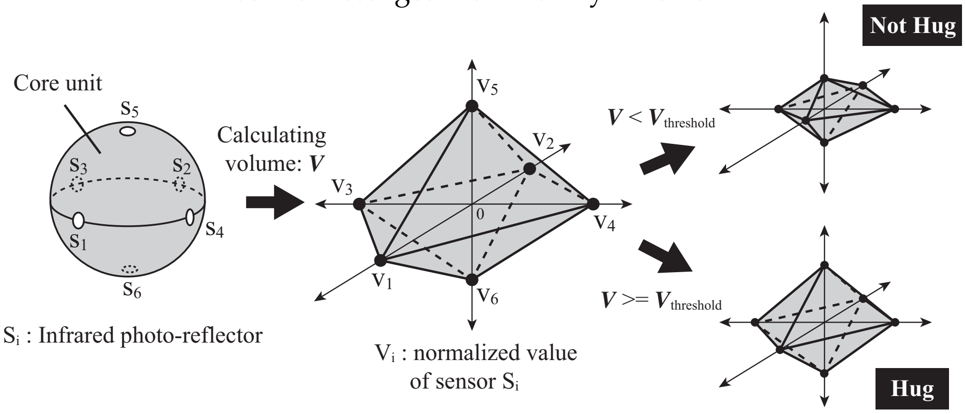


FIGURE 3.3: Hug detection algorithm based on infrared photo reflective sensors.

(1 meter of Adafruit's Neopixel, 30 LEDs), a charging circuit with a lithium battery (1400mAh), a 3.3V regulator, an ON/OFF switch, and a micro USB connector for charging the battery (Figure 3.1 Top-Right). The schematic of the Macaron main board and sensors boards can be found in Appendix A. The plastic case has an access hole for the USB connector (for battery charging) and the switch (Figure 3.1 Top-Left), and from which the LED strap connects to its socket. The LED strap is placed around the circumference of the cushion (Figure 3.1 Bottom-Left).

The communication is managed by Bluetooth using a server client configuration. This allows to connect multiple clients (Macaron's modules), and a server in a notebook computer manages the communication among them. Figure 3.2 illustrates a scenario with two users geographically separated, using Macaron to communicate affective messages driven by hugs.

3.2.1.1 Hug detection method

Each photo reflective sensor in the core unit is represented as a point in the software (Figure 3.3 Left). The point location ranges from the

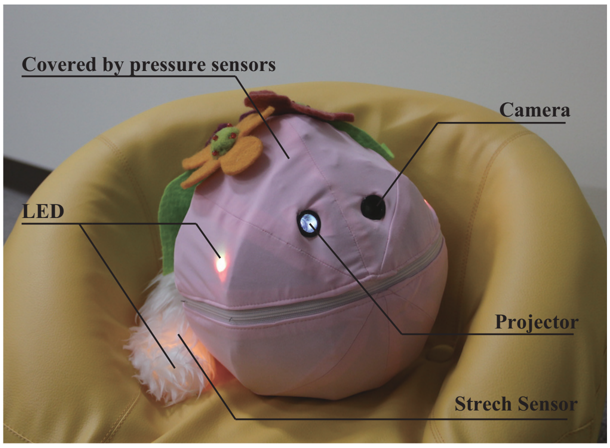


FIGURE 3.4: Pepita: a huggable robot companion with caricatured appearance.

normalized value of the sensor reading. The combination of the six points creates two rectangular pyramids that share the same base. Using Figure 3.3 (center) as a reference, one pyramid is V1V3V2V4V5, and the other is V1V3V2V4V6. When the cushion is deformed the sensor values (ranged from 0 to 1023) increase, thus the distance from the origin of each point, increases. From this, it can be understood that bigger volumes will result not only based on the amount of deformation detected by one sensor, but also by the combination of the points.

Hugs involve embracing the pillow with both arms against the body (Figure 3.1 Bottom-Right). It is expected that hugs will generate polygons with bigger volumes as the deformation is distributed on different sections of the cushion, and measured by different sensors (Figure 3.3 Right). The volume of the polygon in total is the sum of the volume of the two pyramids. To reduce the effects of spike noise, we applied a median filter to the volume value. Also, since hugging involves embracing over a significant interval of time (at least 1 s), we tried to minimize spike noise by adding a condition that removes instant peaks. Then, the implemented hug detection condition was the filtered volume values that go over a selected threshold, keeping the value over the threshold for more than 500 ms.

3.2.2 Paired devices for mediating social signals in remote communication: Pepita

Pepita is a caricatured robotic device (Figure 3.4) designed as a social companion. As shown in Figure 3.5, the system consists of three main components: the robot circuit, a smartphone connected to a projector, and an external computer for remote control, each of them managed by a different algorithm [77].

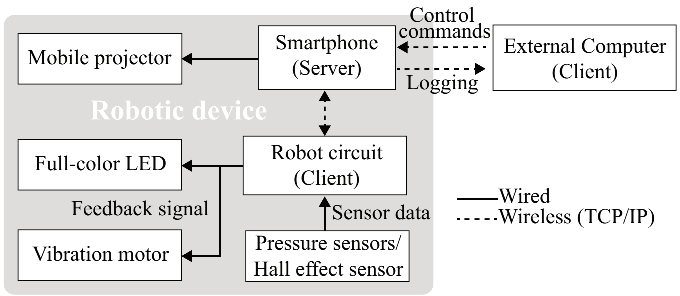


FIGURE 3.5: Device components: composed by a smartphone that communicates with the robot's circuit and an external computer.

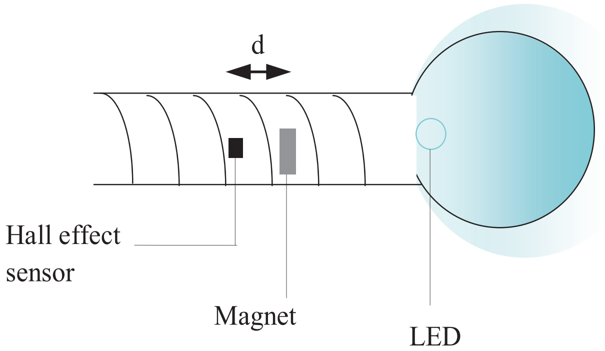


FIGURE 3.6: Stretch sensor structure: variations in the distance between the hall effect sensor and the magnet are used to detect stretching.

The circuit is composed of a microprocessor (Arduino UNO) connected to the smartphone and the external computer by a Wi-Fi module (Seeed Studio, V1.0). A printed circuit board was attached on top of the Arduino and the WiFi shield. Alongside the sensors, the circuit contains three full-color LEDs and one vibration motor. The LEDs were placed on the sides to look like the cheeks and in the tail tip to emphasize the caricatures' features. The vibration motor was attached to the bottom of the plastic case. The robotic device has two types of sensors to detect tangible gestures: pressure sensors covering the robot's body, and a stretch sensor in the tail. The pressure sensors were made to endow the robotic device with hug detection. It is made of 5 mm thick conductive foam (Seiren Electronics Co., Tokyo, Japan) divided into eight petal-shaped pieces covering an 18 cm diameter plastic sphere. Handmade electrodes were made using copper sheets with conductive tape (Seiren Electronics Co., Tokyo, Japan) and attached to the foam pieces. The wires were attached to the copper sheet, and each piece of the sensor was connected in parallel with a 15 Kohm resistor. The stretch sensor was made following the approach from [78] that uses the properties and structure of the material for sensing gestures. A silicone tail containing a magnet, together with a linear Hall effect sensor (Figure 3.6) was used

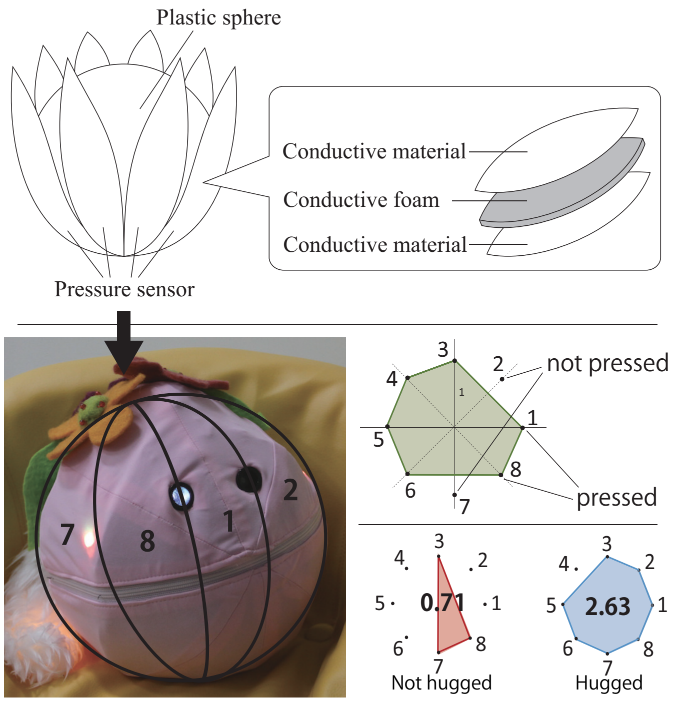


FIGURE 3.7: (Top) Pressure sensor's arrangement: each petal-shaped sensor was made of soft conductive foam and attached around the spherical body of the robot. (Bottom) Hug detection method: when a piece of the sensor is pressed with certain applied pressure, it generates a point placed in a fixed position on the Cartesian plane. When the area of the polygon formed by the generated points exceeds a threshold, the device recognizes the action as a hug.

to detect when the tail was stretched. Silicon was poured into a cast of half of the tail (longitudinal section), and the wires were introduced while the silicon was still soft. After it dried, the other half was poured covering the wires. To prevent the wires from snapping, they were coiled and placed inside the cast before pouring in the silicon. When the tail is pulled, the distance between the magnet and sensor changes, which is detected as changes in the magnetic field.

The second part comprises a smartphone (Galaxy Nexus, Samsung electronics Co., Suwon, South Korea) which is connected to a pico-projector (EAD-R10, Samsung electronics Co., Suwon, South Korea), making it possible to display the screen content on any surface. The projector was placed in one eye, and the smartphone's camera in the other. An Android application manages the content of the screen, which changes according to the interaction with the device. The smartphone works as a server, receiving the sensor data and sending it to the external computer.

The last component of the system is an external computer, which is used to visualize and store the sensor data. All of the elements are connected using TCP/IP sockets. The smartphone works as the server, mediating the communication between the microcontroller and external computer. By doing this, it is possible to send commands from the computer to the device remotely (for example, to start collecting data from the sensors and visualize it on the computer screen).

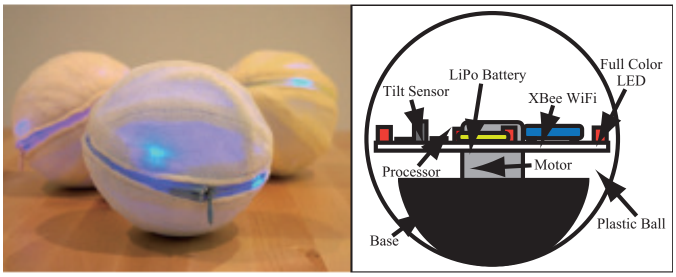
3.2.2.1 Hug detection method

We used the combined readings from all the sensor pieces to implement hug detection on the robot. Similar to the stretch sensor in the tail, this sensor works with the structure to detect hugging. This approach is advantageous because it makes it possible to cover the robot body with sensors using a minimum amount of wiring.

The sensor for hug detection is divided into eight different pieces, and each piece is fixed on a particular position (Figure 3.7 Top). Each of these pieces is represented in the code as a point, and the eight points are used to generate a polygon. The generated polygon's area is used to determine when the robot is being hugged. When any of the sensor pieces is pressed, the value drops to a certain level, and if this value goes below to a set threshold, a point for the polygon is generated. For the threshold, we chose 500 (ADC) value. This value will differ according to the deformation and size of the pieces, which makes it important to tune it. Each of the generated points is placed in a fixed position in the Cartesian plane, separated from the origin by a segment of a fixed value of one (Figure 3.7 Bottom), and the polygon's area is calculated using the following equation:

$$S = \frac{1}{2} \sum_{k=1}^{n} (X_k - X_{k+1})\,(Y_k - Y_{k+1}), \tag{3.1}$$

where $n = 8$, and X and Y refer to the coordinates of each point. With this approach, polygons with bigger areas result when more points are generated. Considering the action of hugging involves embracing the device with both arms, most of the sensor pieces are expected to be pressed when the robot is hugged. For this reason, hugs will generate a larger polygon by activating more points, compared, for example, with the action of pressing the robot with both hands (Figure 3.7 Bottom). Hug detection is made possible by reading the size of the polygon generated when the user manipulates the device. When the area reaches the selected threshold, a hug is detected. With this approach, it is expected to increase the sensing accuracy in distinguishing hugs from other types of manipulations.




```
17:32:35 CHANGE   103 RED
17:32:41 RECEIVE  104 RED
17:32:41 CHANGE   104 GREEN
17:32:46 RECEIVE  104 GREEN
17:32:47 RECEIVE  103 GREEN
17:32:47 CHANGE   103 BLUE
17:32:53 RECEIVE  104 BLUE
17:32:53 CHANGE   104 RED
17:33:01 RECEIVE  104 RED
17:33:05 RECEIVE  104 RED
17:33:06 RECEIVE  103 RED
17:33:06 CHANGE   103 GREEN
17:33:11 RECEIVE  103 GREEN
```

FIGURE 3.8: (**left**) Cololo: the spherical device to communicate presence, (**center**) Hardware overview, (**right**) Data collected by Cololo

3.2.3 Paired devices for mediating social signals in interventions for children with ASD: Cololo

Cololo is a spherical robotic device developed for enhancing remote communication by conveying abstract messages using colored lights (Figure 3.8 (Left) [9]). As shown in Figure 3.8 (Middle), each Cololo device is a 10 cm diameter spherical case that contains: a tilt sensor using a ball switch, an XBee WiFi module, full-color LEDs, an electric motor, and a microcontroller. A counter mass is attached to the rolling axis of the motor, enabling Cololo to “wiggle” when the motor is powered. When a user makes contact with a Cololo device and it is detected by the tilt sensor (manipulation), the manipulated Cololo sends a message to an external server through wireless communication. Then, the external server transfers the message to the paired Cololo and it is represented by colored lights and wiggling movements. The message consists of the ID of the sender with its LED color, and all messages are recorded with time stamps as shown in Figure 3.8 (Right).

Cololo holds several advantages when used as an assistive device for children with ASD [79]. The spherical shape with no sharp corners provides a safe, soft, and friendly appearance that is advantageous when the device is being manipulated [80]. Studies on activities with children with ASD using a ball-shaped robot [81] showed the potential of these robotic toys as an attractive element in the intervention. In addition, the visual cues in the form of colored lights and wiggling movements stimulate the child, and the ability to communicate with the paired device provides a simple way to include others in the game. Based on these considerations, we used Cololo devices for the intervention aimed at facilitating the exchange of intention in play activities that involve turn-taking.

Chapter 4

Evaluation

This section contains the evaluation of each solutions designed for the applications described on Section 3.1. Each study was meant to contribute on the understanding of how the proposed paired devices can facilitate social signaling.

4.1 Social signaling for remote communication: Macaron

Section 3.2.1 introduced the hardware specifications of Macaron, as well as the proposed hug detection approach. This section includes the evaluation of Macaron in terms of 1) hug detection performance, 2) effect of the communication through Macaron. To understand the effect of communication through Macaron, we have divided the evaluation into two separated studies. The first one is a pilot study to investigate the feedback perception. Then, an experimental study explored the effect of mediating social touch through a physically embodied interface, by comparing it with a virtual representation of Macaron.

4.1.1 Hug detection performance

The purpose of this experiment was to test the accuracy of the system detecting hugs among different tangible gestures. Initially, we identified gestures commonly used when manipulating a cushion, and those that could be easily wrongly detected as a hug. The following were the gestures tested in this experiment: *Hug, Press on the right and left sides, Press on the upper and lower areas, Hold it in your arms, and Put it on the table and rest your head*.

12 participants joined the experiment (5 female, 7 male, average age 29.4). They were asked to sit in front of a desk looking toward a wall. On the desk, we set a computer used to display the instructions. A camera placed on the side recorded the complete session. Before starting the test, the participants were explained that the purpose of the experiment is to record interaction data and they were asked to follow the instructions that will appear automatically on the screen. The instructions were written in both, English and Japanese.

TABLE 4.1: Results of the hug detection experiment

Precision	Recall	Accuracy
0.81	0.79	0.79

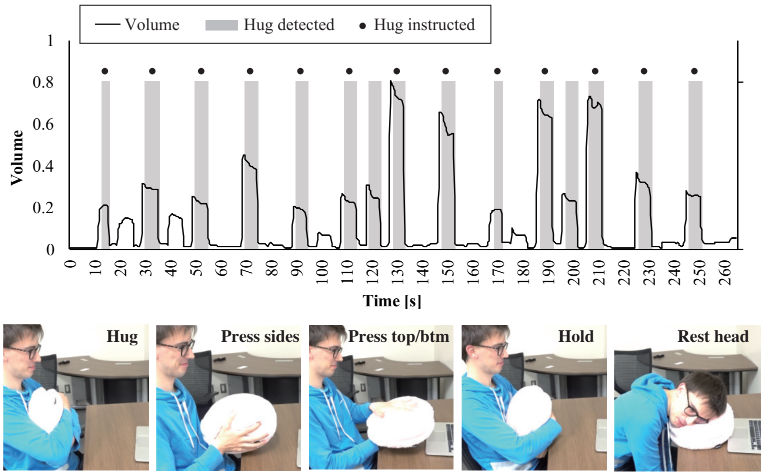


FIGURE 4.1: Data from one participant during one session to test hug detection.

Each instruction was programmed to be displayed for 7s followed by a "release" instruction for 3s. Participants were asked to place Macaron either on the table or their legs during the release instruction. They familiarized with the instructions before starting the experiment; however, they were not instructed how to make the gestures. The idea is to test the participant's criteria when performing the gestures. Moreover, the device did not deliver any feedback during the experiment to avoid instructing the participants. The plastic case was taken out the cushion, rotated and reinserted at the beginning of each session to avoid position dependence. Each participant performed a total of 25 gestures, a hug alternated with one of the other four gestures.

4.1.1.1 Results

In total, we analyzed 156 hugs predictions among 300 cases (156 hugs and 144 other gestures). We tested the detection performance using different thresholds and selected a threshold (volume value) of 0.15. Table 4.1 includes the analysis of precision and recall using the selected threshold. Figure 4.1 shows the data collected from one participant during one session. The plot represents the volume of the polygon for each gesture. The gray areas represent hugs detected by the system, and the dots represent instructed hugs. In this particular session, the system detected two false positives, and both were related to pressing top and bottom.

To investigate the variability of a hug, we avoid verbally instructing the participants about the way to perform the gestures, and we did not include feedback from Macaron to indicate when the hug was detected. As shown in the table 4.1, a threshold of 0.15 had a good performance for this application. We aimed to obtain a highly responsive system by detecting hugs when the participants were doing so, but at the same time, a higher recall value is related to the ability to distinguish hugs from other gestures. We considered that a precision value of 0.8 (or higher), is appropriated for this application.

4.1.2 Pilot study: feedback perception

Beside identifying hug-based expressions, Macaron delivers cues using patterns of colored lights and vibrations. These patterns should contain enough information to allow users to perceive the intent of the partner. In the literature, a study found that automatic communication using emoticons results in less perceived intimacy than user-initiated communication [82] Moreover, a different study verified that user-initiated messages using colored lights were perceived more human and intentional than automatic feedback condition [9]. For this reason, It is necessary to design an experiment that provides the data required to investigate if users can transmit and receive intention when communicating with someone else remotely.

4.1.2.1 Purpose

Following same method of [9], this pilot study aims to verify if the feedback delivered by Macaron based on the implemented interaction rule, allows users to perceive the intent of the communication partner.

4.1.2.2 Materials

The system used in this experiment was composed of one Macaron module, one computer running the server, a video camera to record the session and a set of printed questionnaires. A graphic interface was used on the computer by the experimenter to manage: data collection, setting the feedback condition and simulate a paired device to communicate with the participant.

4.1.2.3 Macaron's interaction rule

Figure 4.2 illustrates Macaron's interaction rule, which is designed for paired devices (A and B). Each of these devices support hug detection (Hug timeline) and delivers visual feedback with colored lights as well as vibration patterns (LED timeline). To simplify the information, the vibration patterns were constant for all conditions, and only the colors were used to define the meaning of the cue. Macaron was designed to support synchronous and asynchronous communication driven by hugs. The implementation of the interaction conditions

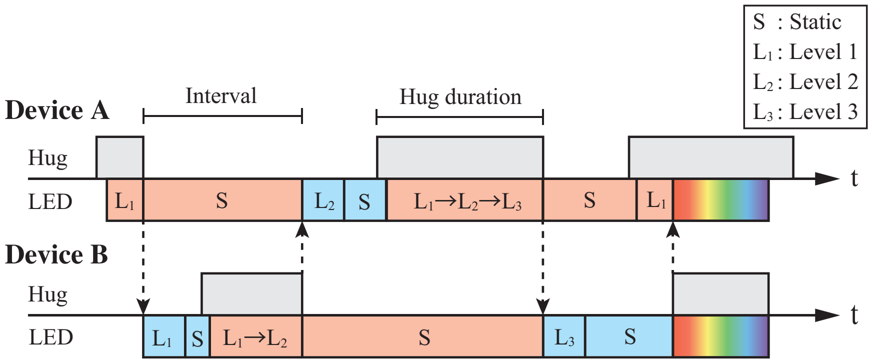


FIGURE 4.2: Interaction rule of two participants using Macaron to communicate.

needs to be carefully designed considering the nature of hugging. For example, different from pressing a button, the action of hugging is not instantaneous, but it involves embracing with both arms for a particular time interval. Based on this:

- Asynchronous condition: hug detection is represented by red-colored lights, and message notification by blue-colored lights. This condition contains three levels of hugs, and the level increases the longer the user keeps hugging the pillow without releasing it. Levels are represented by the blinking frequency of the LED's; thus, when a user is hugging Macaron, it is possible to visualize how the level increases before he or she releases it. The level value is transferred within the message, and it is represented by blue blinking patterns on the paired device. These blue lights have the frequency indicated by the level value received in the message. When the user finishes hugging and releases the pillow, the LED's stop blinking and become static red. Similarly, when the blinking pattern of the message notification finishes Macaron becomes static blue. Changes in the static colors are used especially for those cases when the two users intend to communicate at different times. By this, users can visualize the partner's message notification even if they are not looking at the pillow when the message is received.

- Synchronous condition: it was implemented to visually represent those moments when users try to communicate simultaneously (sync hugs). The feedback was made by a multicolored LED animation. Every time a user is hugging Macaron (while the red blinking is being displayed), and a message from another user is received, the lights pattern changes to a multicolor pattern in both devices. With this, a user can realize the moments the partner is hugging at the same time.

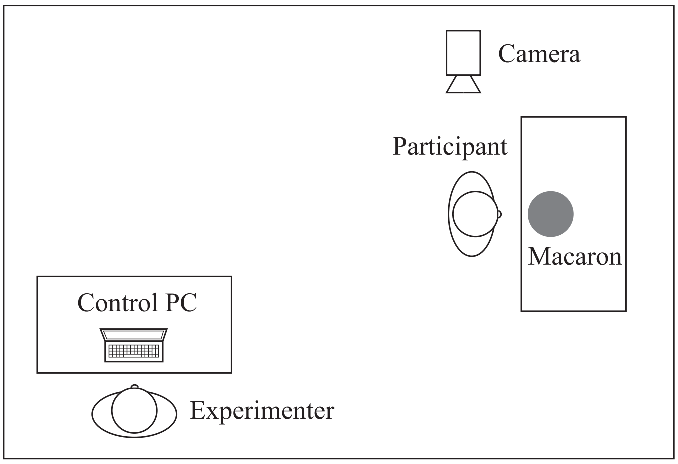


FIGURE 4.3: Layout of the experiment room: participants were facing a wall while interacting with Macaron. The experimenter communicated with Macaron using a computer from the back side of the room.

4.1.2.4 Flow of the experiment

3 participants (female, average age 25) joined this experiment. All the participants had previous experiences of communicating with a person with a close relationship (e.g., family, friends or partner) while living geographically separated. The system used during this experiment was composed of one Macaron module, one computer running the server, a video camera to record the session and a set of printed questionnaires. A graphic interface was used on the computer to manage: data collection, setting the feedback condition and simulate a paired device to communicate with the participant. Each session was video recorded, and it was carried out by two experimenters. Participants were asked to sit in front of a desk and facing a wall during the tasks (Figure 4.3. One experimenter introduced two Macaron (switched off) followed by an instruction video that explained how Macaron works, and the meaning of the LED patterns (the relationship of the actions and the visual cues). The video showed actors in a home environment. With this, it was intended to set Macaron in a scenario of use and make the participant understand the concept without explicitly being explained. Then the participants were told that a person (without describing gender or age) was going to use one of the two Macaron to communicate from another room in the same building. After telling this, one of the experimenters went out the room carrying one of the Macaron modules, and the participant received the other. This step was essential, as it was necessary to make the participant believe there was always a human partner communicating through Macaron. We tested the participant's perception using three types of communication partners (feedback condition), and each condition lasted 4 minutes. The order of the conditions was: 1) Manual, 2) Echo, and 3) Random. The description of the behavior of each condition was as follow:

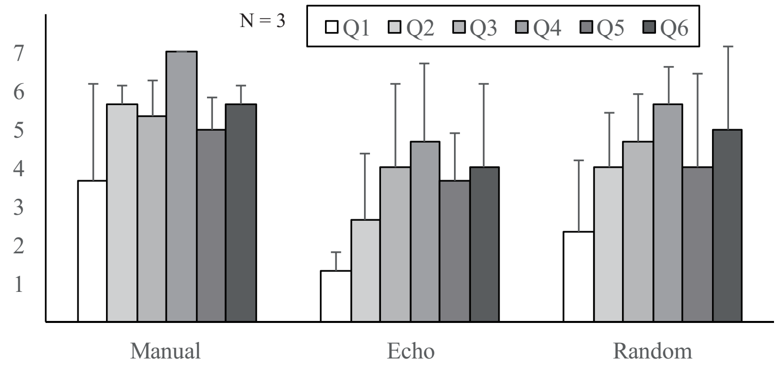


FIGURE 4.4: Data from the questionnaires applied after each feedback condition.

1. Manual: a human freely communicates through Macaron (one experimenter)

2. Echo: the server echoes back a message every time the participant hugs. The time delay in sending the message is 2s, and the message notification is always set on level 1.

3. Random: the server sends a message by randomly selecting a time interval between one message and the next (randomly selection from 1s to 30s). The level of the message notification is chosen randomly as well.

All the conditions were managed by a computer. Echo and Random conditions were preprogrammed, and the manual condition was controlled out by one experimenter (without the participant being aware of it). After each task, participants answered a questionnaire with six items. The first one was Inclusion of Others in Self (IOS), which was proposed as part of a method to evaluate the effect of affective technologies [82]. It is a single item pictorial measure which asks them to rate using a 7 points scale Q1: "Please circle the picture that best describes your current relationship with your communication partner". The next three questions tried to assess how the participant perceived the method of interaction [82]. "What do you think of the method of interaction? The method of interaction was..." and then participants answered using a 7 points scale Q2: from very cold to very warm, Q3: from unsociable to sociable, and Q4: from impersonal to personal. The last two questions were related to the perceived intentionality and humanity, and participants answered using a 7-points scale Q5: How much did you perceive the partner's intent? From not intentional to intentional, and Q6: How much did you perceive the partner's humanity? From machine-like to human-like. After finishing the three conditions, participants were interviewed to obtain more impressions about their experience.

4.1.2.5 Results

The results on Figure 4.4 show the average of the scores from the three participants for each item on each condition. The whiskers represent standard deviation. Further interpretation of these results, as well as the comments collected from the interviews, are included in the discussions section.

4.1.3 Effect of mediating social touch by a physically embodied interface

In the previous experiment, we could verify that participants could communicate through the proposed interaction rule. This leads to the next step which involves the use of Macaron on applications for communication to point out its effect on affective social signaling.

4.1.3.1 Purpose

In remote communication context, the tangible channel is underrepresented in traditional media. Based on the potential benefits of mediating social touch [38], we aim to understand how hug-based messages can be beneficial to convey affective cues remotely. The purpose of this experiment is to investigate the effect of the embodiment representation on participants' perception.

With this comparative study we expect to point out which aspect of the two interfaces have stronger effect on the communication. The method to evaluate this effect includes user's behavioral data extracted from the sensors in the interfaces, physiological data and user's self report.

4.1.3.2 Participants

In total eight pairs (16 participants) joined this experiment. When recruiting, we asked one of the participants from each pair to invite "someone with a close relationship, that you communicate with frequently." A study that evaluated the effect of affective technologies in remote communication pointed out the importance of the relationship of the participants [82]. If we consider that two persons are going to convey messages with affective connotation, it is hard not to understand the importance of their relationship. Table 4.2 contains more detailed information about the pairs. Participants provided this information at the end of their session.

4.1.3.3 Materials

Two different experiments rooms were prepared for this study. The layout is illustrated in Figure 4.5, and it consisted of a desk with a chair and a computer display on which participants will receive the

TABLE 4.2: Descriptions of each pair.

ID	Genders	Type of relationship	Length of relationship
P1	M/F	Romantic (dating)	7 months
P2	M/F	Romantic (Living together)	1 year and 8 months
P3	M/F	Family (siblings)	all life
P4	M/F	Romantic (Living together)	9 years
P5	M/M	Close friends	7 years
P6	F/F	Close friends	1 year
P7	M/F	Romantic (Living together)	2 years
P8	M/F	Romantic (married)	7 years

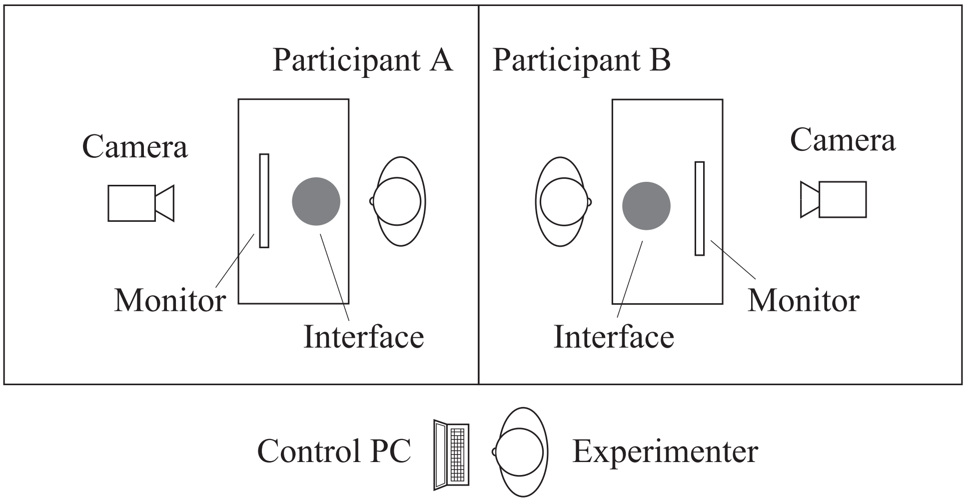


FIGURE 4.5: Layout of the experiment rooms.

instructions and contents of the experiment. Since this experiment tries to simulate a communication scenario, creating a good environment was important. To create a more comfortable and relaxing atmosphere, we reduced the room illumination by setting two small LED lamps with adjustable brightness, and yellow light. The lamps were placed on each side of the displays. Video cameras were set in front of the desks in a way that allows visualizing the participant's use of the interfaces. For synchronization purposes, a mirror was placed on a white panel behind the participants to reflect the contents of the displays.

The computers for data collection were located behind one of the panels in one of the rooms. From another computer, the videos were fed to both participant's displays and the audio through headphones. Each participant wore an E4 sensor from Empatica, to collect their physiological signals during the session. Participants rated their experience using questionnaires placed on the desks. Additionally, each participant interacted with both a Macaron module and a Virtual Macaron module. Macaron module was described on Section 3.2.1, and it displayed cues based on the rule described in Section 4.1.2.3. Figure 4.6 shows the behavior of Macaron, and how it

FIGURE 4.6: Macaron: Paired cushions displaying cues described by the proposed interaction rule.

was used by participants to communicate. Top and middle figures display two users taking turns to take a message, and the lower figure shows users sending messages simultaneously. Virtual-Macaron module is a set of tablet computers as presented in Figure 4.7. They displayed an illustration of a Macaron cushion that displayed the same colored lights patterns as the cushion version, but instead of detecting hug, it detected when the users touched the screen. Based on Macaron interaction rule, Virtual Macaron displays (A) Standby state, (B) Hug detected, (C) Message received, and (D) Sync Hugs.

4.1.3.4 Experiment protocol

Each pair arrived together and entered one of the experiment rooms, where one of the experimenters explained the overview of the experiment. After signing the consent forms, the experimenter asked them to wear the E4 sensors. Then, the experimenter explained the content of the questionnaires, to make participants familiar with them. Following this, participants were asked to sit in front of the displays, each one placed on different rooms.

The overview of the experiment can be found in Figure 4.8. It consisted of 3 conditions: 1) Macaron OFF, 2) Macaron ON, and 3) Screen. During each condition, participants received the instruction

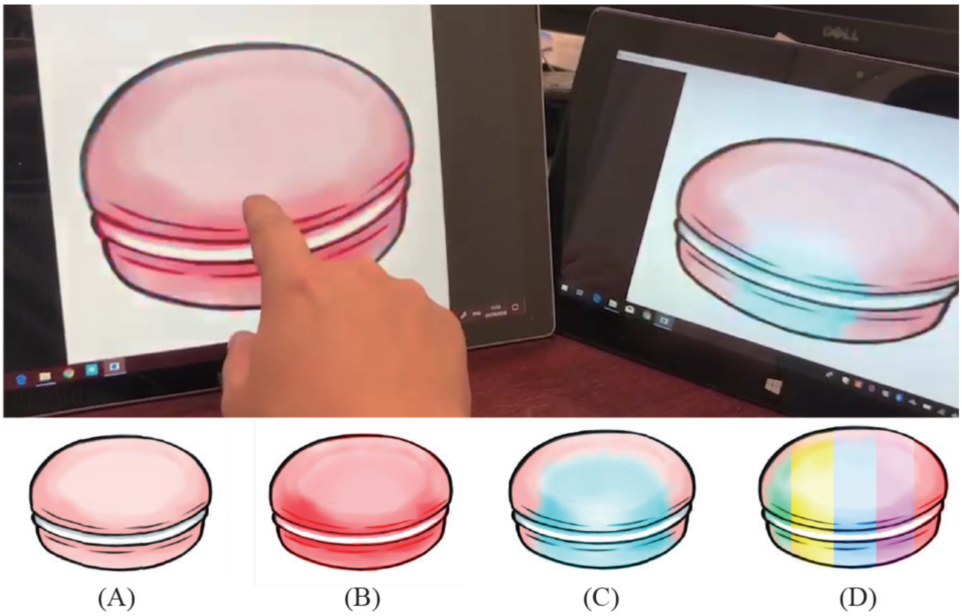


FIGURE 4.7: Virtual macaron: a set of paired touch screen that display a graphical representation of Macaron displaying the same colored patterns.

only through the display. The description of each part of the experiment is as follows:

- Macaron OFF: Participants received a Macaron module the does not deliver any feedback. To the participants, the modules represent a simple cushion. The experimenters explained the participants: "you will watch movies from now. I will ask you to hold this while watching the movies, and you are free to use it as you want". While Macaron was not delivering feedback, it was recording the sensors' data during the complete duration of the activity.

- Macaron ON: participants received a Macaron module (Figure 4.6), and they were instructed about the meaning of the colored lights patterns and showed each participant how to operate the device. The experimenters explained the participants: "you will watch movies from now. I will ask you to hold this while watching the movies, and you are free to use it as you want". The server recorded the messages exchanged through Macaron.

- Screen: participants received a Virtual-Macaron module (Figure 4.7), and they were instructed about the meaning of the colored lights patterns and showed each participant how to operate the device. The experimenters explained the participants: "you will watch movies from now. I will ask you to hold this while watching the movies, and you are free to use it as you

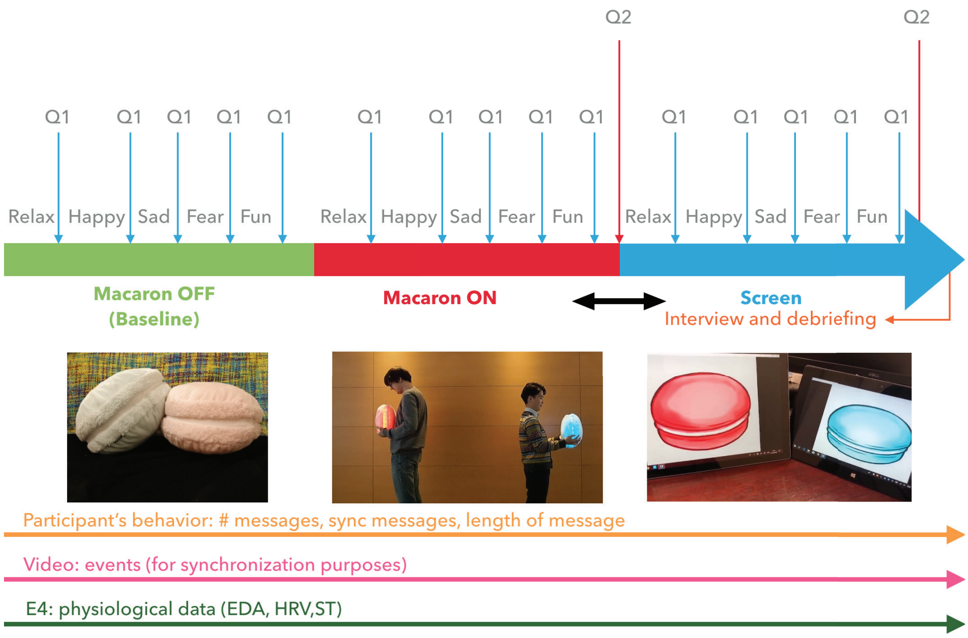


FIGURE 4.8: Flow of the experiment.

want". The messages exchanged through virtual Macaron were recorded by the server.

The experiment followed a within-subjects design; thus all the participants tried all the conditions. The sessions always started by Macaron OFF condition, followed by the two other conditions counterbalanced among the group. Each condition started with a video of 1.5 minutes to induce relaxation. The content of the videos was the same as the ones used on the work of [82]. After this video, participants were asked to answer the form Q1, found in Appendix C, Figure C.1. Q1 form contained an item of Inclusion of Others in Self Scale (IOS) to evaluate the perceived intimacy [83], and the next three items extracted from Subjective Closeness Index (SCI) [84] was used to evaluate the quality of the interaction. These items were used in the past to evaluate an affective technology for communication [82]

Participants had one minute to answer the form Q1. From here we simulated a situation where participants are watching movies at the same time, and they are allowed to communicate or not, through the interface. Figure 4.9 shows two participants watching movies during Macaron ON condition. Participants were asked to imagine they were in different countries watching movies simultaneously, and that they could use the provided interface freely during the session. As illustrated in Figure 4.8, on every condition participants watched a set of four short films. These films were taken from libraries of video material that were validated to elicit emotional responses [85–88] Four emotions were targeted, two positive (Happiness and amusement) and two negatives (Fear and sadness), and three sets of films were made (Appendix D). After each video, participants answered the form Q1, resulting in a total of 15 responses

FIGURE 4.9: Two participants during Macaron ON condition: during the complete length of the experiment participants were in different rooms, watching the movies. They only interacted through the interfaces.

during the complete experiment. Each of the three parts of the experiment ended with the evaluation of the interface through the questionnaire Q2 C, Figure C.2 taken from the study of a similar interface made by [12]. During the three conditions, participants were video recorded, and the data from the device's sensors (information about the manipulation of the interface) and the E4 data were recorded as well. The total duration of each session was about 2 hours.

4.1.3.5 Evaluation method

To understand how each condition affected the participant's communication, we collected data from different sources.

1. Questionnaires: represent the participants' report. Two types of questionnaires were used in this study. Q1 to evaluate the quality of the interaction, in terms of social attributes, perceived intimacy, and awareness. Q2 was used to understand the participant's reaction to the interface.

2. Macaron data: represent participants' behavior, in terms of the number of messages, length of the message, and the number of synchronized messages. This information will be used to understand if there was a difference in the way participants communicate through the two types of interfaces.

3. Physiological data: represent participants' states. Studies in the past have used different methods to evaluate the affective aspect of communication. One of the most popular is self-report, either by questionnaire or interview [13] [12]. A different study

used hormone levels to investigate stress reduction [17]. However, this type of research would benefit from the analysis of physiological data, as it is an alternative tool to evaluate the state of a person. Among the different signals collected by the E4 sensors, we have primarily interested in skin temperature (ST), heart rate variability (HRV) and electrodermal activity (EDA).

4.1.3.6 Analysis

ST data were analyzed by calculating the average and standard deviation. HRV was analyzed EDA data was analyzed using the toolbox from Ledalab,

The statistical analysis was done with SPSS. The data from the events observed on the videos was synchronized with the data from Macaron modules and E4 sensors. From the 2 hours sessions, we analyzed the segments when participants were watching the video stimuli. An example of the synchronized and trimmed data can be found in Appendix F. From the E4 data, we worked with ST, HRV, and EDA.

Skin temperature (ST):

1. Raw data: data is expressed degrees on the Celsius (°C) scale.

2. Analysis: the raw data was trimmed, and the target segments (during each stimulus) were extracted. In total, for each participant, we analyzed 15 data set (3 conditions, 5 videos). From the data, we calculated the average and standard deviation, a method that has been used in the past [89].

3. Interpretation: "During a state of increased exertion, excitement and stress, the muscles are forced to contract, causing a stenosis of vasculature. This leads to a reduction of skin temperature, since the blood circulation of the tissue is reduced" [90]. In other words, increase levels of skin temperature are related to a relaxed state.

Heart rate variability (HRV):

1. Raw data: we calculated HRV from the interbeat interval (IBI) data taken from E4 sensor.

2. Analysis: the raw data was trimmed, and the target segments (during each stimulus) were extracted. In total, for each participant, we analyzed 15 data set (3 conditions, 5 videos). We analyzed HRV by LF/HF ratio [91, 92].

3. Interpretation: "In this model, a low LF/HF ratio reflects parasympathetic dominance. This is seen when we conserve energy and engage in tend-and-befriend behaviors. In contrast, a high

LF/HF ratio indicates sympathetic dominance, which occurs when we engage in fight-or-flight behaviors or parasympathetic withdrawal" [91]. In other words, a higher level of LF/HF ratio relates to higher levels of stress.

Electrodermal activity (EDA):

1. Raw data: data is expressed as microsiemens (uS).

2. Analysis: the raw data was trimmed, and the target segments (during each stimulus) were extracted. In total, for each participant, we analyzed 15 data set (3 conditions, 5 videos). To analyze EDA data we used the toolbox from Ledalab, and the method of Continuous Decomposition Analysis [93], to extract the phasic and tonic components of the signal.

3. Interpretation: "EDA has been closely linked to autonomic emotional and cognitive processing, and EDA is a widely used as a sensitive index of emotional processing and sympathetic activity. The most common measure of this component is the Skin Conductance Level (SCL) and changes in the SCL are thought to reflect general changes in autonomic arousal. The other component is the phasic component and this refers to the faster changing elements of the signal - the Skin Conductance Response (SCR)" [94]. SCL is also known as the tonic component, and SCR is known as the phasic component. In other words, higher levels of SCL and SCR are related to higher arousal.

4.1.3.7 Results

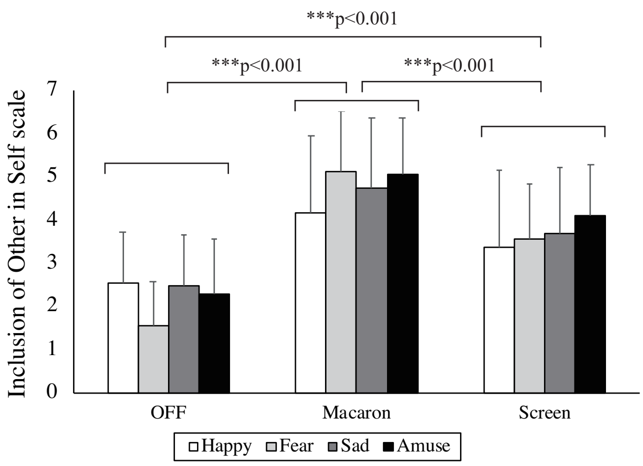


FIGURE 4.10: IOS questionnaire: perceived intimacy.

Regarding the questionnaires, Q1 form was filled 15 times by each , and it provides information about the quality of the interaction on each condition. From this form, Figure 4.10 summarizes the result from the IOS item. We analyzed the data using Three-way within

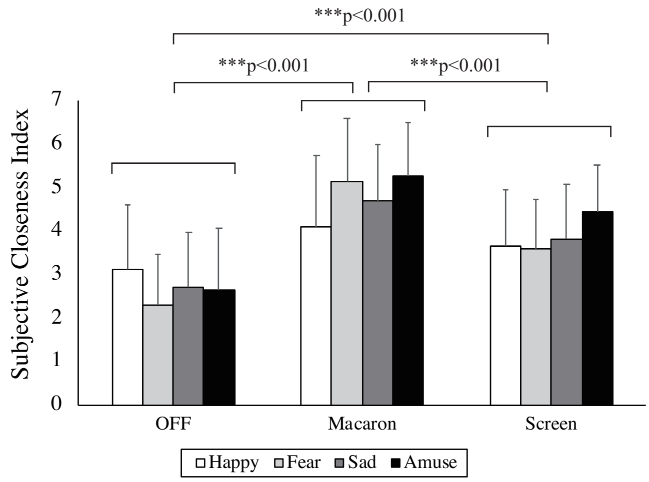


FIGURE 4.11: SCI questionanire: quality of the interaction.

subjects ANOVA; the factors were condition, stimuli, and order. Significant difference was found among conditions ($F(2,192) = 52.287$, $p < 0.001$, $\eta^2 = 0.410$). Order effect was not observed. To investigate the differences, Turkey's multiple comparison test as applied as post-hoc test. The results of each combination were: Macaron OFF - Macaron ON ($MeanDifference(MD) = -2.5469, p < 0.001, 95\%$ CI: [-3.1065, -1.9872]), Macaron ON - Screen ($MD = 1.0938$, $p < 0.001$, $95\%CI$: $[0.5341, 1.6534]$, Macaron OFF - Screen ($MD = 1.4531$, $p < 0.001, 95\%CI : [-2.0128, -0.8935]$)

A similar tendency was observed on the items that evaluate the quality of interaction. From each participant, we used the average of the three items score. Three-way within subjects ANOVA was applied and the main effect regarding: condition ($F(2,540) = 107.357$, $p < 0.001$, $\eta^2 = 0.213$) and stimuli ($F(4,540) = 33.815$, $p < 0.001$, $\eta^2 = 0.210$) were confirmed, while the interaction and order effect were not observed. As shown in the figure 4.11, significant differences between: Macaron condition and Off condition ($MD = -1.5333$, $p < 0.001$, $95\%CI$: $[-1.8425, -1.2242]$), Macaron condition and Screen condition ($MD = 0.6056$, $p < 0.001$, $95\%CI$: $[0.2964, 0.9147]$), and Off condition and Screen condition ($MD = -0.9278$, $p < 0.001$, $95\%CI : [-1.2369, -0.6186]$) were found by Tukey's post-hoc test.

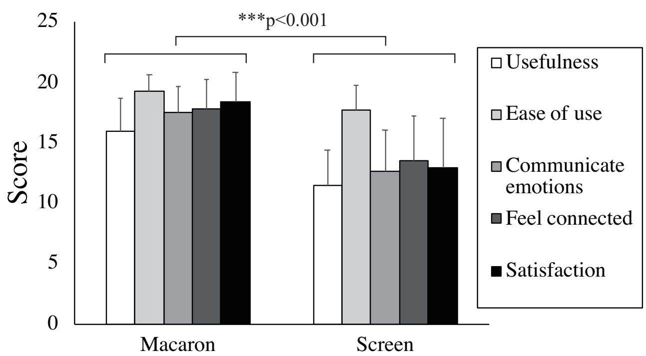


FIGURE 4.12: Reaction to each interface.

Questionnaire Q2 was used two times during the experiment to ask participants to evaluate the interface. The questionnaire consisted of 15 items grouped in 5 categories, and the analysis was done based on these categories. The results from Q2 form are presented in Figure 4.12. Results from Two-way within subjects ANOVA showed that the average score of Macaron condition is significantly higher than that of Screen condition ($F(1, 160) = 88.137$, $p < 0.001$, $\eta^2 = 0.386$). The significant difference due to the order effect was also revealed ($F(1, 160) = 4.960$, $p < 0.05$, $\eta^2 = 0.034$).

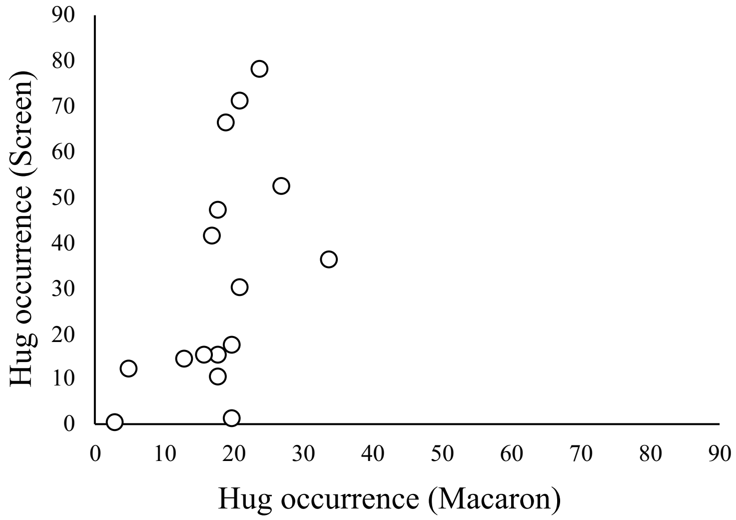


FIGURE 4.13: Number of messages exchanged on Macaron ON and screen condition.

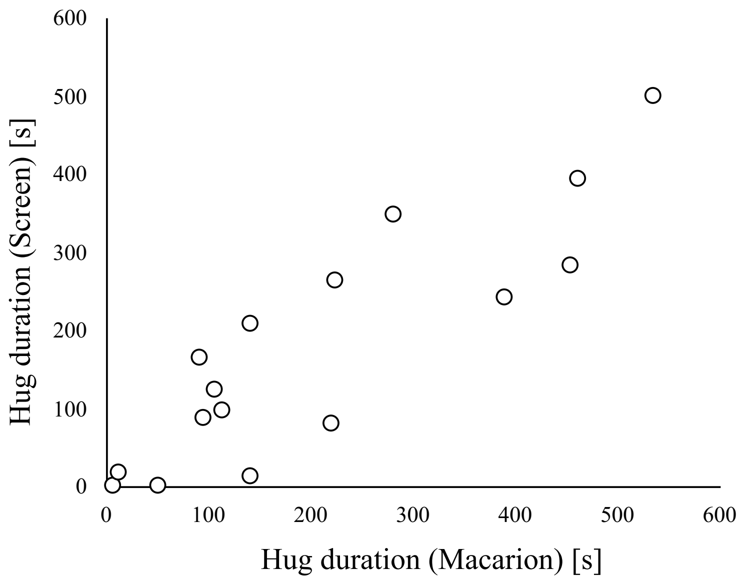


FIGURE 4.14: Duration of time that participants used each interface.

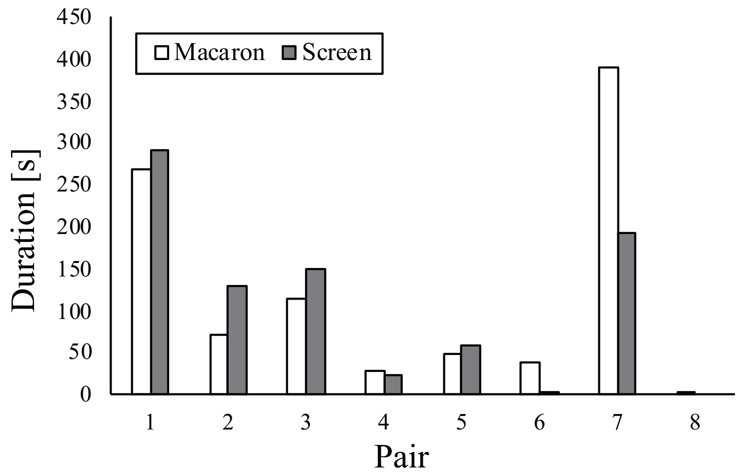


FIGURE 4.15: Duration of synchronized messages.

Regarding the messages exchanged between participants, Figure 4.13 showed the plot of the number of occurrences of messages exchanged for each participant during Macaron ON and screen condition. Participants tended to send more messages in the screen condition compared to Macaron ON condition. The data presented in Figure 4.14 refers to the total duration that participants used each interface. It was observed that each participant used both interfaces more or less the same amount of time. Figure 4.15 showed the total duration of time of synchronized message for each pair. Synchronization time of each pair did not change among Macaron ON and Screen condition.

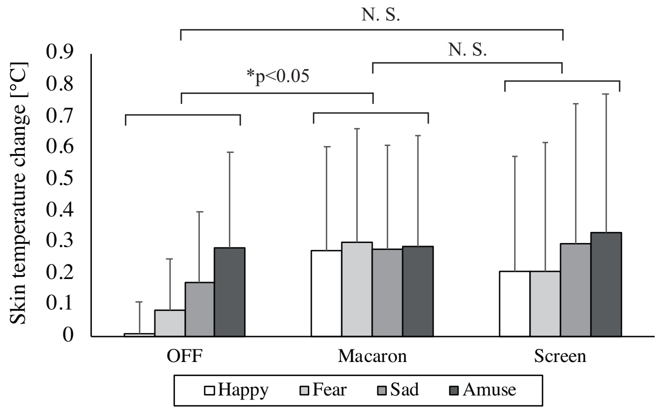


FIGURE 4.16: Average and standard deviation of skin temperature values from all participants during the three conditions.

The analysis from E4 sensors includes skin temperature (ST), heart rate variability (HRV) and skin conductance (EDA). Regarding ST (Figure E.1), the change of ST from the baseline was computed for each segment of the experiment, and compared in terms of condition, stimuli and order with Three-way within subjects ANOVA. The result shows that there was a significant increase of ST in Macaron condition compared to the OFF condition ($MD = -0.149$, $p < 0.05$,

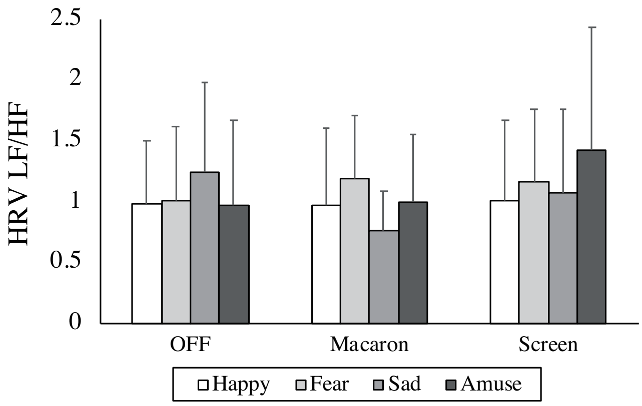


FIGURE 4.17: Average and standard deviation of heart rate variability values from all participants during the three conditions.

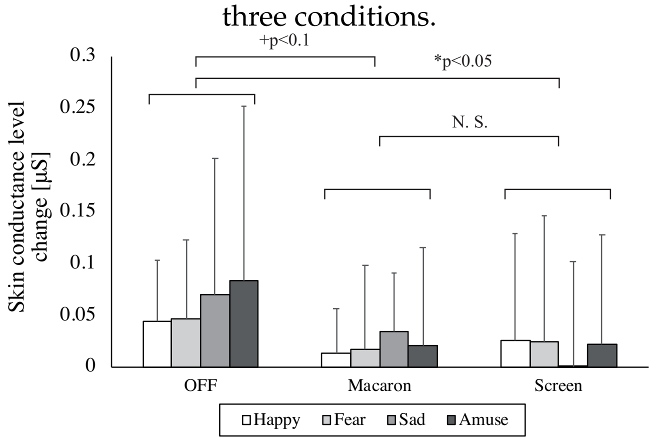


FIGURE 4.18: Average and standard deviation of the tonic component of skin conductance data from all participants during the three conditions.

$95\%CI : [-0.297, -0.0013]$) whereas no difference between the Screen condition and OFF condition was observed.

Regarding HRV, the results are presented in Figure 4.17. No significant difference was found by Three-way within subjects ANOVA.

Regarding the tonic component of EDA (Figure 4.18), there was a tendency of reduction in Macaron condition compared to OFF condition ($MD = 0.0393$, $p < 0.1$, $95\%CI : [-0.0028, 0.0814]$), and the significant decrease in Screen condition ($MD = 0.0427$, $p < 0.05$, $95\%CI : [0.0006, 0.0848]$) were observed by Tukey's multiple comparison test.

On the other hand, phasic component of EDA (Figure 4.19) showed significant difference between: OFF condition and Macaron condition ($MD = -0.0914, p < 0.05, 95\%CI : [-0.1683, -0.0146]$), Macaron condition and Screen condition ($MD = 0.0763$, $p < 0.05$, $95\%CI : [-0.0005, 0.1532]$) by Tukey's multiple comparison test.

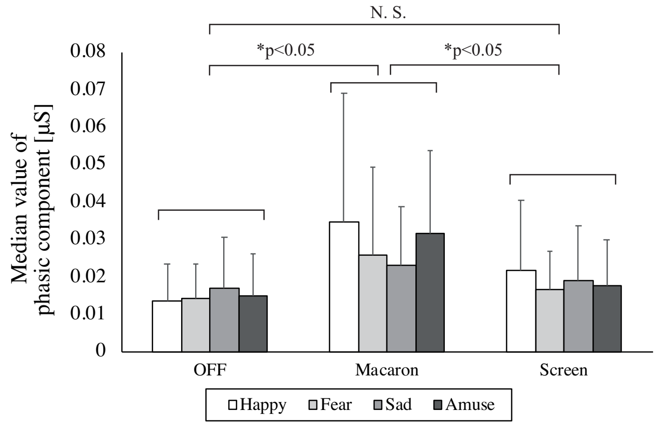


FIGURE 4.19: Average and standard deviation of the phasic component of skin conductance data from all participants during the three conditions.

4.2 Social signaling for remote communication: Pepita

Based on the results observed on the evaluation of Macaron, we propose a different approach of paired device for facilitating social signaling in remote communication. Based on the feedback received from the users during the experiment, it was implemented not only hug detection but also the detection of a touch-based gesture with negative connotation. Thus, together with hugs a negative gesture was included on the device. To expand the expressive capabilities to transfer messages with positive and negative meaning, visual cues made by colored lights were complemented with projected avatars. By doing this, the design can benefit from both, physical and virtual embodiment characteristics. This section includes the description of the devices and its preliminary evaluation of its design.

4.2.1 Purpose

Based on the importance of affective expressions in remote communication, we explore a way to enhance the paired device's expressiveness using projected avatars. This approach is beneficial as the images can be projected on different surfaces, increasing the size of the images, and making it possible to share the images with multiple users on the same physical space. Following this, we propose Pepita, designed to sense affective touch-based gestures and provide visual feedback using projected avatars. It was observed that combination of two types of embodiment (physical and virtual) to enhance the capabilities of a physically constrained robotic device is underrepresented. For this reason, we emphasize the various benefits of an embedded projector into a robotic device designed as a social mediator. This study covered the design criteria, implementation, performance evaluation of the different characteristics of the form and function of Pepita. The study was divided into three main parts: (1) the exploratory study of the different features of the device, (2) design and performance evaluation of sensors for affective interaction employing touch, and (3) design and implementation of affective feedback using projected avatars.

4.2.2 Exploring the Design of Pepita

This section introduces the methods used to investigate three features that we considered relevant in the design of Pepita: (1) the huggable aspect related to the body shape and appearance, (2) the expressiveness to convey affective states and (3) the general impression of the robot's appearance as a character. An online questionnaire using the service provided by (https://www.soscisurvey.de) was used as a tool to collect information, which is a conventional method for comparing different types of robots [95, 96]. The complete questionnaire can be found in Appendix F, Section F.1. Participants were initially contacted by social network service, and after they had agreed to participate, the link for the questionnaire was sent via email. Informed consent was obtained from the participant before starting the questionnaire. Once the participants finished answering the questions, the link was disabled to avoid double responses. The selection criteria were simple; the participants had to be adults who were familiar with technology but not involved in this or any of the projects that were introduced in the questionnaire. A total of 52 participants (age 28.0 ± 4.1 on average, 29 males and 23 females) took part in this study. Nationalities were diverse, separated in the following groups: 4 from North America, 17 from Central America, 16 from South America, 8 from Europe and 7 from Asia.

4.2.2.1 Questionnaire Overview

Different robots and devices have been designed to be hugged, but one common feature is the presence of arms. Pepita has a simple appearance, and we wanted to understand how it was perceived based on its appearance. From the reviewed huggable robots, those with the appearance of a popular cartoon character (e.g., Disney characters) and ones with a repeated type of appearance (e.g., teddy bear) were excluded from the comparison with Pepita. Therefore, we chose three huggable robots to compare to Pepita (Figure 4.20), including one huggable robot with an expressive face [97], one huggable robot with a simple appearance and no face [15], and one robot with the familiar appearance of a teddy bear [98]. In this way, we attempted to determine whether aspects like the presence of a face and a familiar body shape affected the participant's selections. The participants were presented with a photo of each robot and asked to rate the following statements using a 5-point scale: (1) It looks huggable, (2) It

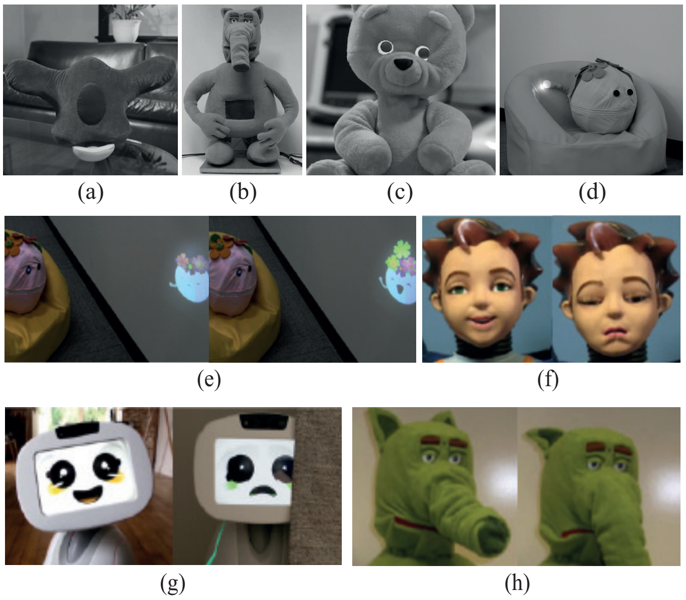


FIGURE 4.20: Photos used in the questionnaire to compare huggable robots according to the appearance: (**a**) The Hug; (**b**) Probo; (**c**) The Huggable and (**d**) Pepita. Photos used in the questionnaire to compare different robot's emotional expressions. Expressions made by display: (**e**) Pepita and (**g**) Buddy. Expressions made by mechanical face: (**f**) Zeno and (**h**) Probo.

looks easy to hug, and (3) It looks appealing to hug. The order of the pictures was balanced among the number of participants to avoid an order effect.

The second part of the questionnaire explored the modalities used by robots to convey affective expressions visually. We compared two categories: facial expressions using mechanical faces (Zeno [99] and Probo [100]) and facial expressions using a display (Buddy [101] and Pepita [102]). Considering we need video stimulus for these items, robots that fall in this category and that had video available to the public were selected. Participants were asked to watch four videos in succession showing a robot displaying happy and sad expressions. The order of these videos was balanced to avoid an order effect. They were presented one at a time, but the participants were freely allowed to replay them. After watching the videos, the participants were presented with a reference photo of each robot that appeared in the videos (Figure 4.20). Using a 5-point scale, the participants rated the following statement:

> Based on your first impression, express using the following scale how acceptable you find the robot's expressions of emotions?

The third part was related to exploring the first impression people had of the device. Along with an online questionnaire, we showed the participants a video of a person interacting with Pepita. In this way, we introduced the concept of Pepita to each participant. In the video, Pepita is on a sofa showing an avatar with a sad expression. Then, the person in the video takes the device and hugs it. After being hugged, Pepita changes the avatar to one with a happy expression. To minimize the context effect, we did not show the facial expressions of the person in the video but focused on the robot displaying its functions. After watching the video, the participants were presented with a photo of Pepita (Figure 3.4) and were asked to express their impressions using a semantic differential structure with a 7-point scale. Since a single question evaluates this item, a 7-point scale was chosen to obtain more information. The paired words selected to describe Pepita are commonly used to evaluate social robots using this structure, and it was applied following an already existing methodology [103]. Additionally, we wanted to collect some qualitative data about the participant's impressions of the robot, for which we asked an open-ended question to determine which features of Pepita positively or negatively impacted their answers. It is important to point out the limitation of this methodology: the user's perception will be different when just looking at a picture of the robot compared to directly interacting with the robot. However, this study had the goal of determining the characteristics of the robot's appearance that affected people's perception of it, and, for this, we solely used visual stimuli.

4.2.2.2 Results

Table 4.3 refers to the huggable aspect, comparing Pepita with the other three huggable robots. One-way analysis of variance (ANOVA) was applied for each aspect and significant differences in each aspect was found ($p < 0.001$, $F(3, 204) = 8.94$, $\eta^2 = 0.13$ in "Looks huggable", $p < 0.001$, $F(3, 204) = 5.92$, $\eta^2 = 0.087$ in "Looks easy to hug", $p < 0.001$, $F(3, 204) = 17.0$, $\eta^2 = 0.25$ in "Looks appealing" where p and η^2 denote significance probability and effect size respectively). Afterwards, to investigate the differences between Pepita and other robots, Tukey–Kramer's multiple comparisons test was used as a post hoc test. As the result, the following combinations showed significant differences: Pepita and The Huggable ($M = -0.673$, $p < 0.01$, 95% CI$[-1.22, -0.122]$) in the "Looks huggable" category, Pepita and The Hug ($M = 0.596$, $p < 0.05$, 95% CI$[0.0414, 1.15]$) in the "Looks easy to hug" category, Pepita and Probo ($M = 0.808$, $p < 0.01$, 95% CI$[0.225, 1.39]$) and Pepita and The Huggable($M = -0.750$, $p < 0.01$, 95% CI$[-1.33, -0.167]$) in the category of "Looks appealing to hug", where M denotes the mean difference and 95% CI represents the 95% confidence interval.

Regarding the question about the emotional expressions, each robot scored as follows: Pepita = 4.00 $\pm$ 0.929, Probo = 2.98 $\pm$ 1.15, Zeno = 3.44 ± 1.04 and Buddy = 4.52 ± 0.754. One-way ANOVA revealed that there is a significant difference among means of participants' answer for different robots ($p < 0.001$, $F(3, 204) = 25.4$, $\eta^2 = 0.27$). Significant differences were found between: Pepita and Probo ($M = 1.06$, $p < 0.001$, 95% CI$[0.562, 1.55]$), Pepita and Zeno ($M = 0.558$, $p < 0.05$, 95% CI$[0.0615, 1.05]$), and Pepita and Buddy ($M = -0.519$, $p < 0.05$, 95% CI$[-1.02, -0.0230]$), using Tukey–Kramer's multiple comparison test.

The box plot presented in Figure 4.21 shows the results of the evaluation of Pepita's appearance using semantic differential. Each item refers to a pair of adjectives. The median values of the pairs Unkind/Kind, Unfriendly/Friendly, Unpleasant/Pleasant, and Awful/Nice were found to be positive (5.5, 5, 4.5, and 5, respectively). On the other hand, the pairs Fake/Natural, Artificial/Lifelike, and Machinelike/Humanlike were found to be negative (3, 3, and 3, respectively), and the pair Unconscious/Conscious had a neutral value of 4.

After evaluating the appearance by a scale, the participants gave open-ended responses regarding those aspects that positively or negatively impacted their perception of the robot. To summarize the different answers, these were simplified using single words and grouped in categories:

- Positive aspects (mentions): Shape (8), Projector (7), Color (7), Cute (5), Size (5), Flowers (5), Tail (4), Kind (3), Huggable (3), Interactive (2)
- Negative aspects (mentions): Scary eyes (9), Shape (7), Face (6), Texture (6), Artificial (4), Appearance (4), Tail (3), Hard (2), Not huggable (2), Quality (2)

Among the positive aspects, the shape, color, and the projector had the strongest impact. Comments like the "projected avatars are great", "robots using avatars are interesting", "expressive avatars", and "avatars that display emotions" were collected from the positive aspects. About the shape, some participants expressed that it would be "easy to carry and put in a bag" or it was "round and easy to manipulate". Concerning the color, we found a positive acceptance of bright colors. The majority of the negative aspects were oriented toward the eyes as a considerable number of participants found them to be strange. Some participants expressed that "the robot seems to have one dead eye", "there is only one eye working," or "the eyes are scary". Regarding the appearance, some participants said that "it is not fluffy enough" or "the face looks weird". These comments provide us with some insights regarding the design that will be further analyzed in the discussions.

TABLE 4.3: Results of the comparison among huggable robots based on their appearance. (SD: standard deviation)

n = 52	Looks Huggable (mean ± SD)	Looks Easy to Hug (mean ± SD)	Looks Appealing (mean ± SD)
Pepita	3.21 ± 1.04	3.56 ± 1.05	2.77 ± 1.19
Probo	3.10 ± 1.23	3.31 ± 1.12	1.96 ± 0.96
The huggable	3.88 ± 0.95	3.83 ± 1.01	3.52 ± 1.18
The hug	2.83 ± 1.05	2.96 ± 1.14	2.42 ± 1.20

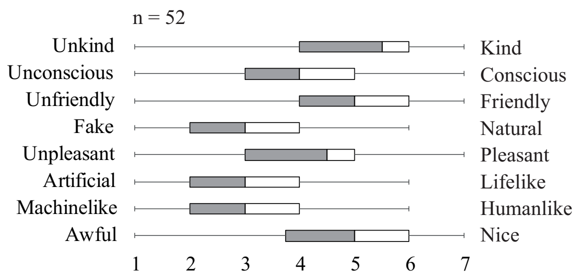


FIGURE 4.21: Results using semantic differential to evaluate the impressions of Pepita's appearance

4.2.3 Hug Detection Performance

The hug sensor had to maintain the simplicity that this design requires. This approach involved working with the sensor's structure and the shape of the robot body. The purpose of this experiment was to evaluate the performance of the pressure sensor designed to detect hugs. Because the device is spherical, there are many possible ways to manipulate it, and the sensors should be able to distinguish hugs from other kinds of touch-based interaction that involve pressing it. To evaluate this, we first observed which tactile gestures led to the majority of detection mistakes. Gestures like petting, slapping, or rotating were too different from hugging, and easily differentiated. However, those gestures that involved pressing using both hands had a higher probability of being incorrectly detected as hugs.

4.2.3.1 Experiment Setup

Ten participants joined this experiment voluntarily, and they provided informed consent before starting the session. The sessions started by asking the participants to follow a set of instructions displayed on a screen. The instructions were the following:

- Hug,
- Press with both hands on the right and left sides,
- Press with both hands on the upper and lower areas.

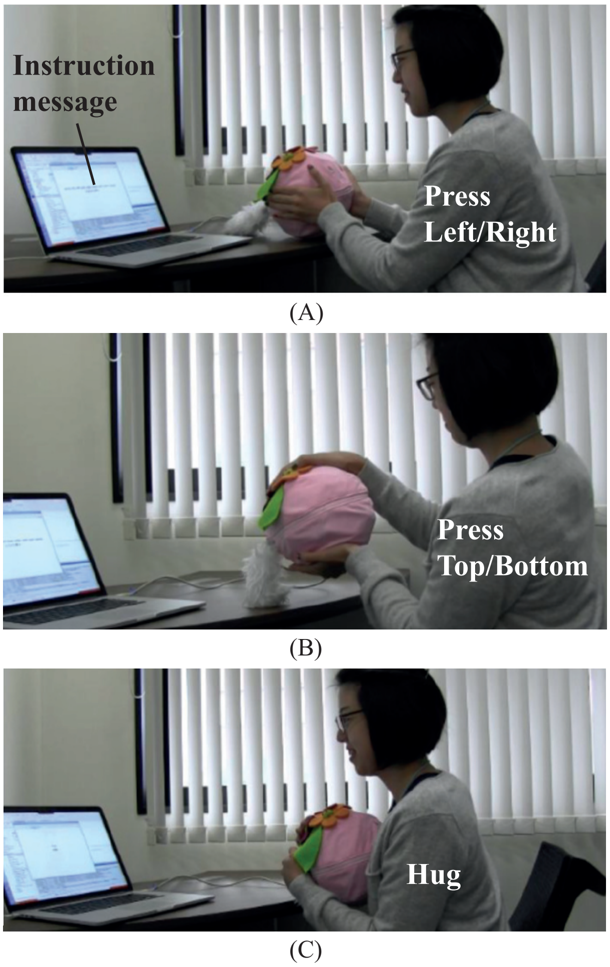


FIGURE 4.22: (**A**) participant during the "press left/right" instruction; (**B**) participant during the "press upper/lower" instruction and (**C**) participant during the "Hug" instruction.

Each instruction was displayed for 10 seconds before changing to the next one. The hug instruction was alternated with the other two instructions, which resulted in five repetitions of the hug instruction, two repetitions of pressing on the left and right sides, and two repetitions of pressing on the upper and lower areas. Participants were asked to sit down in front of a computer and hold the device in their hands (Figure 4.22). They were facing the wall, and the experimenter was standing behind them. During this experiment, participants were not given instructions on how to hug the device (i.e., to apply more pressure, press in specific places, or hold it in a particular way), and the device did not provide feedback when a hug was detected.

4.2.3.2 Results

Figure 4.23 shows the results obtained for the hug detection. The system had a hug detection accuracy of 81.8% when the robot was

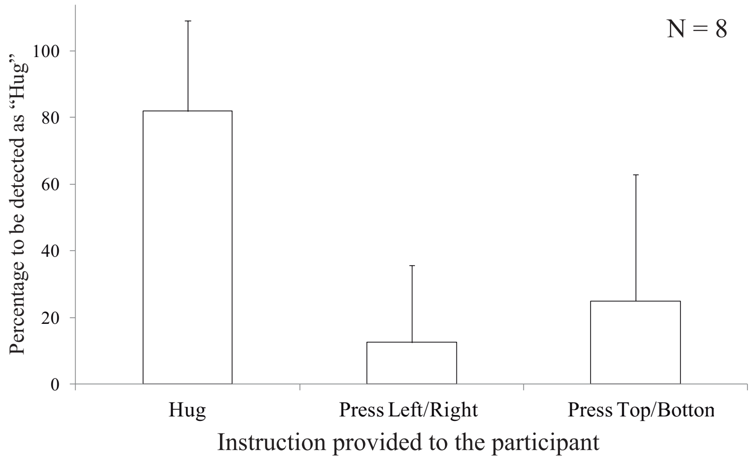


FIGURE 4.23: Hug detection performance.

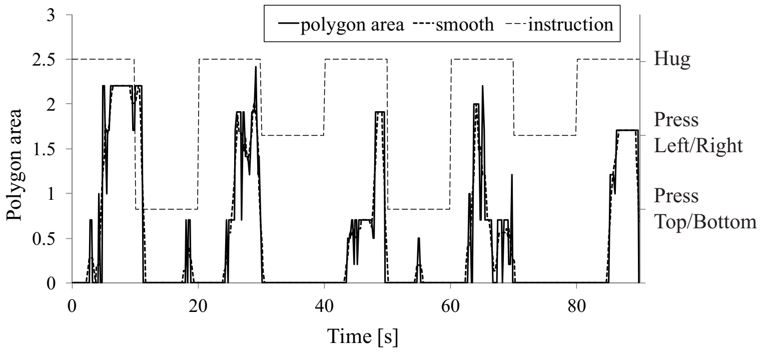


FIGURE 4.24: Data from one participant's session. The peaks of the dotted line represent the intervals when the hug instruction was displayed.

hugged. Regarding false positives, when a participant pressed on the upper and lower areas, this was detected as a hug 23.3% of the trials, and when they pressed on the sides, 12.5% of the trials were considered a hug. Figure 4.24 shows the data from one session, and the data collected during each instruction. The instruction for hugging was displayed five times during each session, and in the figure, the interval for the hug instruction is represented by the peaks of the dotted line. After collecting the data, we chose the polygon area that gave us better results representing hugs, and, in this case, it was 1.4. The hug detection performance resulted in precision = 0.84, recall = 0.82, and F-measure = 0.83.

4.2.4 Force Test for the Hug Sensor

The conductive material used to make the eight electrodes of the hug sensor has specific properties described by the manufacturer. However, different aspects alter the relation force/voltage. First, a combination of materials was used to make the electrodes, and these were

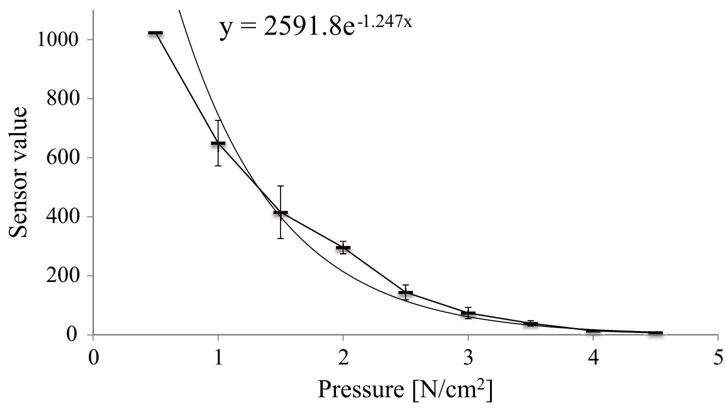


FIGURE 4.25: Sensor values for each applied pressure.

cut and shaped to cover the spherical robot body. With this experiment, we expect to explore the relationship between the applied force and the output voltage of the proposed hug sensor.

In order to do this, we tested the conductive foam using a force gauge. We used plastic circular figures with a fixed area to calculate the applied pressure. Different forces were applied, and we collected ten samples for each. From the data obtained, we will convert the voltage values collected from the participants during the hug detection test into pressure values, in order to understand how much force is necessary for detecting hugs with the system.

4.2.4.1 Results

Figure 4.25 shows the sensor values for each applied pressure. The bars represent the standard deviation, and the dotted line is the approximated curve. The parameters of the curve were defined by the least squares method. We used this curve to change the voltage values obtained during the hug detection test into applied pressure values. Following this, the data on Figure 4.24 was converted into the data observed in Figure 4.26, showing the average pressure applied during the hug detection test of one participant. Based on this test, we understood that, to generate a point for the polygon based on the selected threshold (500 (ADC value)), the user needs to apply about 1.3 N/cm^2 on the sensor piece.

4.2.5 Tail Pulling Detection Performance

We designed a stretch sensor in the tail of Pepita. The tail is made of silicon, and a linear hall effect sensor is used to measure the variations of distance from a magnet when the tail is stretched. The wiring inside the silicon was coiled to avoid ruptures. With this approach,

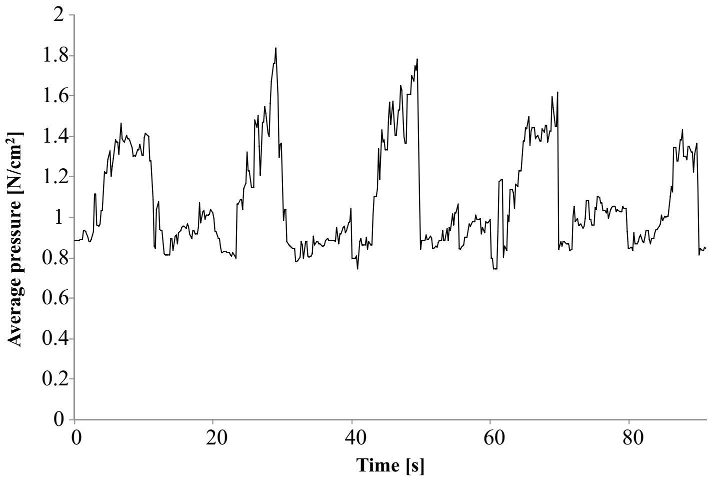


FIGURE 4.26: Data from one participant's session showing average pressure used to detect a hug with the device.

it was possible to work with the structure and material of the tail to design a simple interaction. The purpose of this experiment was to test the performance of the sensor to detect pulling behavior among other gestures.

4.2.5.1 Experiment Setup

Fourteen participants participated in this test voluntarily. The instructions were presented automatically on a screen. Participants familiarized with the silicon tail before starting the test. Each gesture instruction was presented for 3 s, followed by a release instruction for 3 s as well. The instructions were:

- Pull,
- Shake,
- Grasp.

In total, each instruction was presented five times each, and feedback from the system was disabled.

4.2.5.2 Results

After collecting the data from the participants, we selected a threshold that resulted in high detection performance: 595 (ADC value), about 2.9 [V]. A precision and recall analysis was performed using the selected threshold and the data collected from the 14 participants. The performance of the gesture detection resulted in Precision = 0.84, Recall = 0.93, and F-measure = 0.88.

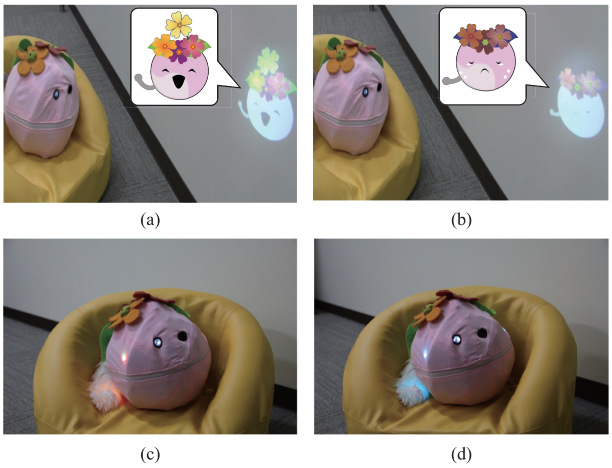


FIGURE 4.27: Representation of the robot's affective expressions: (**a**) happy with Avatar; (**b**) sad with Avatar; (**c**) happy with LED; (**d**) sad with LED.

4.2.6 Affective Feedback Using Projected Avatars

With this experiment, we explored an alternative method to represent affective expressions visually by the robotic device. We evaluated the affective expressiveness of projected images in comparison with a more conventional visual feedback method, such as colored light patterns. The expressions for a physically constrained robot developed by [30] represented four emotions: happy, sad, angry, and relaxed. Their proposed methodology suggested that it is possible to represent all the emotions that a user can perceive by implementing only these four emotional expressions, as they are found in one of each quadrant of valence-arousal. With this design, any emotion on the same quadrant is considered to be similar to the representative one (e.g., happy and delight), and it is easier to distinguish it from others from different quadrants (e.g., angry and calm). This also implies that increasing the variety of emotions to be expressed does not always benefits the quality of interaction, on the contrary, it could confuse the user especially in the case of physically constrained robotic devices. To avoid this issue, we adopted a simplified set of the robot's expressions such as one positive and one negative expression. For this reason, we are designing visual representations of a positive affective expression (happy-like) and negative affective expression (sad-like) using both avatars and colored lights.

The avatars were animated in a sequence that displayed a character that shared features with the physical robot. The avatars featured facial expressions as well as images such as blooming flowers for happy, dry flowers for sad, or changes in the color of leaves (Figure 4.27).

To design the expressions using only colored lights, we followed the methodology proposed by [32], and used the following function:

$$f(t) = \begin{cases} -\frac{1}{2}\cos\left(\frac{2\pi t}{xT}\right) + \frac{1}{2} & 0 < t \leq \frac{xT}{2}, \\ 1 & \frac{xT}{2} < t \leq \frac{T}{2}, \\ \frac{1}{2}\cos\left(\frac{2\pi t}{xT}(t - \frac{T}{2})\right) + \frac{1}{2} & \frac{T}{2} < t \leq \frac{T}{2}\left(1 + x\right), \\ -1 & \frac{T}{2}, (1 + x) < t \leq T. \end{cases} \tag{4.1}$$

(4.2)

Positive expressions like happiness are related to yellow light with a high frequency and square waveform. Negative expressions like sadness are related to blue light with a low frequency and sinusoidal waveform. In this experiment, we set T = 900 ms and x = 0.25 for the happy state, and T = 3350 ms and x = 0.75 for the sad state. These selected values were similar to the ones proposed by [32], and they have already been proven to be effective at conveying these affective expressions.

4.2.6.1 Questionnaire Overview

In this study, we had the goal of answering the following questions:

1. Can each visual element displayed by the robotic device represent the intended affective expression?
2. Comparing the LED and projected avatars, which is more efficient at representing the selected affective expression?
3. When the robot is projecting avatars, is it perceived as one entity (the robot and its avatar) or two separate entities (a robot and an avatar)?

In the experiment described next, we compared the effect of colored lights with the effect of projected avatars when they were used to convey the robot's affective expressions. To evaluate the use of avatars in this application, we developed an online questionnaire using the service provided by (https://www.soscisurvey.de). The complete questionnaire can be found in Appendix F, Section F.2. Twenty-six participants who were not familiar with the robot (18 males and 8 females, age 26.9 ± 3.7 on average) took this questionnaire. The nationalities of the participants were grouped as: North America = 1, Central America = 3, South America = 15, Europe = 1 and, Asia = 1. The participant's cultural background is a factor that can potentially impact the perception of emotions represented by color [104, 105]. In the literature, common colors are combined with other parameters to represent a robot's emotional expressions [30, 32, 106], and, for this reason, the cultural aspect is not necessarily controlled as the feedback perception is far from being only related to colors. Since the purpose of this experiment is not to evaluate the effect of the cultural background on the perception of the robot, there was no restriction regarding nationality to join this experiment.

Participants were contacted via social network and receive the access link via email. The questionnaire was not open but could only be accessed after receiving an invitation. The questionnaire consisted of two sets of two videos each, followed by questions. The order of the videos was counterbalanced to reduce the order effect. Two of the videos showed the robot projecting an avatar with emotional facial expressions (Figure 4.27a,b), and the other two showed the robot displaying LED color patterns (Figure 4.27c,d). Each participant received one of the combinations, and the combinations were proportionally balanced among the group of participants.

The videos presented the robot displaying one type of visual feedback without being contextualized by the environment or interaction. A previous study showed evidence that the context could affect the participant's recognition of the robot's expressions [107]. For this reason, we avoided influencing the participant's choices by adding elements related to interaction, such as by showing the happy state after hugging or the sad one after hitting. In this questionnaire, we attempted to evaluate only the perception of the visual elements. After watching the videos, the participants were asked to rate the following statements using a 5-point rating scale:

- In my general impression, I consider that the perceived behavior of the robot makes reference to a happy-like behavior.
- In my general impression, I consider that the perceived behavior of the robot makes reference to a sad-like behavior.

The second part of the questionnaire was related to the perception of the robot's embodiment. Because we were using two types of embodiment (the projected robot and physical robot), we intended to clarify the entity for which the participants perceived the affective expressions.

The concept of using multimodal interfaces to benefit from different types of embodiment has been explored in the past [108], where an avatar was used as a complement of a physical robot, and they were combined into one entity. The avatar was designed with an appearance similar to the physical robot, and it was implemented using a migration system, which involved having either the avatar or robot active at any given time. Following a similar approach, a different study explored user's perception when interacting with an artificial pet with two types of embodiment (virtual and physical robot), that transferred from one embodiment to the other, leaving only one of them active at a time [109, 110]. These studies pointed out the importance of making the user perceive that they were interacting with the same entity that migrated from one embodiment to the other.

In our approach, both types of embodiment, the robot and the avatar, were active simultaneously instead of one at a time. For this reason, we included one last question at the end of the questionnaire to try to understand the perception of the robot embodiment:

From the following statements, choose the one that most closely reflects your perception of the robot body interface:

- I perceive the robot body interface as two entities: an avatar and a robot,
- I perceive the robot body interface as one entity: the robot and its avatar.

4.2.6.2 Results

Based on the answers obtained from the questionnaire, we attempted to answer three questions related to this robotic device. The first question tried to determine whether both the LED and projected avatars could convey the intended affective expressions. The results presented on Table 4.4 show a clear difference for both types of stimuli. The stimulus for the happy state made with projected avatars obtained a score of 4.73 for the happy state compared to 1.46 for the sad state. In the case of the stimulus made with colored lights, the happy state obtained 3.62 compared to 2.19 for the sad state. The stimulus for the sad state made with projected avatars obtained a score of 4.35 for the sad state compared to 1.27 for the happy state. Colored lights obtained 3.69 for the sad state and 2.23 for the happy state. For all the combinations, the participants could perceive the represented affective state showed on the videos using both LEDs and avatars. Note that, to confirm the effect of the order of stimulus on participants' score, we applied a Kruskal–Wallis test for each stimuli, and no significant difference due to the order has been found for all stimulus (Avatar (Happy): $p > 0.5,\ \chi^2 = 1.65,\ \eta^2 = 0.066$, Avatar (Sad):$p > 0.5,\ \chi^2 = 2.30,\ \eta^2 = 0.092$, LED (Happy): $p > 0.5,\ \chi^2 = 1.63,\ \eta^2 = 0.065$, LED (Sad): $p > 0.5,\ \chi^2 = 1.91,\ \eta^2 = 0.076$).

To answer the second question, we compared the stimuli made with the projected avatars and LEDs to display the affective state. When presenting the stimuli for the happy state, participants' perception of a happy state obtained 4.73 with projected avatars compared to 3.62 with LEDs. In the case of the stimuli for the sad state, participants' perception of a sad state obtained 4.35 compared to 3.69 with LEDs.

The third study was related to the question of embodiment, and the results showed that 20 participants perceived the projected avatar as part of the physical robot, five participants perceived that the avatar and robot were different entities, and one participant did not include this answer.

TABLE 4.4: Mean and standard deviation of participant's answer as "Happy" or "Sad" to each stimulus.

n = 26	Projector		LED	
	Happy	Sad	Happy	Sad
Score as "Happy" (mean± SD)	4.73 ± 0.59	1.46 ± 0.80	3.62 ± 0.88	2.19 ± 0.83
Score as "Sad" (mean± SD)	1.27 ± 0.52	4.35 ± 0.92	2.23 ± 0.85	3.69 ± 0.72

4.2.7 Application Scenarios for Pepita

This section includes 1) Description of the Social Context of Pepita and 2) Combining Robots with Projectors.

4.2.7.1 Description of the Social Context of Pepita

Based on the features selected for the design of Pepita, one of the possible applications is as a mediator of remote communication. This involves a paired robots configuration, which means that each user will communicate using one identical robot. The current implementation is driven by the user's tactile gesture (hug/pulling the tail), and it communicates with the user using projected avatars. These avatars are displayed by a running application in the robot's smartphone. This feature allows for easily assigning different avatars to the different users, as a kind of ID. The robot will display its avatar with a happy expression every time the user hugs it. At the same time, every time a robot is hugged, it is translated as a message and displayed on the paired robot. As time passes without a user hugging a robot, the avatar will change to showing a sad expression. The same robot will display the partner's avatar every time it receives a hug from the paired robot. Similarly, as time passes without receiving more messages, the avatar of the paired robot will change to display a sad expression. It is expected that the users will be aware and feel motivated to interact with the robot to improve its affective state, which is also a message for the partner. Since this type of communication does not contain any detailed content, the purpose of the message is open to interpretation.

Pepita is envisioned to be placed in two scenarios. On the first one, the robot works as a standalone communication device, and each message works as a kind of presence indicator (Figure 4.28A). The second scenario involves Pepita working as a complement of verbal communication, and, in this context, Pepita works as a tangible emoticon (Figure 4.28B). With this, it is expected to enhance the transmission of affective messages and foster a sense of co-presence.

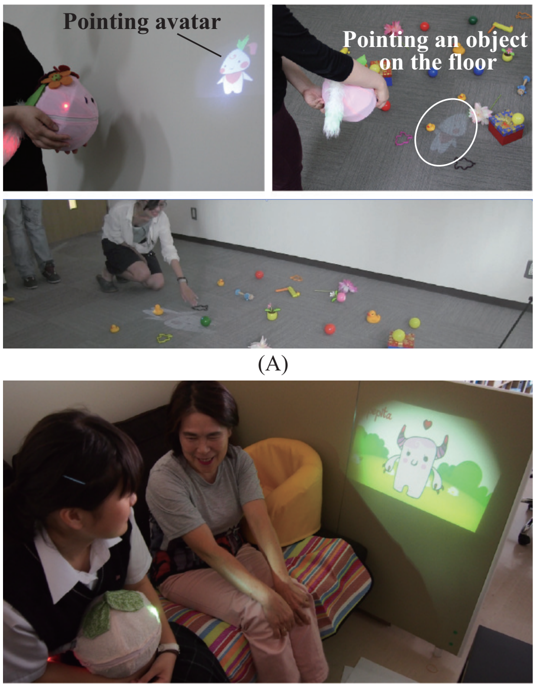
Two user scenarios are proposed to understand how Pepita mediates expressions of affection in remote communication. The first one

FIGURE 4.28: Social context for the proposed companion robot Pepita: (**Top**) as a tangible emoticon; (**Bottom**) as a presence indicator.

reflects the use of Pepita as a standalone communication device, and, on the second one, Pepita works as a complement of verbal communication:

1. Jane is a college student living away from her family. Every morning before going to the university, she leaves a message to her parents by hugging her robot. Jane observes how her avatar displays a happy expression and then sets the robot on the sofa. She comes back home in the evening and observes her parents' avatar displaying a sad expression, which she understands as "I received a hug from my parents a long time ago", she takes her Pepita and hugs it, observing her avatar appearing with a happy expression, and conveying to her parents that she is now home.

2. On the weekend, Jane is talking with her mother by phone sharing stories of the recent days. She takes her robot and hugs it sending a message. Her father, who is in the living room watching TV, observes their robot reacting showing Jane's robot's "happy" avatar. Then, he asks to speak to her and say hello. While talking to her, he hugs back.



(A)

(B)

FIGURE 4.29: Enhancing the robot's expressiveness with projected images: (**A**) avatars to make pointing gestures using the robot in remote communication; (**B**) scenario of a robot sharing information with multiple users in the same place and remotely.

4.2.7.2 Combining Robots with Projectors

Designing robots with a screen for a face is becoming popular because it makes it possible to display a broad range of expressions that are easy to identify. Screens can easily be related to computers, and computers are machines like robots. For this reason, screens can be considered to be appropriate parts of robots. Projectors work in a way similar to screens because they make it possible to display multiple kinds of information. The difference between them is related to the way the user interacts with the interface. Embedded screens are constrained to the size of the robot's body, and thus are commonly small. This causes users to become immersed in the contents of the screen. On the other hand, projectors can be used to display large images that can be shared by multiple people sharing the same space. It is interesting to explore the social aspect of robots with projectors, and their effect on human interaction.

Projected images have a substantial visual impact and they allow to keep the robot structure simple. Moreover, projected images create opportunities to share the experience with more than one user sharing the same space, using a small robot body. By combining robots and projectors, it is also possible, for example, to share photos and memories with people in the same place and those in different locations, opening different possibilities for future designs (Figure

4.29B). Enhancing a robot's capabilities using a projector is not limited to facial expressions. Robots with no limbs like Pepita have limited body's capabilities, and projectors can be used to display animations of body gestures. For example, avatars with pointing gestures projected from the robot could potentially allow users to perform actions together regardless of whether they share the same space (Figure 4.29A).

However, projected images are hard to see in a bright environment, and they can be occluded when the robot is hugged. One possible solution to overcome these limitations could be by notifying the user when a message arrives by combining LEDs and vibrations. By doing this, the user can adapt and move the robot to a darker place if necessary, or uncover the lens of the projector. Other limitations are related to the heat generated by the lamp which could be a problem if it is used continuously for extended periods of time. Future similar projects need to consider the problem heat can represent for the application carefully.

4.3 Social signaling for children with ASD: Cololo

Paired robotic toys Cololo, introduced in Section 3.2.3, were designed to mediate the exchange between a child diagnosed with ASD and his therapist, during a game for training turn-taking skills. As summarized in Section 3.1.2, children with ASD has difficulties in understanding social cues, and their interest in mechanical toys and visual information make them a good alternative as a tool in interventions. Cololo identify when the user, child or therapist, manipulates the toy (sense behavioral cues by touch channel), and transfer this action as a message to the paired device, represented by patterns of colored lights and vibrations. These were the characteristic selected for the proposed approach as explained in Section 2.2. The following section introduces how Cololo was used in the interventions, and the results from the experiments that explored the effect of Cololo on children's behavior.

4.3.1 The need and purpose of this study

To facilitate turn-taking on children with ASD, we proposed a system using paired devices Cololo [111, 112]. This system was used to facilitate a fundamental exchange of the intention to communicate such as social signals in turn-taking interventions. The design of Cololo follows a concept similar to that of [23], which employs robotic toys with colored lights to engage children in turn-taking activities. However, our approach targets individuals with low-functioning autism, and in order to investigate the context-independent effect of visual stimuli on the turn-taking of children with ASD. Therefore, the interaction rules of the robotic toys were simplified and used to indicate the exchange of turns and represented using paired devices.

In the past we investigated the effect of the intervention with Cololo on children's play and social behavior using a psychological approach [113]. We compared the effect of Cololo's visual feedback condition against no-feedback condition and found that visual feedback was related with an increased number of contact with toys and an increased number of gaze shifting toward the therapist's toy. However, no investigation was done regarding the effect of different Cololo interaction rules on play/social behavior of children and its potential for the facilitation of turn-taking. Moreover, the use of Cololo as a measurement device to describe the behavior of children automatically and quantitatively during the intervention has not been discussed.

In this study, we propose two interaction rules for the intervention using Cololo, the "Two-sided lighting rule" and the "One-sided lighting rule", in order to clarify the effect of different sensory feedback methods (i.e., interaction rules) on children's behavior. These rules are implemented based on the results of a pilot experiment with a previous interaction rule of Cololo [111] to overcome its limitation and be beneficial for facilitating turn-taking behavior. Thus, the purposes of this study are: (1) To evaluate how the devices influence turn-taking behaviors of children by comparing the effects of the two different interaction rules; and (2) To explore the potential for and the limitations of describing children's play behavior in turn-taking interventions using quantitative data from the robotic devices. To investigate this, an experimental study was performed with a therapist and children with low-functioning autism to compare the data obtained from the human analysis with the data obtained from Cololo. Using this data, the effect of the two rules was explored by quantifying the behavior of children related to turn-taking.

The contributions of this work are as follows:

1. The introduction of different interaction rules using paired devices for turn-taking.

2. Investigation of the effect of different interaction rules using paired devices on the turn-taking behavior of children with ASD.

3. Investigation on the capacity and viability of using Cololo for automated quantitative measurement of children's and therapist behavior during therapy sessions.

4.3.2 Cololo in Turn-Taking Interventions

Turn-taking is an important and fundamental skill that involves the spontaneous exchange of initiative during social interaction. Turn-taking involve being able to understand and produce social signals, which are an essential part of communication. Different approaches particularly targeted training turn-taking skills and using robotic toys for this purpose has shown positive effects [21–23, 61]. Robotic toys can respond to the child's actions with engaging and constant stimuli, which is beneficial to guide the child during a game that involves turn-taking.

Following this, this study explores the use of paired robotic toys Cololo in facilitating turn-taking. Considering the lack of awareness of social signals used in turn-taking, Cololo is used to convey simple cues and therefore facilitate the exchange of actions during a game for training turn-taking skills. On the sessions with Cololo, there are two devices/modules, one held by the therapist and one by the child, and they take turns to manipulate the toys and change their colors. The change of color indicates the event of "taking a turn", and the color itself has no specific meaning. Manipulation of the toys includes any action that activates the tilt sensor inside the device, including shaking, kicking it, rolling it, among others. When the user manipulates the device during the user's turn, the device sends a message to the paired device and "passes the turn". The therapist uses these two devices to assess and train turn-taking skills, based on the timing the child manipulates the toy. The child must learn when to interact with the ball by realizing when it is his/her turn aided by the cues delivered from the devices. The way these cues are delivered is defined by interaction rules.

4.3.3 Interaction Rules and proposed theory

The interaction rules of Cololo are illustrated on Figure 4.30. The left-most of Figure 4.30 represents the previous Cololo interaction rule in which both devices always respond when manipulated (indicated by black arrows) and provide feedback through the change of light color and movements. However, this rule does not include the concept of a "turn-holder", and it was used only for the pilot experiment (Section 4.3.5.2).

Figure 4.30A represents the "Two-sided lighting rule" in which the feedback with colored lights is always enabled for both devices and indicates the turn-holder through wiggling movements. A "turn" or "turn holder" refers to the device that is enabled to send a message to the paired device. When a device is manipulated while it is holding the turn, the device sends a message to the server to provide the associated visual feedback, and then the other device will hold the turn for the next manipulation. On the other hand, when the device that is not holding the turn is manipulated, the devices will not provide any feedback, and the manipulation will be recorded as a "failed attempt". Since LEDs of both devices are always enabled, this rule would be engaging to children with ASD who are known to be sensitives as well as positively responsive to visual stimuli.

Figure 4.30B represents the "One-sided lighting rule" in which the colored-lights feedback is disabled (i.e., "off") while the device is not holding the turn. In this rule, the device that is holding the turn is identifiable by lights and wiggling movements. Removing the feedback made by lights from the receiving device might be beneficial to indicate the turn-holder and reduce distractions.

On both rules, the turn-holder is indicated by the wiggling device, and the changes of turns are indicated by changes of colors. The difference between these two rules is the role of the feedback made by lights, as it can be used to also indicate the turn-holder or not. Based on the considerations above, we proposed an explanation based on the relationship between the two interaction rules and children's behaviors related to turn-taking as follows:

1. Lighting up both devices (lights do not indicate the turn holder), will elicit a higher number of manipulations because of its relatively higher amount of visual stimuli.
2. Lighting up one device (lights indicate the turn-holder) will reduce the number of the child's manipulations during the therapist's turn (i.e., reduce the number of failed attempts of turn-taking).
3. Lighting up one device (lights indicate the turn-holder) will lead to a higher occurrence of turn-taking by the child.

4.3.4 Turn-taking Behavior Analysis Using Data From Cololo

On the interventions with Cololo, turn-taking is described based on the manipulations of the toys and timing. Using Cololo, the history of messages exchanged between the participant and the therapist can be recorded as shown in Figure 3.8 (right), to describe and analyze the participant's behavior during a turn-taking intervention. Each message contains the following information:

1. Time: the time a message from any of the toys is received by the server.
2. Label: "RECEIVE" indicates that the tilt sensor in a device detected a manipulation and then the server received a notification. If the manipulation was made during its turn, a new line will be added in the log with the label of "CHANGE", and the server then forwards the message to the paired device.

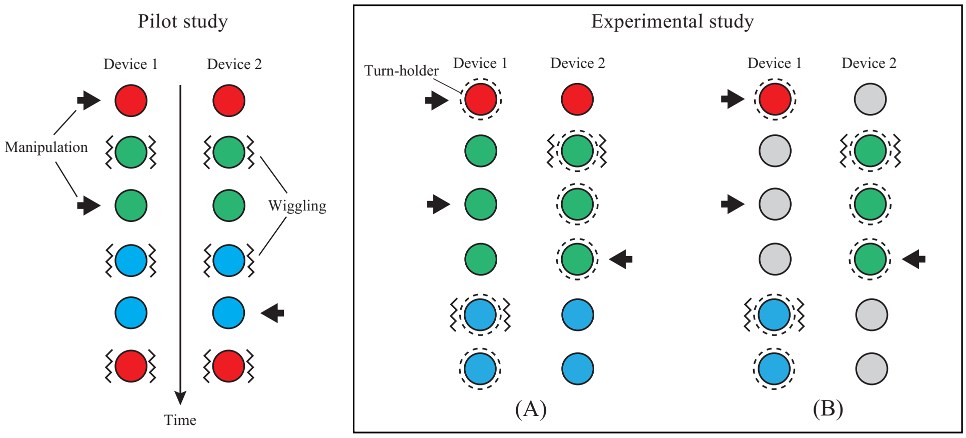


FIGURE 4.30: Three interaction rules for Cololo: Previous rule [111] in which both devices always respond when manipulated (indicated by black arrows) and provide feedback through the change of light color and movements, (**A**) Two-sided lighting rule. The feedback with colored lights is always enabled for both devices and indicates the turn-holder through wiggling movements. A "turn" or "turn holder" refers to the device that is enabled to send a message to the paired device, (**B**) One-sided lighting rule. The colored-lights feedback is disabled (i.e., "off") while the device is not holding the turn. In this rule, the device that is holding the turn is identifiable by lights and wiggling movements.

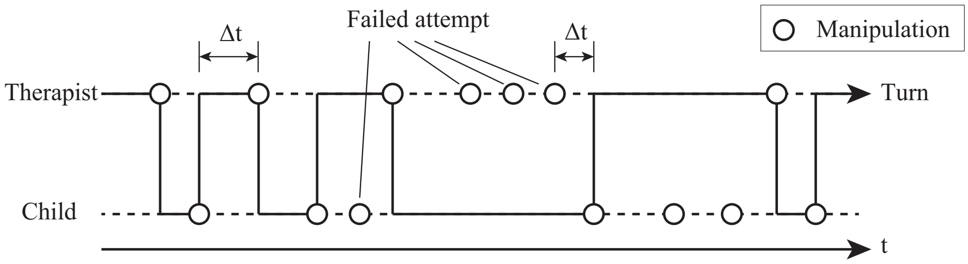


FIGURE 4.31: Criteria used to analyze the data from Cololo. Each mark represents manipulations of the toys made by two users. They are separated in two time lines and used to describe the exchanges made by the child and the therapist during the interventions.

3. ID: the identification number of the sender device. There is a number for the therapist's device and for the participant's device.

4. Color: the information related to the light color of the sender.

Figure 4.31 illustrates an example of how Cololo data are visualized. Each dot represents a manipulation detected by the device. Turn-taking, on the other hand, is a combination of different behaviors (e.g., taking actions, impulse control, and gaze orientation among others) and it cannot be represented using only the data from Cololo. However, since the turn-taking behavior is observed as a time series of manipulations, the temporal feature such as the interval between manipulations (Δt) can be used to parametrize the turn-taking in the Cololo data. Turn-taking described by the data from Cololo will, therefore, be constituted by the messages exchanged within the time interval (Δt) between the paired devices. Considering that the sessions using Cololo involve the therapist and the child exchanging turns to change the color of the devices, it is essential to select an appropriated time interval that allows the system to exclude those exchanges that less likely will be turn-taking.

4.3.5 Evaluating the effect of paired devices Cololo in facilitating social signaling

In this section, a pilot study conducted for establishing the two-sided lighting rule and one-sided lighting rule is first explained, followed by the experimental study for verifying the hypotheses regarding the effect of sensory feedback using the two proposed rules on children's behaviors.

4.3.5.1 Experimental Setup

Experiments were held at the playground in Keio University, which consists of a carpeted room with a two-way mirror behind which the parents could observe the sessions. In the sessions with Cololo, the therapist and the child played in the center of the room most of the time, and the session was video-recorded using two cameras. The data from Cololo were collected by a computer placed in a separated room. In addition to the therapist that facilitated the intervention with Cololo, two other observers analyzed the sessions. Considering that the visual feedback from the devices was difficult to see in a fully illuminated room, the room's main lights were switched off, and two lamps were placed in two corners of the room.

At the beginning of every session, the therapist introduced Cololo and showed the child how to play with them. The child was then given the opportunity to initiate playing with Cololo. The game consisted of taking turns to change the color of the balls by manipulating them. Since Cololo looks like a ball, children manipulated it as one (rolling it on the floor, shaking it, throwing it or pushing it against the other). When the child intended to manipulate Cololo during his turn, he was praised by the therapist (e.g., exclamations such as "great!"). When the child was playing with his device during the therapist's turn, the therapist gave indications to guide his attention back to the activity (e.g., "look at me, here I go!"). When the child lost complete interest in the toys, the therapist attempted to capture his interest by showing him how to play (e.g., by rolling the ball in front of him and saying, "let's play!").

This study was approved by Keio University, Faculty of Letter's Institutional Review Board and was, therefore, completed in accordance with the ethical standards established in the 1964 Declaration of Helsinki. Written informed consent from the parents of all participants was obtained. All participants had a diagnosis of autistic disorder, PDD-NOS, or ASD by an external medical doctor. Diagnosis of Pervasive Developmental Disorders (PDDs) was further confirmed using the Pervasive Developmental Disorders Autism Society Japan Rating Scale.

4.3.5.2 Pilot Study

One participant was selected for this intervention, a male (6 years and 11 months old) diagnosed with autism and Attention Deficit Hyperactivity Disorder (ADHD). The purpose of the pilot study was to observe the way the participant interacted with Cololo under the previous interaction rule while playing with the therapist.

Through this pilot study, we observed that the visual sensory feedback delivered from Cololo engaged the participant in the activity. However, it was difficult to elicit play behaviors related to turn-taking because in the previous interaction rule, Cololo provided visual feedback every time the participant manipulated the device regardless of who the turn-holder was. The therapist used different instructions in an attempt to try to guide the participant's gaze away

TABLE 4.5: Participants' profile.

Participant	Chronological Age (year; month)	Developmental Quotient	PARS	CARS
P1	4;7	43	46	44
P2	4;5	39	21	36
P3	5;8	70	26	32.5
P4	3;8	44	21	N/A

(A) (B)

FIGURE 4.32: (**A**) Participant during a session with the Two-sided lighting rule, (**B**) Participant during a session with the One-sided lighting rule.

from his device, and give him the opportunity to look at the therapist taking his turn to change the color of the devices. But since the participant observed that the device always changed color every time he manipulated it, it was difficult to use the device feedback to elicit turn-taking.

Based on the findings from the pilot study, the two-sided lighting rule and the one-sided lighting rule were designed by implementing the concept of the turn-holder explicitly (i.e., only one Cololo could send a message at a time), and indicating the turn-holder via visual and haptic cues (i.e., colored lights and/or wiggling movements).

4.3.5.3 Experimental Study: Comparison of Two Interaction Rules

Four male participants were recruited as volunteers through the Department of Psychology at Keio University. Their profiles can be found in Table 4.5. Informed consent was obtained from their parents before the children were included in the study. All participants tried each interaction rule in separated sessions on different days. They tried the two-sided lighting rule on the first session and the one-sided lighting rule on the second session. Figure 4.32A shows a participant in a session with the two-sided lighting rule, and Figure 4.32B shows a participant in a session with the one-sided lighting rule.

The interventions lasted 5 minutes, but in those cases where the therapist could not engage the child in the activity, the sessions were terminated before that. On both conditions the turns were mediated by the devices, allowing one user, either the therapist or the participant, to take a turn each time. The purpose of this experiment was: (1) To evaluate how the devices influence children's behaviors related to turn-taking by comparing the two modified rules, and (2) To explore the potential and limitations of describing turn-taking interventions using the data from the robotic toys.

4.3.5.4 Effect on Children's Behavior During Turn-taking Interventions

The sessions were analyzed using both, the data from human coders and the data from Cololo. Three human coders (C1, C2, and C3) watched the videos from the intervention and were asked to count the number of gaze shifting and manipulations of the toy. During the interventions with Cololo, the participant and the therapist were exchanging turns to manipulate the toy. While taking turns to manipulate the toys and change their color, the participant is expected to be able to look at the therapist while waiting his turn. This skill is of importance for turn-taking interventions, as looking at the peer helps to understand context and intentions. The role of Cololo is to facilitate this exchange by delivering visual cues to guide the child's gaze and to indicate the moment to manipulate the toy.

For counting the number of gaze shifting instances, the videos from the sessions were divided into segments of 10 s, and the human coders were asked to judge if gaze shifting occurred or not during each segment by assigning a numerical value such as 0 or 1. Scoring intervals is a common practice for analyzing social behaviors in interventions for children with ASD [114, 115]. Note that, more than one instance of gaze shifting in the same segment still counted as 1. The instructions given to the human coders for identifying gaze shifting were as follows:

- Condition: During the therapist's turn and when a message is being exchanged. Indicators: Each time the child directs his gaze from the therapist's facial region or the therapist's device to his device.
- Condition: During the child's turn and when a message is being exchanged. Indicators: Each time the child directs his gaze from his device to the therapist's facial region or the therapist's device.

For counting the number of toy manipulations, the human coders were asked to use the definition presented below. Moreover, identified manipulations were categorized as a "successful" or a "failed" attempts of turn-taking, according to the result of the action of manipulating the toy. Indicators of clear attempts to manipulate the device were described as observing the participant's gaze on one of the devices or the therapist's face while directly manipulating the toy (i.e., shaking, throwing, pushing, or rolling with either hands or feet).

- Successful attempt of turn-taking: described by clear attempts to manipulate the device that result in a device changing color (the child manipulates his toy during his turn)
- Failed attempt of turn-taking: described by clear attempts to manipulate the toy that did not result in a change of color (the child manipulates his toy during the therapist's turn)

4.3.5.5 Turn-Taking Analysis by Cololo

In our previous study, two human coders (C4 and C5) analyzed the videos interventions to identify when turn-taking was observed [112]. The therapist provided a definition of turn-taking that was suitable for interventions with Cololo. It was defined as a social exchange mediated by the device's visual cues and it was counted as an occurrence:

- Each time the child took his turn: described by clear attempts to send a message by manipulating Cololo, with the gaze on one of the devices or the therapist's face (Pattern A).
- Each time the child waited without manipulating Cololo during the therapist's turn: the child is looking at the therapist or paired device, showing no intention of manipulating the device (Pattern B).

Using Cololo's log file, it was possible to extract the number of exchanged messages and the time interval between each of them. However, not all the messages recorded by Cololo represent turns exchanged between the child and the therapist. For example, during the sessions, there were cases where the child started playing with his toy without looking once to the therapist. According to the definition, even if Cololo is recording the child's manipulations of the toy, these do not represent turn-taking. For this reason, to compare the data from the human analysis with the data from Cololo, it is necessary to exclude as much as possible, those exchanges that are not turn-taking. Since the information collected by Cololo is the time interval between the exchanged messages, this is the information that we are going to extract from the human analysis. The process to extract the time intervals from the human analysis had two steps: (1) The selection of a time interval (Tmin–Tmax) that contains the majority of observed turn-taking, and (2) Filter Cololo data by excluding the messages outside the selected time interval, and compare Cololo data to the human analysis.

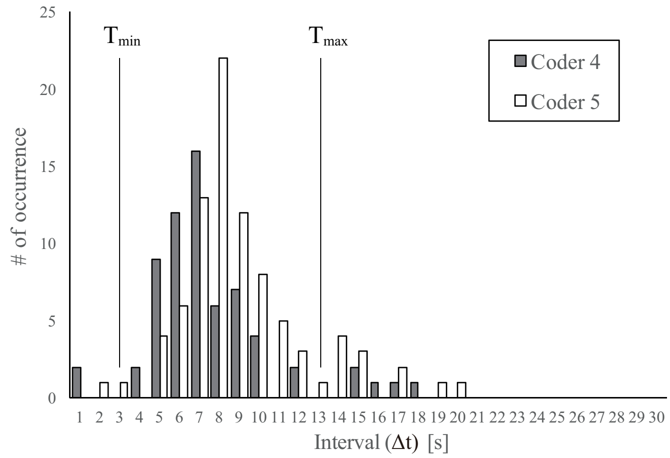


FIGURE 4.33: Length of the time intervals between the messages exchanged during turn-taking interventions.

To select the time interval (Tmin–Tmax), two human coders (C4 and C5) counted the number of turn-taking instances using the previously mentioned definition. They counted from the moment they observed the child's intention to manipulate the toy until the paired device was manipulated as a response (Pattern A), and also from the moment the therapist intended to manipulate the toy until the child's manipulation of the toy (Pattern B). From this information, we obtained the time intervals between all the exchanged messages that were considered turn-taking by the coders. It was observed that both coders followed a similar pattern that formed the curve represented in Figure 4.33. The selected time interval Tmin–Tmax was between 3 and 13 s, which is where coders C4 and C5 observed the majority of turn-taking.

The second part of the analysis involved asking different coders (C1, C2, and C3) to do the same analysis and count turn-taking instances based on the definition. The coders analyzed the sessions of the 4 participants with both interaction rules (Human data). With the previously selected interval Tmin–Tmax, we filtered the data from Cololo as follows: those exchanges of messages within the interval of 3 s and 13 s, were discarded (Cololo data). With this, it was expected that the number of messages exchanged that were less likely to be considered turn-taking by human analysis would be reduced. The results from this analysis compare the human data with Cololo data to explore the agreement between these, and to understand how much it is possible to describe turn-taking using the data from the toy's manipulations.

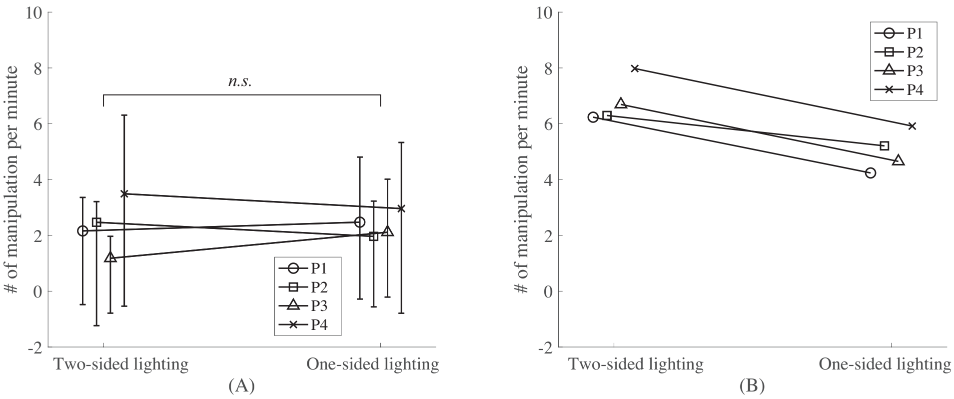


FIGURE 4.34: (**A**) Average and range of the number of toy manipulations counted by three coders (C1, C2 and C3), (**B**) Number of events detected by Cololo.

4.3.6 Results

In this section, results from the video analysis by human coders and Cololo data are presented. Three coders were asked to count the occurrences of the participant's manipulation and gaze-shifting behavior based on the definitions explained in Section 4.3.5.4. To compensate for the differences in the sessions' duration, the number of occurrences of each behavior was normalized by dividing it by the session length (in minutes). The results presented in Section 4.3.6.4 refer to the use of the data collected from the interaction using Cololo to describe the exchanges between the child and the therapist..

4.3.6.1 Effect on Manipulation

Figure 4.34 shows the number of manipulations per minute for each participant during the sessions with two-sided lighting and one-sided lighting, counted by the human coders (Figure 4.34A) and the Cololo devices (Figure 4.34B). In the case of the human analysis, the number of total manipulations refers to the sum of the number of successful and failed attempts of turn-taking.

Markers and error bars in Figure 4.34A represent the average of the three coders and maximum/minimum values, respectively. For the results obtained by the human coders, a one-way repeated measures ANOVA was applied to investigate the statistical difference in the number of manipulations between the two-side lighting rule and the one-sided lighting rule. The results indicate that no significant difference was observed regarding: the rule factor ($F(1,9) = 0.03, p > 0.05$, $\eta^2 = 0.003$), and the interaction between the rule factor and the coder factor ($F(2,9) = 0.83$, $p > 0.05, \eta^2 = 0.16$). The result from Cololo data shows that for all participants there is a slight decrease in the number of manipulations during the sessions with one-side lighting.

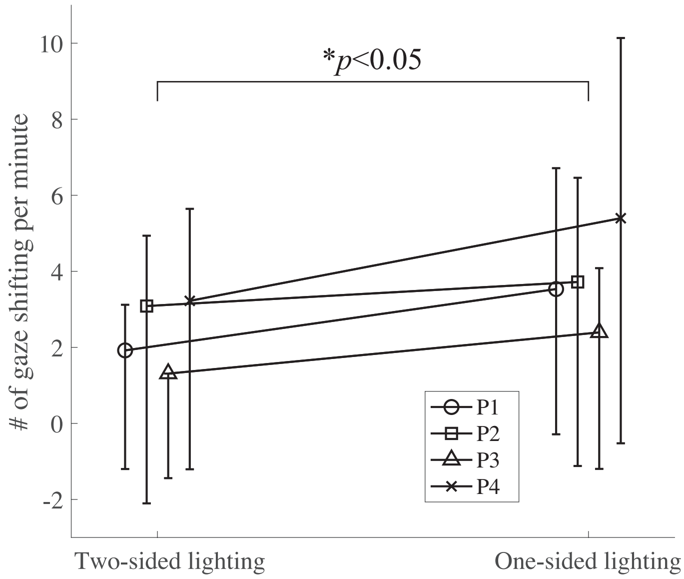


FIGURE 4.35: Average and range of the number of gaze-shifting per minute counted by three coders (C1, C2 and C3).

4.3.6.2 Effect on Gaze Shifting

Since gaze shifting is not as straight-forward for human coders to identify as manipulation is, Fleiss' Kappa was calculated to evaluate the inter-rater agreements before the analysis. The Fleiss's Kappa is a typical statistical measure to assess the agreement among more than two raters. Since there is no ground-truth data for gaze shifting behavior, Fleiss's Kappa was calculated to evaluate the reliability of human coders' observation. The 75th, 50th and 25th percentile of the distribution of Fleiss's Kappa values were 0.43, 0.35 and 0.1 respectively. According to this result, in most of the sessions, the result from three coders showed moderate agreements.

Figure 4.35 shows the number of gaze shifting instances per minute for all participants counted by human coders. Markers and error bars represent the average of the three coders and maximum/minimum values, respectively. Two-way repeated measures ANOVA was applied and a significant increase of gaze shifting in one-sided lighting compared to two-sided lighting rule was observed ($F(1,9) = 30.94$, $p < 0.05$, $\eta^2 = 0.775$). The interaction effect between factors was not observed ($F(2,9) = 2.043$, $p > 0.05$, $\eta^2 = 0.312$).

4.3.6.3 Effect on Failed Attempts of Turn-taking

The results of the number of failed attempts (the child manipulates the toy during the therapist's turn) per minute are shown in Figure 4.36. Figure 4.36A,B represent the results from human coders and Cololo data, respectively. It was represented as the ratio of failed attempts to the total of manipulations of the toy. Markers and error

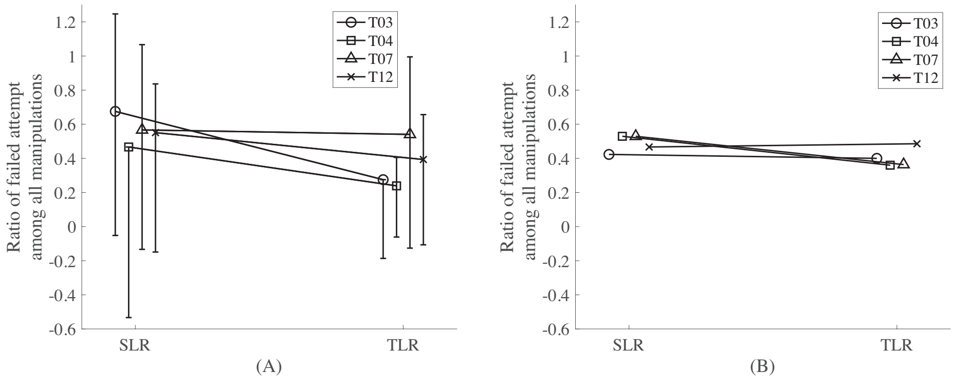


FIGURE 4.36: (**A**) Average and range of the failed attempts ratio among all manipulations identified by three coders (C1, C2 and C3), (**B**) Number of events detected by Cololo.

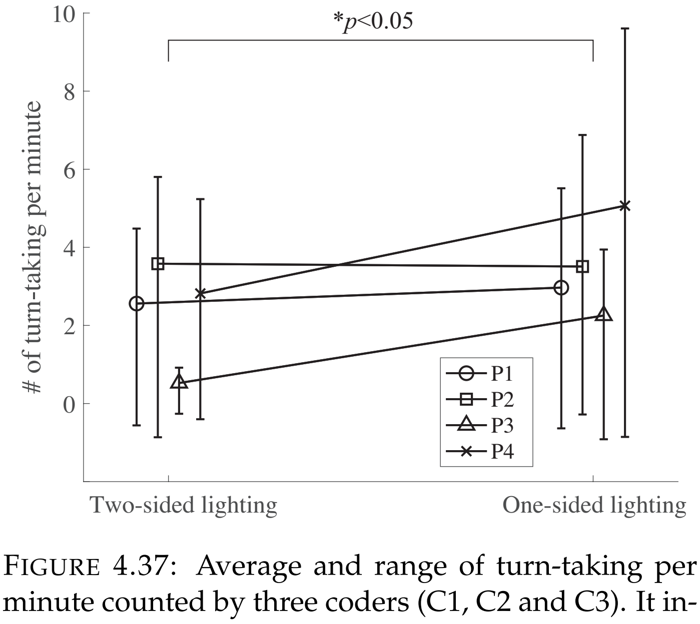


FIGURE 4.37: Average and range of turn-taking per minute counted by three coders (C1, C2 and C3). It includes the data from interventions with Cololo with the two modified interactions rules.

bars in Figure 4.36A represent the average of the three coders and maximum/minimum values, respectively. A significant reduction in failed attempts for the one-sided lighting rule was observed by Two-way repeated measures ANOVA ($F(1,9) = 7.145$, $p < 0.05$, $\eta^2 = 0.443$), without the interaction effect between factors ($F(2,9) = 0.426$, $p > 0.05$, $\eta^2 = 0.086$). This reduction could also be observed in the analysis made with Cololo data from most of the participants (Figure 4.36B).

Finally, it was evaluated the effect of different rules on the number of turn-taking behaviors observed during the interventions. Results presented in Figure 4.37 show the number of turn-taking instances per minute counted by the coders for the four sessions with both modified rules. Markers and error bars represent the average of the three coders and maximum/minimum values, respectively. Two-way repeated measures ANOVA was applied and a significant increase of turn-taking with the one-sided lighting rule compared to the two-sided lighting rule was observed ($F(1,9) = 10.511$, $p < 0.05$, $\eta^2 = 0.539$). The interaction effect between factors was not observed ($F(2,9) = 2.341$, $p > 0.05$, $\eta^2 = 0.3412$).

4.3.6.4 Turn-Taking Analysis by Cololo

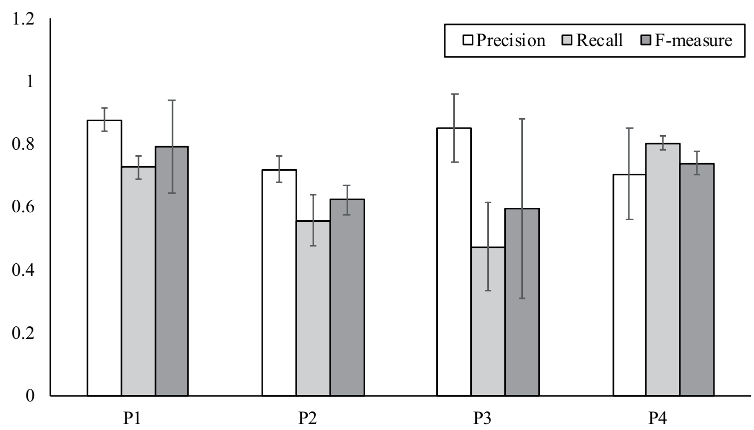


FIGURE 4.38: Agreement between the system and the human analysis (C1, C2 and C3). The data from the toys was filtered using the *T*-value.

Figure 4.38 shows the average of the analysis of precision and recall between the data from the coders (C1, C2, and C3), and the data from Cololo filtered with the *T*-value described in the Section 4.3.5.5. Taking the data from one participant's session as a sample, we overlapped the data from Cololo with the human analysis (Figure 4.39). Each mark represents the toy's manipulation by the participant and the therapist after being filtered using the selected *T*-value, and the shadowed areas represent turn-taking instances (pattern A and B) observed by one of the coders. Exchanges of messages (two marks

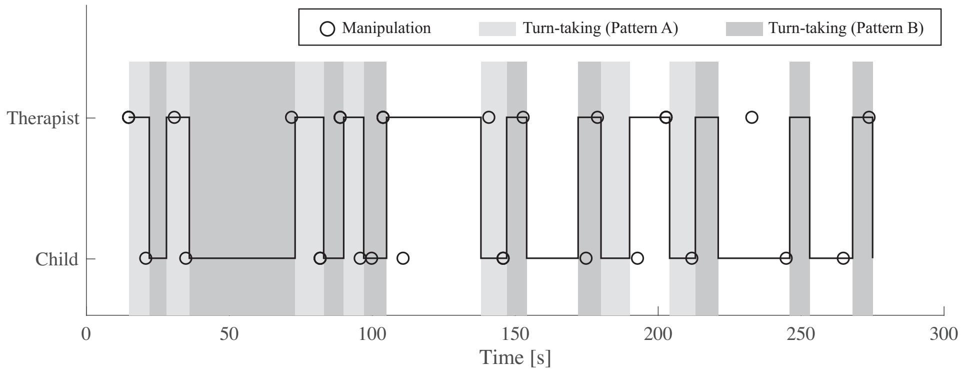


FIGURE 4.39: Marks represent the filtered data from the toys, and the shadowed areas the turn-taking (Pattern A and Pattern B) observed by one of the coders. This is an example of the results of the analysis on which it can be seen how the sensor's data overlap the turn-taking intervals indicated by human analysis.

from two devices) that overlap within the areas can be considered as a moment when Cololo could indicate turn-taking as observed by the coder.

Chapter 5

Discussions

Through this research, we have explored an alternative for conveying cues with a similar role as those of social signals. We started by defining target populations that can benefit from this approach. Then we defined the requirements for the devices: paired devices with a physical embodiment, that sense haptic cues and convey simple visual and haptic cues.

5.1 Approach of facilitating social signaling using paired devices with visual and haptic cues

By following the proposed approach, we presented three possible solutions. Two of them, Macaron and then Pepita, were designed for facilitating social signaling in remote communication. The third one, Cololo, was designed to be used for facilitating the exchange in turn-taking interventions for children with ASD. The discussion presented in this section is meant to verify the hypotheses introduced in Section 1.2.

5.1.1 H1: Social signals can be represented by visual and haptic cues using paired devices for minimal communication

This hypothesis refers to the capability of the proposed paired devices in sensing and conveying signals. Through the evaluation of the three proposed interfaces, we could understand the potential and limitation of our approach. In remote communication, it was necessary to detect hugs as a touch-based gesture with affective connotation, and two different sensors were designed.

The first one was implemented on Macaron. To investigate the variability of a hug, we avoid verbally instructing the participants about the way to perform the gestures, and we did not include feedback from Macaron to indicate when the hug was detected. As shown in the table 4.1, a threshold of 0.15 had a good performance for this application. We aimed to obtain a highly responsive system by detecting hugs when the participants were doing so, but at the same time, a higher recall value is related to the ability to distinguish hugs from other gestures. We considered that a precision value of 0.8 (or higher), is appropriated for this application.

With this study, it is important to point out the implications of the two types of errors. False negatives are easier to address considering that in a real scenario using Macaron, the system will provide feedback (colored light patterns and vibrations) to indicate when the hug is detected and thus, a message is sent. The user can visualize this and adapt his or her behavior until the system detects the intention. However, it is important to consider as well the negative implications; a too unresponsive system might lead to loss of motivation or trust from the user. The other type of error is related to the number of false positives. An example of false positive can be found on figure 4.1, on which in two opportunities the system detected a hug when the participant was pressing on the upper and lower areas of the cushion. The implications of this type of error can be related to messages sent when the user did not intend to. From this study, we concluded that each of these errors should be analyzed and addressed carefully, to find a threshold to allow the system better perform for the selected application.

The second hug detection approach required a bigger plastic case to contain a projector and a mobile phone. For this reason, a new sensor was designed to adapt to the new interface shape. This sensor measured applied force using conductive material, and a sensor with simple structure was designed. We tested the detection of hugs among 2 other kinds of touch-based interactions. We chose to test pressing on the sides, top, and bottom because these are common ways to handle a spherical robot, and they are more similar to hugging. Even though the sensor could not provide the level of detailed tactile information that would be possible using sensors with a higher resolution, the results proved that the sensors worked well for the intended interaction (Figure 4.23). The highest number of false positives was obtained when pressing on the upper and lower sides. We expected this result because the poles of the sphere contained eight sensors in an area that could be covered by the hand.

To better evaluate the accuracy of the system, precision, recall, and F-value were calculated. The results showed that the system is accurate and sensitive in detecting hugs. However, it is important to understand the potential implications related to failures in the detection. One type of error occurs when the system fails on detecting an actual hug (false negative). Among the two types of errors, this is the easiest to address because users can realize when it occurs based on the lack of response from the system (e.g., not receiving feedback right after hugging), and then correct it by hugging again. The other type of error is when different types of gestures are incorrectly detected as hugs (false positive). Unintended hug detection should be avoided as it might impact the user's credibility on the system. This sensor was designed especially for this interface, with the criteria of using soft materials and reducing the sensor complexity as much as possible. We are satisfied with the results so far as we were able to detect hugging using only eight electrodes while maintaining high performance. These values could be improved by adding different kinds of sensors (like sound or temperature) to complement the pressure sensor readings. A limitation of the current approach is related to the way the sensors are mapped into a two-dimensional polygon. This makes it impossible, for example, to detect hugs if the robot is tilted 90 degrees. Pepita was designed with features such as eyes, and decoration on the top, to give the user an idea of the orientation to hold the device, but solutions in the future should tackle this issue. Moreover, for the current implementation, we only used spatial information at each time frame (the area of the polygons generated when the sensors are pressed), and the temporal information was neglected. Therefore, using a model which involves temporal features or past events such as Hidden Markov Model or Recurrent Neural Network would improve the accuracy of detection, and it should be considered for future implementations.

With Pepita we worked with another types of touch based gesture besides hugging. This second gesture was supposed to convey a negative meaning, and we chose tail pulling (Section 4.2.5.2). The sensor was designed to work with the structure and material of the tail. The magnet and the hall effect sensor were placed at a distance that resulted in clear signal changes when the tail was being pulled. To test the detection performance, we chose three other gestures that could be wrongly detected as pulling. The results of the sensor performance show that the detection was accurate enough for the proposed application (higher than 0.8). This approach offers a feasible solution with a sensor of simple structure that supports detection of pulling.

Beside sensing cues from the user, the devices were designed to represent this action using visual and haptic cues. With the study of Macaron, we explored the participant's perception of Macaron as a communication device, based on three different communication partners. Specifically, we tried to verify if the user was able to transmit and receive intention through the device when communicating remotely. This perception will be influenced by the interaction conditions and the way the device delivers feedback to the user [82] [9]. Figure 4.4 shows a small tendency on which the human condition obtained higher scores on the six items. Even though we worked with a limited sample, we could test the structure of the experiment and verify that the interaction rule worked to collect the information necessary to investigate the effect of the device further.

Regarding improvements for the interaction rule, we observed that the levels of a hug (increased frequency of blinking according to the length of the hug were not clear for the participants. Therefore the interaction rule was simplified to represent the following information:

1. Red blinking lights: hug detected
2. Blue blinking lights: message received
3. Multicolor lights: sync hugs

Dynamic states were simplified by removing the levels, and static colors and vibration patterns were left unchanged from the rule described in section 4.1.2.3. Besides understanding the characteristics that influence the user's perception of the rule, it was observed that users could perceive the partner's intent using only visual and haptic cues. This is based on the timing of the message and the ability to synchronized messages. Participants reported that the predictability of the echo condition strongly influenced on the partner's perception (humanity, intent or distance). Moreover, a higher number of synchronized messages was related to the partner's humanity. We understood by this that different feedback patterns affected the user's perception of the partner's intent and humanity.

In the case of Pepita, to limit the robot body to simple features, we combined it with projected avatars, which enhanced its communication capabilities without increasing the complexity of the hardware. For this reason, we evaluated the use of a small projector embedded in the robot body to convey the affective states related to two emotions. In this study, we explored the benefits of using a projector compared to abstract visual elements made by colored lights. The results related to the perception of colored lights patterns were in line with those of a previous study [32]. The participants perceived a happy state when the robot was blinking yellow lights and a sad state with blue lights (Table 4.4). These results also showed that avatars were more effective conveying affective expressions compared to the feedback made with colored lights. During the study, the robot was not moving or showing any behavior—only visual elements were being displayed. For a robot with limited features and an inexpressive face, there are few options to convey affective expressions visually. Colored lights are commonly used and explored, as they are a simple and effective solution, but the perception of these abstract representation of emotions can be different from one person to the other. Projected avatars with affective expressions have the potential of reducing the user's misinterpretation by providing a clear message.

Regarding the robot embodiment, the results showed that most of the participants could perceive that the avatar was reflecting the expressions of the physical robot as only one entity. This exploratory study showed that, even though the avatar was displaying the affective expressions, since they were projected from the robot, it was possible for the participants to perceive them as part of the robot's behavior instead of an external agent. Based on these results, we are encouraged to keep exploring the possibility of using images projected from a robot to convey its expressions in a robot-like manner.

Another evaluation of the capabilities of paired devices in sensing and conveying cues was performed with children with ASD using Cololo. The analysis explored the capabilities of describing turn-taking using only the data of the toy manipulation. To achieve this, we selected a interval Tmin–Tmax of 3 to 13 s (Figure 4.33), and with this value the data from Cololo was filtered, leaving only the exchange of messages within Tmin and Tmax. This filtered data set was tested by comparing it with the analysis of different coders (C1, C2, and C3) obtaining in average an agreement of 0.72 between the human analysis and the analysis with Cololo (Figure 4.38). These results have different implications. The data from the toys cannot describe turn-taking, but it can be used as an indicator of turn-taking. For example, the plot in Figure 4.39 shows that the data from the toys can follow the human analysis to a high degree, but not all the turn-taking indicated by Cololo was also indicated by the human coder. To describe turn-taking-like behavior more accurately by computers, one possible approach is the combination of different sensors. However, with this approach, it was possible to describe children's behavior during play activities that involve the manipulation of toys. This type of data is: (1) Difficult to obtain by traditional evaluation methods such as video coding, and (2) Valuable to understand how children play with the toys. While the information obtained from interactive systems such as Cololo will not replace the therapist's judgment, providing quantitative measurements with good reliability will facilitate the understanding of children's responses in different social situations.

These results showed that the proposed paired devices were capable of sensing users cues (hugging, manipulations of the toys). These actions were translated into visual cues that: 1) Convey affective meaning, 2) Convey the partner's intent and 3) Guide the user in the exchange. With this, we understood that H1: Social signals can be represented by visual and haptic cues using paired devices for minimal communication, was verified.

5.1.2 H2: Social signaling can be facilitated by the communication through visual and haptic cues

With this hypothesis, we try to understand if these cues had the role of social signals or not, based on the observed effect on the users. We observed that, while the interaction with the interface is simple, there are several implications related to the interaction of two users through paired devices. With the proposed paired devices, we are replacing behavioral cues commonly used in face to face interaction, by visual and haptic cues delivered by paired devices. These cues were simple indicators of others' actions. We observed that, in both of the applications targeted by this research, the cues produced tangible changes in the participants. Therefore, we could assume that these simple cues delivered by paired devices, worked as social signals. If this is true, then what is the effect produced by these new type of social signals in human interaction?

The evaluation explored different aspects. In the application for remote communication, we evaluated Macaron. With this study, we intended to understand if the cues delivered by paired devices in remote communication context, worked as social signals. To investigate this, we performed a comparative study with a more traditional type of communication using graphical representations on tablets. With this evaluation, we expected to evaluate the role of the cues, and also point out the benefits of mediated social touch using physically embodied interfaces. Participants reported feeling a higher perceived intimacy on the Macaron condition compared to the Screen condition 4.10. Moreover, the quality of the interaction was also higher in Macaron condition (Figure 4.11). When comparing the reaction to the interfaces, Macaron was rated higher in categories such as "Feeling connected" and "Communicate emotions," features that are desirable for a mediator of affective content (Figure 4.12). The information extracted from the messages exchanged talks about the user's behavior on each condition and the role of the cues. The number of occurrences on the screen condition was as expected, higher than the Macaron condition (Figure 4.13). This is because of the nature of the gestures (hugs and pressing screen with fingers). However, we observed that the duration of time that participants were using both interfaces was very similar; thus, the messages sent by Macaron were made using longer gestures than those sent by Virtual Macaron (Figure 4.14). In both cases, we observed a positive correlation between Macaron and Virtual Macaron, which refers that the communication behavior of each participant did not change according to the interface.

Another component of the analysis of the role and effect of Macaron was made with physiological data. We faced different challenges in being able to analyze the reaction to interfaces using this data. The challenges involve not only analyzing the data, but also interpretation. In terms of stress reduction, indicators such as increase of ST did not show differences between types of embodiment, but it was observed a difference between the Macaron OFF condition and Macaron ON and Screen (Figure E.1) In other words, participants showed to be more relaxed when they were communicating, either by Macaron or Virtual Macaron, reflected as an increased level of ST. HRV can also be used as an indicator of stress levels [91]. The analyzed data from HRV did not show a significant difference between conditions (Figure 4.17) However, as we observed in a study that used HRV to describe emotional state [116], it is necessary to use powerful stimuli, especially strong and negative ones. Another interesting observation was on the EDA data. While participants showed a higher tonic component level during the Macaron OFF condition (no communication), they showed a higher phasic component level when they were communicating either by Macaron or Virtual Macaron, compared to Macaron OFF.

This effect was even stronger on Macaron ON, compared to the screen condition. When participants were communicating with their partner, they displayed, in general, a lower level of arousal (tonic component) but higher arousal as a response to the stimuli from the movies (phasic). This effect is attributed to being able to perceive the partner through the simple cues made by colored lights, and with a stronger effect when these are presented through a physical embodied interface such as Macaron. These results showed that the cues delivered by paired devices worked on changes of behaviors measured by the way participants used the devices, as well as internal state measured by physiological signals. Therefore, these cues worked as social signals in this remote communication scenario.

In the application for children with ASD, instead of comparing different embodiment, we compared different interaction rules (cues patterns). From the different behaviors exhibited during a turn-taking game, Cololo was able to sense only the manipulation of the toys. Based on the timing of these manipulations, the exchange between the child and therapist on turn-taking interventions was identified. To guide the child during turn-taking, instead of relying solely on common social signals using in face-to-face scenarios, Cololo delivered cues using colored lights and vibrations delivered based on different interaction rules.

Initially we assumed that lighting up both devices (lights do not indicate the turn-holder), will elicit a higher number of manipulations because of its relatively higher amount of visual stimuli. As shown in Figure 4.34A, no significant difference was found between the effect of the two-sided lighting and one-sided lighting on the number of manipulations. Based on the results, our assumption was not supported, as coders did not observe a significant difference between the two conditions in the number of times the participants manipulated Cololo. However, the data from Cololo showed that all the participants reduced the number of manipulations during the session with the one-sided lighting rule. This difference between the human analysis and the system is understandable, as Cololo register an event each time the sensor is activated (by tilting the toy) regardless it is a toy manipulation as it was previously defined. The data from Cololo while lacking context, it provides an image of the level of engagement of the child to play with the toy. It was not clear that the level of engagement to play with Cololo was due to a learning outcome or as a consequence of the novelty of the toys, which is a limitation of the study. The current results compare the effect of each rule on the manipulations of the toys, which for the therapist is an indicator of "engagement to play with the toy".

We also assumed that lighting up one device (lights indicate the turn-holder) will reduce the number of the child's manipulations during the therapist's turn (i.e., reduce the number of failed attempts of turn-taking). This assumption was supported by the results presented in Figure 4.36A. We consider that since children with ASD tend to show a strong interest in visual stimuli, illuminating only one device at a time was effective to lead the gaze shifting. This could be related to the reduction of the number of failed attempts in the sessions with the one-sided lighting rule. Because in the one-sided lighting condition, the participant looked at the therapist's device more often than during the two-sided lighting condition, as it was described by the results on Figure 4.35, this was helpful to clarify who was holding the turn at a time. A similar tendency was displayed by the data collected from Cololo (Figure 4.36B). Occurrences of failed attempts of turn-taking can indicate both that the child is engaged enough to play with the toys but can also be a negative indicator of a child not being able to wait for his partner to take the turn. The human analysis and the device analysis followed a similar trend, a reduction of failed attempts at turn-taking under the one-sided lighting rule, that implies the reliability of the effect.

To facilitate the number of turn-taking instances during the sessions with Cololo, both increasing the number of successful attempts of turn-taking and reducing the failed attempts are essential aspects. We initially assumed that more colored-lights feedback would elicit more manipulation based on the insights from our previous study [113]. However, in the sessions with one-sided lighting, the participants manipulated the device as much as they did during the sessions with two-sided lighting. Moreover, indicating the turn-holder by both lights and movements was beneficial for reducing the failed attempts which also resulted in an increased number of turn-taking instances observed by the human coders (Figure 4.37). For this reason, we assumed that lighting up one device (lights to indicate the turn-holder) would lead to a higher occurrence of turn-taking by the child. The results in Figure 4.37 supported this assumption, and overall the one-sided lighting rule was more effective than the two-sided lighting rule for facilitating turn-taking behavior of children with ASD. On interventions with Cololo, switching off one of the toys to make emphasis on the turn-holder resulted in more turn-taking behaviors. This effect was also reflected on the previous results, as the one-sided rule elicited less failed attempts of turn-taking (the child waited for his turn without manipulating the toy) and more gaze shifting (the child looked at the therapist during the therapist's turn).

The results from these studies point out to the same direction: that employing paired devices with visual and haptic cues it is possible to facilitate social signaling. This is reflected in changes such as inner states (measured by self-report and physiological data) or changes of behaviors (frequency of occurrences). By this, it is concluded that H2: Social signaling can be facilitated by the communication through visual and haptic cues, is verified.

5.1.3 H3:The effect of the represented social signals is conditioned by interaction factors: timing/order, interface and feedback

We assume that there are parameters that condition the effect of the cues as social signals. On the study with Pepita, we explored characteristics of the interface such as the appearance on the user's perception and acceptance. The first part of the evaluation consisted of a general evaluation of different characteristics of the design of Pepita. Regarding the huggable aspect, we compared Pepita with three other huggable robots (Table 4.3). Among them, Probo ([97] was designed with a caricatured appearance, The Huggable [98] looks like a teddy bear, and The Hug [15] has a simple appearance with big arms and no facial features. We could see that, even though Pepita has no arms, it had a score similar to the other three interfaces. However, the teddy bear appearance effectively evoked the feeling of being huggable, which was reflected in higher scores. Evaluating huggable aspects using only photos can lead to limited results since the impression can change when directly touching the robot. Because the role of the appearance was being investigated, we found appropriated to use photos to compare the proposed design with other huggable robots. In this study, we concluded that common features in huggable interfaces (e.g., open arms) are not essential, but making the robot look familiar based on an already existing idea of huggability could be beneficial to make the robot more appealing to be hugged.

Then, methods for conveying the robot's affective expressions were evaluated. We compared robots that used mechanical facial expressions (Probo [100] and Zeno [99]) and two robots that used a display to manage the robot's facial expressions (Pepita [102] and Buddy [101]). The results showed that there was a significant difference between Pepita and the other three robots. Participants favored both of the robots that presented expressions using a display over the two that used mechanical facial expressions. However, Buddy's expressions were significantly more acceptable than Pepita's expressions. Embedded displays (Buddy) can be used to present facial expressions as a part of the robot's body, while projected displays open the possibility of designing new ways of interacting with robots. Moreover, projected images can be shared by different users and have a strong visual impact. On the other hand, robots with mechanical facial expressions have specific applications in which it is necessary to make robots that imitate a human's behavior to a higher degree. In this study, we found that, regardless of the application, using avatars projected from the robot's body was considered to be as acceptable as other more traditional ways of conveying a robot's affective expressions.

Regarding the perception of the robot's appearance, the results in Figure 4.21 show that, in general, the participants gave Pepita positive ratings in the following four aspects: "kind", "friendly", "pleasant", and "nice". On the other hand, a tendency toward the attributes "artificial", "machine-like", and "fake" reflected that participants perceived Pepita as an artificial agent. These results encourage us to use a similar design in a future study given that rating Pepita as an artificial character is not necessarily undesirable. Based on these results, we concluded that it is important to design this type of character to have a robot-like means of expression that matches the expectations of the user. Moreover, to obtain more insights related to which factors positively and negatively impacted the perception of Pepita, we asked the participants an open-ended question. Their positive comments expressed a general acceptance of the images projected from the robot as an alternative means of representing the robot's expressions. We observed that the huggable aspect was mentioned by some participants to be positive, but not as strong as other features. This outcome becomes more evident after reading some of the negative comments that pointed out that the robot did not look huggable enough. For this reason, we plan to work on the robot's softness to try to reduce this negative aspect. Another critical factor that needs to be improved in this design is the negative influence of the eyes. The majority of negative aspects included the appearance of the eyes. Because Pepita has a projector placed in one of the eyes, one eye lights up, and one is off. This characteristic was perceived to be unnatural because it seemed that one eye was "not working" or "dead." Thus, we plan to find a different location for the projector that does not negatively affect the perception of the robot's appearance.

The study with Macaron pointed out different features. As it was mentioned before, participants were able to perceive the partner's intent through the cues delivered by Macaron. However, the perception of intent was conditioned by aspect such as timing or order of the messages. Participants reported that the predictability of the echo condition strongly influenced on the partner's perception (humanity, intent or distance). Moreover, a higher number of synchronized messages was related to the partner's humanity. We understood by this that different feedback patterns affected the user's perception of the partner's intent and humanity. Besides the timing of the cues, we observed that differences in the embodiment representation of the paired devices could also affect the perceived intimacy and social aspect of communication, measured by participants' report. This positive effect was also reflected on the participant's inner state, measured by physiological signals, such as a higher level of relaxation (based on ST, SCL), or a stronger response to emotional stimuli (based on SCR). Additionally, for each participant, the number of occurrences and duration of the messages did not differ between Macaron and Virtual Macaron. This was reflected in the positive correlation observed in the data tendency. Based on this it was concluded that the effect on the user was not due to communication styles but because of the properties of the interface.

In the case of Cololo, we understood that different feedback pattern resulted on different observed behaviors. We could observe that the cues delivered by paired devices could guide children during the exchange in interaction. In the context of collocated interaction during a game, cues delivered by Cololo worked as social signals, resulting in changes in the participant's behavior (an increase of desirable behaviors or reduce negative ones). Each interaction rule resulted on different observed behaviors, in terms of manipulations of the toys, gaze shifting, turn taking and failed attempts of turn-taking. Based on this we identified simple guideline for paired devices for facilitating the exchange in communication. With this study we have pointed out the benefit of paired devices structure, using visual and haptic cues as social signals, to facilitate the exchange in communication between the child and therapist. For children with ASD that also has difficulties with verbalization, this proposed system worked to facilitate the exchanges between them and the therapist, and also to describe aspects of their play behavior. This paired devices configuration using simple visual cues could be incorporated into different styles of games that involve taking turns. Although limited interaction rules and their effect on turn-taking behavior was studied with Cololo, the findings suggest an insight to design robotic toys for facilitating turn-taking. To facilitate a targeted behavior, it is necessary to design how to display the visual cues carefully. We consider that the proposed interaction rules for paired devices can be applied to other paired devices that communicate and provide feedback using colored lights and movements.

5.2 Limitations and challenges encountered

Different limitations of social signaling through paired devices with visual and haptic cues were pointed out through this investigation. It was observed that different from face to face interaction, remote communication mediated by technology comes together with particular restrictions. This is because sensors are not 100% accurate, which means that the encoded information from the paired devices is already different from the way we interact when we are in front of each other. This means that the decoded message represented by visual and haptic cues does not represent all the sender's behavioral cues. According to each application, designers should carefully address these issues to avoid adverse effects on the user's interaction.

The selection of paired devices as a feasible solution was made based on having two identical interfaces that share the same affordance might be beneficial as users can receive cues based on both, actions made by themselves and by the partner's actions. For children with ASD, this is a desirable feature. Paired devices that deliver cues with the role of social signals can guide the child in a similar way that people do in face to face interaction. During face to face interaction, people commonly observe others. They can understand based on social signals, when is the moment to let others "take the turn" in the interaction. Considering that this task is complicated for some children, paired devices can support a similar exchange, by delivering cues that can engage children in the exchange with others. Once children look up to others and play with others, even if simple lights guide them, the opportunity of social interaction is opened. The barrier of social interaction is then breached.

Similarly, we understood that this characteristic of paired devices as mediators is also beneficial in remote communication scenarios. In this context, even though they are not able to see each other, users can imagine what the partner is doing as the devices are identical. When a user sees the device delivering cues because a message was received, it is possible to imagine that the partner is hugging on the other side of the line. Even though the message has minimal information, it is possible to set it in context by understanding the characteristics of paired devices. However, this can be affected strongly by the relationship between the two persons' communication. In the experiment, we created a situation on which we elicited emotional responses by films, and we provided the opportunity to communicate or not, through Macaron. We found that participants with a long history of relationship attributed different meanings to the messages according to the content of the movies. Among them, participants sent more messages when they found a movie they watched together in the past, or when they knew the partner liked or disliked a particular content. The message changed from being an indicator of action, to be a message with much more meaning. This new attributions to the messages are hard to be made by people are not familiar with each other.

Regarding the evaluation method used for the study with autistic children, further studies to evaluate the effect of feedback delivered by paired devices should include larger samples. Moreover, they must seek to eliminate ordering and novelty effects through blocked and longitudinal study designs. In the case of the evaluation of Macaron, we used different tools to understand the user's experience. Among them, questionnaires needed to be explained at the beginning of the session due to the use of relevant keywords. When recruiting participants for this experiment, we made sure that everyone could understand verbal and written English instruction. Even though these questionnaires were taken from the literature, the long sessions and the limited time to answer came with difficulties for the participants to answer the questions.

This was one of the motivation to try to understand the user experience from another perspective, and the analysis with physiological data was included. Again, through the revision of the available literature, we observed that there are many limitations to this approach, which should be carefully addressed. For example, skin conductance can be affected by different environmental factors, like if the participant is hungry or not, if he or she is struggling with work, the room temperature, health condition of the participants, sensors accuracy, among many others. Based on this and considering that experimental studies allow us to control variables, we addressed some of these environmental factors for better performance. However, we believe that this is far from perfect as we are merely using this as an evaluation tool, instead of understanding the potentials of this measurement. To increase the validity of this method, we combined the analysis of different physiological signals extracted from the same wearable sensors.

Chapter 6

Conclusions

This research started with an understanding of the problems that affect social signaling in human communication. From the literature review, we extracted two different scenarios on which this approach might be beneficial: on remote communication and interventions for children with ASD. While the role of the device is different in these two scenarios, fundamentally the human problem is the same: limited amount/kind of social signals. For applications in remote communication, we proposed Macaron as a possible solution. We implemented a method for hug detection, and from the results, it is important to point out the implications of the two types of errors. False negatives are easier to address considering that in a real scenario using Macaron, the system will provide feedback (colored light patterns and vibrations) to indicate when the hug is detected, and thus, a message is sent. The user can visualize this and adapt his or her behavior until the system detects the intention. However, it is important to consider as well the negative implications; a too unresponsive system might lead to loss of motivation or trust from the user. The other type of error is related to the number of false positives. An example of false positive can be found in figure 4.1, on which in two opportunities the system detected a hug when the participant was pressing on the upper and lower areas of the cushion. The implications of this type of error can be related to messages sent when the user did not intend to. From this study, we concluded that each of these errors should be analyzed and addressed carefully, to find a threshold to allow the system to better perform for the selected application.

Regarding the evaluation of the perception of the cues delivered by Macaron, during the interview, participants expressed they perceived that three partners were different and that the partner felt more human during the first task (manual). Two participants doubted that the second partner was human (echo) since it was too much predictable. Regarding the third condition (random) while it was not predictable, they perceived it as less responsive thus, with less intention. When we asked the reason, two participants answered because "the partner did not try to synchronize the hugs". This aspect of the interaction rule is of particular importance as it seems to affect the perception of the partner. The results of this test showed that the way the feedback is delivered could affect the user's perception of the device. The proposed three conditions for the three partners were successfully perceived different, which was reflected in the six items and interview. With this methodology, we plan to further investigate the user's perception of Macaron during remote communication tasks. Future work includes a similar study with a larger sample and with a counterbalanced order of feedback conditions. After we confirm that the proposed interaction rule is clear enough to allow people to communicate through Macaron, we will be ready to explore its effect on the human-human remote communication of affective messages.

The last evaluation of Macaron explores the effect on social signaling in remote communication. We compared Macaron with a virtual representation and observed that:

1. In terms of the user's self-report, the interaction was rated more positive, social, personal, and in general more positive, compared to the interaction through the screen. These results go in line with the questionnaire to evaluate the reaction to the interface. Participants rated Macaron higher in items as "feeling of connectedness" or "communicate emotions," features that point out the affective connotation of the communication thought Macaron.

2. In terms of user's behavior, we observed that even though participants sent a higher number of messages using the screen interface, this did not reflect in a more positive experience. The positive features mentioned by the users by the questionnaires can then be attributed to the type of interface and not the number of messages.

3. In terms of the user's physiological state, data from ST showed participants were more relaxed when they were communicating while watching movies. This phenomenon was reflected on higher ST observed on Macaron ON and screen compared to Macaron OFF. Moreover, the analysis of EDA showed that participants had a stronger SCR under the stimuli, compared to the other two conditions.

Based on the results from the study with Macaron, we proposed some modifications for the design of mediators for remote communication. Pepita was designed to sense and convey affective information. The current implementation senses two types of touch-based gestures, hugs as a positive message, and pulling the tail as a negative message. The system translates these actions into visual feedback made by projected avatars, designed to convey a positive affective expression such as happy, and a negative affective expression such as sad. This report includes contents such as design criteria, system overview, and touch-based gestures' recognition methods.

The results offer some guidelines to improve the current proposed design. The hug detection method had a good performance for this application, but comments from the participants pointed to the necessity to improve the huggable aspect of the robotic device. We compared the proposed robot with other huggable robots and concluded that while a simple appearance with no arms is acceptable, making the robot look similar to other familiar huggable elements (like stuffed animals or cushions) can be beneficial to make the robot appealing to be hugged. For this reason, future work involves working on the huggable aspect to elicit and support natural hugging behavior. The proposed approach using an array of sensors made of foam could benefit from the addition of an extra layer of cushion between the case and the sensor to increase the softness.

Regarding the appearance, the caricatured features were found to be funny and acceptable in general, but the projector's lamp in one eye had a strong adverse effect. From this, we plan to relocate the projector in a different part of the robot. Projected avatars to express the robot's affective expressions were clear and effective conveying information. Moreover, participants perceived the avatars as part of the robot (as one entity). In the current design, we worked with a hug-driven change of states, and thus minimal expressions were implemented. Considering that the main benefit of projected avatars is different information that can be easily displayed, more interaction rules leading to more expressions are a good starting point for future works. Before evaluating its effect on remote communication contexts, it was necessary to implement a robot based on robust design criteria to explore which of the characteristics are appropriate for a social mediator of affective messages. Future work will involve the use of two identical robots to mediate communication using touch-based gestures as the input and projected images as the output, in order to explore its role in enhancing the sense of co-presence compared to traditional telecommunication devices.

For applications for developmental disorders, we initially investigated how robotic toys are being designed and used in ASD interventions. We separated them according to two main features: toys for stimulating behaviors and toys for describing behaviors. We considered these two characteristics as the primary functions of the proposed system Cololo. The purpose of the study was: (1) to evaluate how the devices influence children's turn-taking behavior by comparing the effect of two different interaction rules, and (2) to explore the potential for and the limitations of describing children's play behavior in turn-taking interventions using quantitative data from the robotic devices. We followed a similar approach to the ones using multiple robotic toys for turn-taking [23], but it was simplified into a system using paired devices. The devices work with interaction rules meant to help children to identify turns, which can be only achieved by a paired/multiple devices configuration. The approach of using paired devices benefits from being able to deliver sensory feedback based not only on children's actions but also on the therapist's actions. This creates new opportunities for interventions. Then, we described the hardware characteristics and introduced simple interaction rules for facilitating turn-taking behaviors. Four hypotheses were proposed based on the relationships between the interaction rules and children's behaviors related to turn-taking.

The evaluation consisted of a pilot study to observe and decide which modifications of the previous rule were necessary. This was followed by an experimental study to compare the effect of the two modified rules: the two-sided lighting rule and the one-sided lighting rule. Moreover, we explored the use of Cololo's data as an indicator of turn-taking behaviors. Results showed that regarding the number of manipulations, there was no significant difference between two-sided lighting and one-side lighting. However, the number of failed attempts of turn-taking and gaze shifting were reduced in the sessions with one-sided lighting. These results are related to the increased number of turn-taking instances in sessions with one-sided lighting. Following this, we evaluated the use of the data from the toys' manipulation to indicate those exchanges that have a higher probability of being classified as turn-taking. We obtained an agreement of 0.72 between the human coder and the system. These results indicate the potential for and limitation of using this approach to describe children's play behavior in turn-taking interventions.

This investigation ended with the discussion of the different factors that affected social signaling when we used paired devices with visual and haptic cues in minimal communication. By verifying the three proposed research hypotheses, we found that it is possible to facilitate social signaling in human-human communication using paired devices with visual and haptic cues. We observed that even if the message contains little information (an event happened or not), by just including it or complementing interaction by it, we could observe changes in human's behavior and states. Considering the little content of the message, this is highly dependent on the context of the interaction. We conclude that minimal communication is enough for delivering context and intent, by using paired devices with visual and haptic cues.

Bibliography

[1] John B Nezlek, Deborah S Richardson, Laura R Green, and Elizabeth C Schatten-Jones. “Psychological well-being and day-to-day social interaction among older adults”. In: *Personal Relationships* 9.1 (2002), pp. 57–71.

[2] Masako Ishii-Kuntz. “Social interaction and psychological well-being: Comparison across stages of adulthood”. In: *The International Journal of Aging and Human Development* 30.1 (1990), pp. 15–36.

[3] Gillian M Sandstrom and Elizabeth W Dunn. “Social interactions and well-being: The surprising power of weak ties”. In: *Personality and Social Psychology Bulletin* 40.7 (2014), pp. 910–922.

[4] Iris B Mauss, Amanda J Shallcross, Allison S Troy, Oliver P John, Emilio Ferrer, Frank H Wilhelm, and James J Gross. “Don’t hide your happiness! Positive emotion dissociation, social connectedness, and psychological functioning.” In: *Journal of personality and social psychology* 100.4 (2011), p. 738.

[5] Alessandro Vinciarelli, Hugues Salamin, and Maja Pantic. “Social signal processing: Understanding social interactions through nonverbal behavior analysis”. In: *Computer Vision and Pattern Recognition Workshops, 2009. CVPR Workshops 2009. IEEE Computer Society Conference on*. IEEE. 2009, pp. 42–49.

[6] Sebastian Loth and Jan P De Ruiter. “Understanding Social Signals: How Do We Recognize the Intentions of Others?” In: *Frontiers in psychology* 7 (2016), p. 281.

[7] Alessandro Vinciarelli and Alex Sandy Pentland. “New social signals in a new interaction world: the next frontier for social signal processing”. In: *IEEE Systems, Man, and Cybernetics Magazine* 1.2 (2015), pp. 10–17.

[8] Alessandro Vinciarelli, Maja Pantic, and Hervé Bourlard. “Social signal processing: Survey of an emerging domain”. In: *Image and vision computing* 27.12 (2009), pp. 1743–1759.

[9] Kenji Suzuki. “1bit communication network and devices for intention exchange”. In: *Proc. of Int. Conf. on Computational Intelligence, Robotics and Automation, 2005*. 2005.

[10] Wilma A Bainbridge, Justin W Hart, Elizabeth S Kim, and Brian Scassellati. "The benefits of interactions with physically present robots over video-displayed agents". In: *International Journal of Social Robotics* 3.1 (2011), pp. 41–52.

[11] Robert Kowalski, Sebastian Loehmann, and Doris Hausen. "cubble: A multi-device hybrid approach supporting communication in long-distance relationships". In: *Proceedings of the 7th International Conference on Tangible, Embedded and Embodied Interaction*. ACM. 2013, pp. 201–204.

[12] Johan Larsson. *Tangiball: Designing a Product for Conveying Presence*. 2014.

[13] Thomas Visser, Martijn H Vastenburg, and David V Keyson. "Designing to support social connectedness: The case of Snow-Globe". In: *International Journal of Design* 5.3 (2011).

[14] Fotios Papadopoulos, Kerstin Dautenhahn, and Wan Ching Ho. "Exploring the use of robots as social mediators in a remote human-human collaborative communication experiment". In: *Paladyn, Journal of Behavioral Robotics* 3.1 (2012), pp. 1–10.

[15] Carl DiSalvo, Francine Gemperle, Jodi Forlizzi, and Elliott Montgomery. "The hug: an exploration of robotic form for intimate communication". In: *Robot and human interactive communication, 2003. Proceedings. ROMAN 2003. The 12th IEEE international workshop on*. IEEE. 2003, pp. 403–408.

[16] Allan Fong, Zahra Ashktorab, and Jon Froehlich. "Bear-with-me: an embodied prototype to explore tangible two-way exchanges of emotional language". In: *CHI'13 Extended Abstracts on Human Factors in Computing Systems*. ACM. 2013, pp. 1011–1016.

[17] Kaiko Kuwamura, Kurima Sakai, Takashi Minato, Shuichi Nishio, and Hiroshi Ishiguro. "Hugvie: Communication Device for Encouraging Good Relationship through the Act of Hugging". In: *Lovotics* 1.1 (2013). ISSN: 2090-9888. DOI: 10.4172/2090-9888.10000104.

[18] Yasir Tahir, Justin Dauwels, Daniel Thalmann, and Nadia Magnenat Thalmann. "A User Study of a Humanoid Robot as a Social Mediator for Two-Person Conversations". In: *International Journal of Social Robotics* (2018), pp. 1–14.

[19] Shogo Nabe, Stephen J Cowley, Takayuki Kanda, Kazuo Hiraki, Hiroshi Ishiguro, and Norihiro Hagita. "Robots as social mediators: coding for engineers". In: *Robot and Human Interactive Communication, 2006. ROMAN 2006. The 15th IEEE International Symposium on*. IEEE. 2006, pp. 384–390.

[20] Hisato Fukuda, Yoshinori Kobayashi, Yoshinori Kuno, Akiko Yamazaki, Keiko Ikeda, and Keiichi Yamazaki. "Analysis of multi-party human interaction towards a robot mediator". In: *Robot and Human Interactive Communication (RO-MAN), 2016 25th IEEE International Symposium on*. IEEE. 2016, pp. 17–21.

[21] Aude Billard, Ben Robins, Jacqueline Nadel, and Kerstin Dautenhahn. "Building robota, a mini-humanoid robot for the rehabilitation of children with autism". In: *Assistive Technology* 19.1 (2007), pp. 37–49.

[22] Joshua Wainer, Ben Robins, Farshid Amirabdollahian, and Kerstin Dautenhahn. "Using the humanoid robot KASPAR to autonomously play triadic games and facilitate collaborative play among children with autism". In: *IEEE Transactions on Autonomous Mental Development* 6.3 (2014), pp. 183–199.

[23] Jeroen C. J. Brok and Emilia I. Barakova. "Engaging Autistic Children in Imitation and Turn-taking Games with Multiagent System of Interactive Lighting Blocks". In: *Proceedings of the 9th International Conference on Entertainment Computing*. Seoul, Korea, 2010, pp. 115–126.

[24] Matthew J Hertenstein, Rachel Holmes, Margaret McCullough, and Dacher Keltner. "The communication of emotion via touch." In: *Emotion* 9.4 (2009), p. 566.

[25] Betsy App, Daniel N McIntosh, Catherine L Reed, and Matthew J Hertenstein. "Nonverbal channel use in communication of emotion: how may depend on why." In: *Emotion* 11.3 (2011), p. 603.

[26] Lena M Forsell and Jan A Åström. "Meanings of hugging: From greeting behavior to touching implications". In: *Comprehensive Psychology* 1 (2012), pp. 02–17.

[27] Marc Hassenzahl, Stephanie Heidecker, Kai Eckoldt, Sarah Diefenbach, and Uwe Hillmann. "All you need is love: Current strategies of mediating intimate relationships through technology". In: *ACM Transactions on Computer-Human Interaction (TOCHI)* 19.4 (2012), p. 30.

[28] Kim Baraka, Stephanie Rosenthal, and Manuela Veloso. "Enhancing human understanding of a mobile robot's state and actions using expressive lights". In: *Robot and Human Interactive Communication (RO-MAN), 2016 25th IEEE International Symposium on*. IEEE. 2016, pp. 652–657.

[29] Tomoko Yonezawa, Naoto Yoshida, and Kento Kuboshima. "Design of Pet Robots with Limitations of Lives and Inherited Characteristics". In: *Proceedings of the 9th EAI International Conference on Bio-inspired Information and Communications Technologies (Formerly BIONETICS)*. BICT'15. New York City, United

States: ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2016, pp. 69–72. ISBN: 978-1-63190-100-3. DOI: 10.4108/eai.3-12-2015.2262417. URL: http://dx.doi.org/10.4108/eai.3-12-2015.2262417.

[30] Sichao Song and Seiji Yamada. "Expressing Emotions through Color, Sound, and Vibration with an Appearance-Constrained Social Robot." In: *HRI*. 2017, pp. 2–11.

[31] Michael Hemphill. "A note on adults' color–emotion associations". In: *The Journal of genetic psychology* 157.3 (1996), pp. 275–280.

[32] Kazunori Terada, Atsushi Yamauchi, and Akira Ito. "Artificial emotion expression for a robot by dynamic color change". In: *RO-MAN, 2012 IEEE*. IEEE. 2012, pp. 314–321.

[33] Thecla Schiphorst, Frank Nack, Michiel KauwATjoe, Simon De Bakker, Lora Aroyo, Angel Perez Rosillio, Hielke Schut, Norm Jaffe, et al. "PillowTalk: Can we afford intimacy?" In: *Proceedings of the 1st international conference on Tangible and embedded interaction*. ACM. 2007, pp. 23–30.

[34] Johan Redström, Maria Redström, and Ramia Mazé. *IT+ textiles*. Edita Publishing Oy, 2005.

[35] Joseph B Walther and Judee K Burgoon. "Relational communication in computer-mediated interaction". In: *Human communication research* 19.1 (1992), pp. 50–88.

[36] Ronald E Rice and Gail Love. "Electronic emotion: Socioemotional content in a computer-mediated communication network". In: *Communication research* 14.1 (1987), pp. 85–108.

[37] Daantje Derks, Agneta H Fischer, and Arjan ER Bos. "The role of emotion in computer-mediated communication: A review". In: *Computers in Human Behavior* 24.3 (2008), pp. 766–785.

[38] Jan BF Van Erp and Alexander Toet. "Social touch in human–computer interaction". In: *Frontiers in digital humanities* 2 (2015), p. 2.

[39] Jocelyn Smith and Karon MacLean. "Communicating emotion through a haptic link: Design space and methodology". In: *International Journal of Human-Computer Studies* 65.4 (2007), pp. 376–387.

[40] Mohamad A Eid and Hussein Al Osman. "Affective haptics: Current research and future directions". In: *IEEE Access* 4 (2016), pp. 26–40.

[41] Hidenobu Sumioka, Aya Nakae, Ryota Kanai, and Hiroshi Ishiguro. "Huggable communication medium decreases cortisol levels". In: *Scientific reports* 3 (2013).

[42] Trevor Van Gorp and Edie Adams. *Design for emotion*. Elsevier, 2012.

[43] Florian'Floyd' Mueller, Frank Vetere, Martin R Gibbs, Jesper Kjeldskov, Sonja Pedell, and Steve Howard. "Hug over a distance". In: *CHI'05 extended abstracts on Human factors in computing systems*. ACM. 2005, pp. 1673–1676.

[44] Jill Boucher. *interventions with children with autism-methods based on play*. 1999.

[45] Amanda Boutot, Tracee Guenther, and Shannon Crozier. "Let's play: Teaching play skills to young children with autism". In: *Education and Training in Developmental Disabilities* (2005), pp. 285–292.

[46] Simon Baron-Cohen. "Autism and symbolic play". In: *British Journal of Developmental Psychology* 5.2 (1987), pp. 139–148. ISSN: 2044-835X. DOI: 10.1111/j.2044-835X.1987.tb01049.x. URL: http://dx.doi.org/10.1111/j.2044-835X.1987.tb01049.x.

[47] Christopher Jarrold, Jill Boucher, and Peter Smith. "Symbolic play in autism: A review". In: *Journal of Autism and Developmental Disorders* 23.2 (1993), pp. 281–307. ISSN: 1573-3432. DOI: 10.1007/BF01046221. URL: https://doi.org/10.1007/BF01046221.

[48] Emma Williams, Vasudevi Reddy, and Alan Costall. "Taking a Closer Look at Functional Play in Children with Autism". In: *Journal of Autism and Developmental Disorders* 31.1 (2001), pp. 67–77. ISSN: 1573-3432. DOI: 10.1023/A:1005665714197. URL: https://doi.org/10.1023/A:1005665714197.

[49] Jill Boucher. "Editorial: interventions with children with autism - methods based on play". In: *Child Language Teaching and Therapy* 15.1 (1999), pp. 1–5. DOI: 10.1177/026565909901500101. eprint: https://doi.org/10.1177/026565909901500101. URL: https://doi.org/10.1177/026565909901500101.

[50] Sharon Behl Wulff. "The symbolic and object play of children with autism: A review". In: *Journal of Autism and Developmental Disorders* 15.2 (1985), pp. 139–148. ISSN: 1573-3432. DOI: 10.1007/BF01531600. URL: https://doi.org/10.1007/BF01531600.

[51] Brian Scassellati, Henny Admoni, and Maja MatariÄ. "Robots for Use in Autism Research". In: *Annual Review of Biomedical Engineering* 14.1 (2012). PMID: 22577778, pp. 275–294. DOI: 10.1146/annurev-bioeng-071811-150036. eprint: https://doi.org/10.1146/annurev-bioeng-071811-150036. URL: https://doi.org/10.1146/annurev-bioeng-071811-150036.

[52] Ben Robins, Nuno Otero, Ester Ferrari, and Kerstin Dautenhahn. “Eliciting requirements for a robotic toy for children with autism-results from user panels”. In: *RO-MAN 2007-The 16th IEEE International Symposium on Robot and Human Interactive Communication*. IEEE. 2007, pp. 101–106.

[53] Brooke Ingersoll, Laura Schreibman, and Quy H. Tran. “Effect of Sensory Feedback on Immediate Object Imitation in Children with Autism”. In: *Journal of Autism and Developmental Disorders* 33.6 (2003), pp. 673–683. ISSN: 1573-3432. DOI: 10.1023/B:JADD.0000006003.26667.f8. URL: https://doi.org/10.1023/B:JADD.0000006003.26667.f8.

[54] Jean. Piaget. *Play, Dreams, and Imitation in Childhood*. Norton library p. 579; p. 1962. Norton, 1962. ISBN: 9780393001716. URL: https://books.google.co.jp/books?id=C71APgAACAAJ.

[55] Christena Blauvelt Harper, Jennifer B. G. Symon, and William D. Frea. “Recess is Time-in: Using Peers to Improve Social Skills of Children with Autism”. In: *Journal of Autism and Developmental Disorders* 38.5 (2008), pp. 815 –826.

[56] Sarah R. Rieth, Aubyn C. Stahmer, Jessica Suhrheinrich, Laura Schreibman, Joanna Kenned, and Benjamin Ross. “Identifying Critical Elements of Treatment: Examining the Use of Turn Taking in Autism Intervention”. In: *Focus on Autism and Other Developmental Disabilities* 29.3 (2014), pp. 168 –179.

[57] Paul Yoder and Wendy L Stone. “Randomized comparison of two communication interventions for preschoolers with autism spectrum disorders.” In: *Journal of consulting and clinical psychology* 74.3 (2006), p. 426.

[58] Catherine Aldred, Jonathan Green, and Catherine Adams. “A new social communication intervention for children with autism: pilot randomised controlled treatment study suggesting effectiveness”. In: *Journal of Child Psychology and Psychiatry* 45.8 (2004), pp. 1420–1430.

[59] Peter Mundy and Mary Crowson. “Joint attention and early social communication: Implications for research on intervention with autism”. In: *Journal of Autism and Developmental disorders* 27.6 (1997), pp. 653–676.

[60] Kerstin Dautenhahn. “Socially intelligent robots: dimensions of human–robot interaction”. In: *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 362.1480 (2007), pp. 679–704.

[61] Pegah Soleiman, Hadi Moradi, Maryam Mahmoudi, Mohyeddin Teymouri, and Hamid Reza Pouretemad. “The Use of RoboParrot in the Therapy of Children with Autism Children: In Case of Teaching the Turn-Taking Skills”. In: (2011).

[62] Emilia Barakova, Jan Gillessen, and Loe Feijs. “Social Training of Autistic Children with Interactive Intelligent Agents”. In: *Journal of Integrative Neuroscience* 8.1 (2009), pp. 23–34.

[63] Brian Scassellati. “How social robots will help us to diagnose, treat, and understand autism”. In: *Robotics research*. Springer, 2007, pp. 552–563.

[64] James M. Rehg. “Behavior Imaging: Using Computer Vision to Study Autism”. In: *MVA’11*. 2011, pp. 14–21.

[65] Salvatore M Anzalone, Sofiane Boucenna, Serena Ivaldi, and Mohamed Chetouani. “Evaluating the engagement with social robots”. In: *International Journal of Social Robotics* 7.4 (2015), pp. 465–478.

[66] Fahd Albinali, Matthew S Goodwin, and Stephen S Intille. “Recognizing stereotypical motor movements in the laboratory and classroom: a case study with children on the autism spectrum”. In: *Proceedings of the 11th international conference on Ubiquitous computing*. ACM. 2009, pp. 71–80.

[67] Atsushi Funahashi, Anna Gruebler, Takeshi Aoki, Hideki Kadone, and Kenji Suzuki. “Brief Report: The Smiles of a Child with Autism Spectrum Disorder During an Animal-assisted Activity May Facilitate Social Positive Behaviors?Quantitative Analysis with Smile-detecting Interface”. In: *Journal of Autism and Developmental Disorders* 44.3 (2014), pp. 685–693. DOI: 10.1007/s10803-013-1898-4.

[68] Yadong Pan, Masakazu Hirokawa, and Kenji Suzuki. “Measuring K-degree facial interaction between robot and children with autism spectrum disorders”. In: *24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2015)*. 2015, pp. 48–53.

[69] Ognjen Rudovic, Jaeryoung Lee, Miles Dai, Bjorn Schuller, and Rosalind Picard. “Personalized machine learning for robot perception of affect and engagement in autism therapy”. In: *arXiv preprint arXiv:1802.01186* (2018).

[70] Fabrizio Taffoni, Valentina Focaroli, Domenico Formica, Eugenio Gugliemelli, Flavio Keller, and Jana M Iverson. “Sensor-based technology in the study of motor skills in infants at risk for ASD”. In: *2012 4th IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob)*. IEEE. 2012, pp. 1879–1883.

[71] Diego Rivera, Antonio García, Bernardo Alarcos, Juan R Velasco, José Eugenio Ortega, and Isaías Martínez-Yelmo. “Smart toys designed for detecting developmental delays”. In: *Sensors* 16.11 (2016), p. 1953.

[72] Mario Vega-Barbas, Iván Pau, Javier Ferreira, Evelyn Lebis, and Fernando Seoane. “Utilizing smart textiles-enabled sensorized toy and playful interactions for assessment of psychomotor development on children”. In: *Journal of Sensors* 2015 (2015).

[73] Ami Klin. “Attributing social meaning to ambiguous visual stimuli in higher-functioning autism and Asperger syndrome: The social attribution task”. In: *The Journal of Child Psychology and Psychiatry and Allied Disciplines* 41.7 (2000), pp. 831–846.

[74] Francesca GE Happé. “Communicative competence and theory of mind in autism: A test of relevance theory”. In: *Cognition* 48.2 (1993), pp. 101–119.

[75] Eleuda Nunez, Masakazu Hirokawa, Hiroya Igarashi, and Kenji Suzuki. “Robotic Device to Mediate Human-Human Hug-Driven Remote Communication”. In: *International Conference on Social Robotics*. Springer. 2017, pp. 198–208.

[76] Kakehi Gota, Yuta Sugiura, Anusha Indrajith Withana, Calista Lee, Naohisa Nagaya, Daisuke Sakamoto, Maki Sugimoto, Masahiko Inami, and Takeo Igarashi. “FuwaFuwa: detecting shape deformation of soft objects using directional photoreflectivity measurement”. In: *SIGGRAPH Emerging Technologies*. 2011.

[77] Eleuda Nunez, Masakazu Hirokawa, and Kenji Suzuki. “Design of a Huggable Social Robot with Affective Expressions Using Projected Images”. In: *Applied Sciences* 8.11 (2018).

[78] Ronit Slyper, Ivan Poupyrev, and Jessica Hodgins. “Sensing through structure: designing soft silicone sensors”. In: *Proceedings of the fifth international conference on Tangible, embedded, and embodied interaction*. ACM. 2011, pp. 213–220.

[79] Eleuda Nunez, Soichiro Matsuda, Masakazu Hirokawa, Junichi Yamamoto, and Kenji Suzuki. “Effect of Sensory Feedback on Turn-Taking Using Paired Devices for Children with ASD”. In: *Multimodal Technologies and Interaction* 2.4 (2018).

[80] Jussi Suomela and Tomi Ylikorpi. “Ball-shaped robots: an historical overview and recent developments at TKK”. In: *Field and service robotics*. Springer. 2006, pp. 343–354.

[81] Tamie Salter, Neil Davey, and François Michaud. “Designing & developing QueBall, a robotic device for autism therapy”. In: *The 23rd IEEE International Symposium on Robot and Human Interactive Communication*. IEEE. 2014, pp. 574–579.

[82] Joris H Janssen, Wijnand A Ijsselsteijn, and Joyce HDM Westerink. “How affective technologies can influence intimate interactions and improve social connectedness”. In: *International Journal of Human-Computer Studies* 72.1 (2014), pp. 33–43.

[83] Arthur Aron, Elaine N Aron, Michael Tudor, and Greg Nelson. "Close relationships as including other in the self." In: *Journal of personality and social psychology* 60.2 (1991), p. 241.

[84] Ellen Berscheid, Mark Snyder, and Allen M Omoto. "Measuring closeness: The relationship closeness inventory (RCI) revisited". In: *Handbook of closeness and intimacy* (2004), pp. 81–101.

[85] Lisanne M Jenkins and David G Andrewes. "A new set of standardised verbal and non-verbal contemporary film stimuli for the elicitation of emotions". In: *Brain Impairment* 13.2 (2012), pp. 212–227.

[86] Bartolini Ellen Elizabeth. "Eliciting Emotion with Film: Development of a Stimulus Set". In: (2011).

[87] Juliana Doria Bednarski. "Eliciting seven discrete positive emotions using film stimuli". PhD thesis. Vanderbilt University, 2012.

[88] James J Gross and Robert W Levenson. "Emotion elicitation using films". In: *Cognition & emotion* 9.1 (1995), pp. 87–108.

[89] Richard A McFarland. "Relationship of skin temperature changes to the emotions accompanying music". In: *Biofeedback and Self-regulation* 10.3 (1985), pp. 255–267.

[90] J Koch. *Biofeedback-A manual for the therapeutic practice*. 2003.

[91] Fred Shaffer and JP Ginsberg. "An overview of heart rate variability metrics and norms". In: *Frontiers in public health* 5 (2017), p. 258.

[92] Willem J Kop, Stephen J Synowski, Miranda E Newell, Louis A Schmidt, Shari R Waldstein, and Nathan A Fox. "Autonomic nervous system reactivity to positive and negative mood induction: The role of acute psychological responses and frontal electrocortical activity". In: *Biological psychology* 86.3 (2011), pp. 230–238.

[93] Mathias Benedek and Christian Kaernbach. "A continuous measure of phasic electrodermal activity". In: *Journal of neuroscience methods* 190.1 (2010), pp. 80–91.

[94] Jason J Braithwaite, Derrick G Watson, Robert Jones, and Mickey Rowe. "A guide for analysing electrodermal activity (EDA) & skin conductance responses (SCRs) for psychological experiments". In: *Psychophysiology* 49.1 (2013), pp. 1017–1034.

[95] Manja Lohse, Frank Hegel, and Britta Wrede. "Domestic applications for social robots: An online survey on the influence of appearance and capabilities". In: (2008).

[96] Frank Hegel, Manja Lohse, Agnes Swadzba, Sven Wachsmuth, Katharina Rohlfing, and Britta Wrede. "Classes of applications for social robots: a user study". In: *Robot and Human interactive Communication, 2007. RO-MAN 2007. The 16th IEEE International Symposium on*. IEEE. 2007, pp. 938–943.

[97] Jelle Saldien, Kristof Goris, Selma Yilmazyildiz, Werner Verhelst, and Dirk Lefeber. "On the design of the huggable robot Probo". In: (2008).

[98] Sooyeon Jeong, Kristopher Dos Santos, Suzanne Graca, Brianna O'Connell, Laurel Anderson, Nicole Stenquist, Katie Fitzpatrick, Honey Goodenough, Deirdre Logan, Peter Weinstock, et al. "Designing a socially assistive robot for pediatric care". In: *Proceedings of the 14th international conference on interaction design and children*. ACM. 2015, pp. 387–390.

[99] David Cameron, Samuel Fernando, Emily Collins, Abigail Millings, Roger Moore, Amanda Sharkey, Vanessa Evers, and Tony Prescott. "Presence of life-like robot expressions influences children's enjoyment of human-robot interactions in the field". In: *Proceedings of the AISB Convention 2015*. The Society for the Study of Artificial Intelligence and Simulation of Behaviour. 2015.

[100] Jelle Saldien, Kristof Goris, Bram Vanderborght, Johan Vanderfaeillie, and Dirk Lefeber. "Expressing emotions with the social robot Probo". In: *International Journal of Social Robotics* 2.4 (2010), pp. 377–389.

[101] R Hasselvander and F De Visme. `http://www.bluefrogrobotics.com/en/home/`. 2015.

[102] Eleuda Nuñez, Kyohei Uchida, and Kenji Suzuki. "PEPITA: A Design of Robot Pet Interface for Promoting Interaction". In: *International Conference on Social Robotics*. Springer. 2013, pp. 552–561.

[103] Christoph Bartneck. "Who like androids more: Japanese or US Americans?" In: *Robot and Human Interactive Communication, 2008. RO-MAN 2008. The 17th IEEE International Symposium on*. IEEE. 2008, pp. 553–557.

[104] Xiao-Ping Gao, John H Xin, Tetsuya Sato, Aran Hansuebsai, Marcello Scalzo, Kanji Kajiwara, Shing-Sheng Guan, J Valldeperas, Manuel Jose Lis, and Monica Billger. "Analysis of cross-cultural color emotion". In: *Color Research & Application* 32.3 (2007), pp. 223–229.

[105] Marina V Sokolova and Antonio Fernández-Caballero. "A review on the role of color and light in affective computing". In: *Applied Sciences* 5.3 (2015), pp. 275–293.

[106] Johannes Feldmaier, Tamara Marmat, Johannes Kuhn, and Klaus Diepold. "Evaluation of a RGB-LED-based Emotion Display for Affective Agents". In: *arXiv preprint arXiv:1612.07303* (2016).

[107] Jiaming Zhang and Amanda Sharkey. "Contextual recognition of robot emotions". In: *Towards autonomous robotic systems* (2011), pp. 78–89.

[108] Elena Corina Grigore, Andre Pereira, Jie Jessica Yang, Ian Zhou, David Wang, and Brian Scassellati. "Comparing Ways to Trigger Migration Between a Robot and a Virtually Embodied Character". In: *International Conference on Social Robotics*. Springer. 2016, pp. 839–849.

[109] Paulo Fontaínha Gomes, Elena Márquez Segura, Henriette Cramer, Tiago Paiva, Ana Paiva, and Lars Erik Holmquist. "ViPleo and PhyPleo: Artificial pet with two embodiments". In: *Proceedings of the 8th International Conference on Advances in Computer Entertainment Technology*. ACM. 2011.

[110] Elena Márquez Segura, Henriette Cramer, Paulo Fontaínha Gomes, Stina Nylander, and Ana Paiva. "Revive!: reactions to migration between different embodiments when playing with robotic pets". In: *Proceedings of the 11th international conference on interaction design and children*. ACM. 2012, pp. 88–97.

[111] Eleuda Nunez, Soichiro Matsuda, Masakazu Hirokawa, Junichi Yamamoto, and Kenji Suzuki. "Paired robotic devices to mediate and represent social behaviors". In: *24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2015)*. 2015, pp. 722–727.

[112] Eleuda Nuñez, Soichiro Matsuda, Masakazu Hirokawa, Junichi Yamamoto, and Kenji Suzuki. "An approach to facilitate turn-taking behavior with paired devices for children with Autism Spectrum Disorder". In: *25th IEEE International Symposium on Robot and Human Interactive Communication, RO-MAN 2016, New York, NY, USA, August 26-31, 2016*. 2016, pp. 837–842. DOI: 10.1109/ROMAN.2016.7745216. URL: https://doi.org/10.1109/ROMAN.2016.7745216.

[113] Soichiro Matsuda, Eleuda Nunez, Masakazu Hirokawa, Junichi Yamamoto, and Kenji Suzuki. "Facilitating social play for children with PDDs: effects of paired robotic devices". In: *Frontiers in psychology* 8 (2017), p. 1029.

[114] Karen Pierce and Laura Schreibman. "Increasing complex social behaviors in children with autism: Effects of peer-implemented pivotal response training". In: *Journal of applied behavior analysis* 28.3 (1995), pp. 285–295.

[115] Karen Pierce and Laura Schreibman. "Multiple peer use of pivotal response training to increase social behaviors of classmates with autism: Results from trained and untrained peers". In: *Journal of applied Behavior analysis* 30.1 (1997), pp. 157–160.

[116] Kwang-Ho Choi, Junbeom Kim, O Sang Kwon, Min Ji Kim, Yeon Hee Ryu, and Ji-Eun Park. "Is heart rate variability (HRV) an adequate tool for evaluating human emotions?–A focus on the use of the International Affective Picture System (IAPS)". In: *Psychiatry research* 251 (2017), pp. 192–196.

Appendix A

Design of Macaron's Circuit

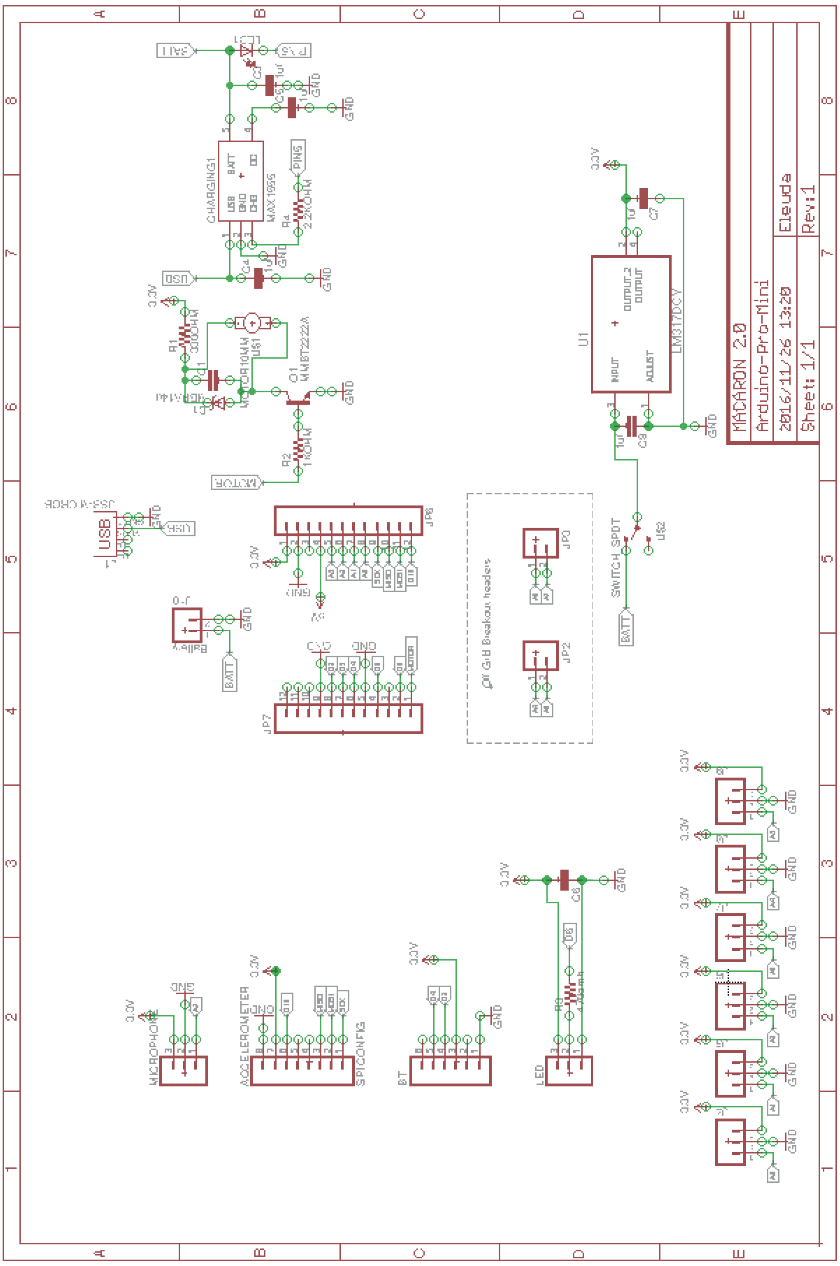


FIGURE A.1: Macaron: schematics of the main board.

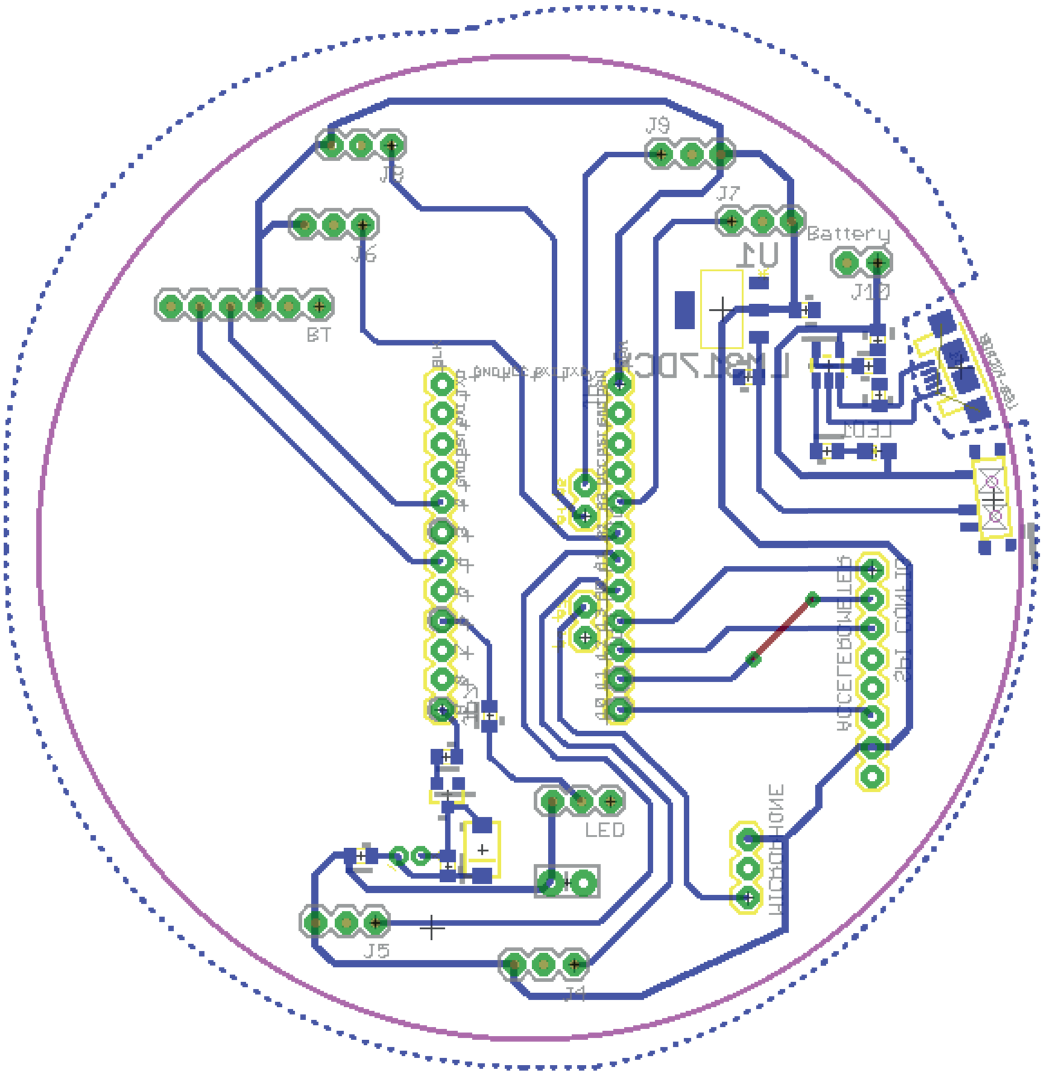
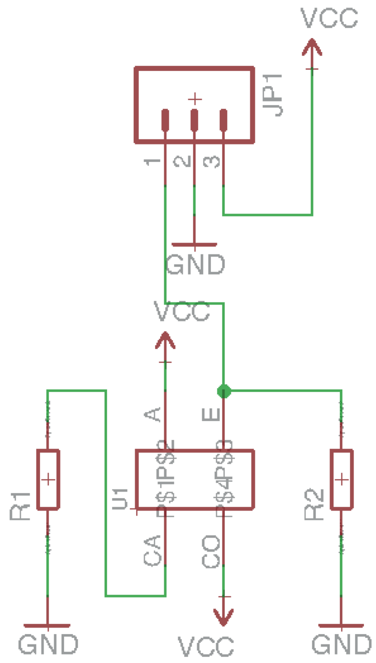
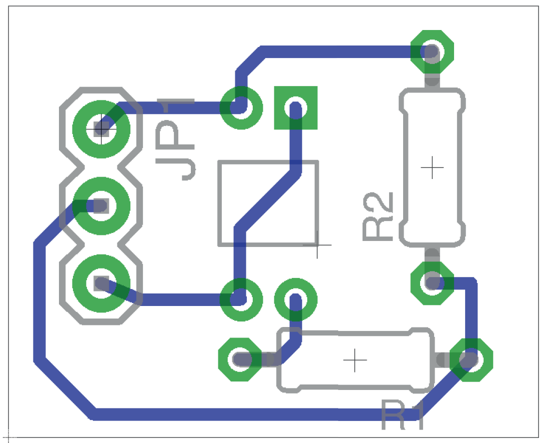


FIGURE A.2: Macaron's main board layout: the components were distributed on a circular PCB board that is placed inside a plastic case.



(A)



(B)

FIGURE A.3: Schematics and layout for Macaron's sensors board: these contain the circuit for the photoreflective sensors, which is wired to the sockets on the main board.

Appendix B

Design of Macaron's Plastic Case

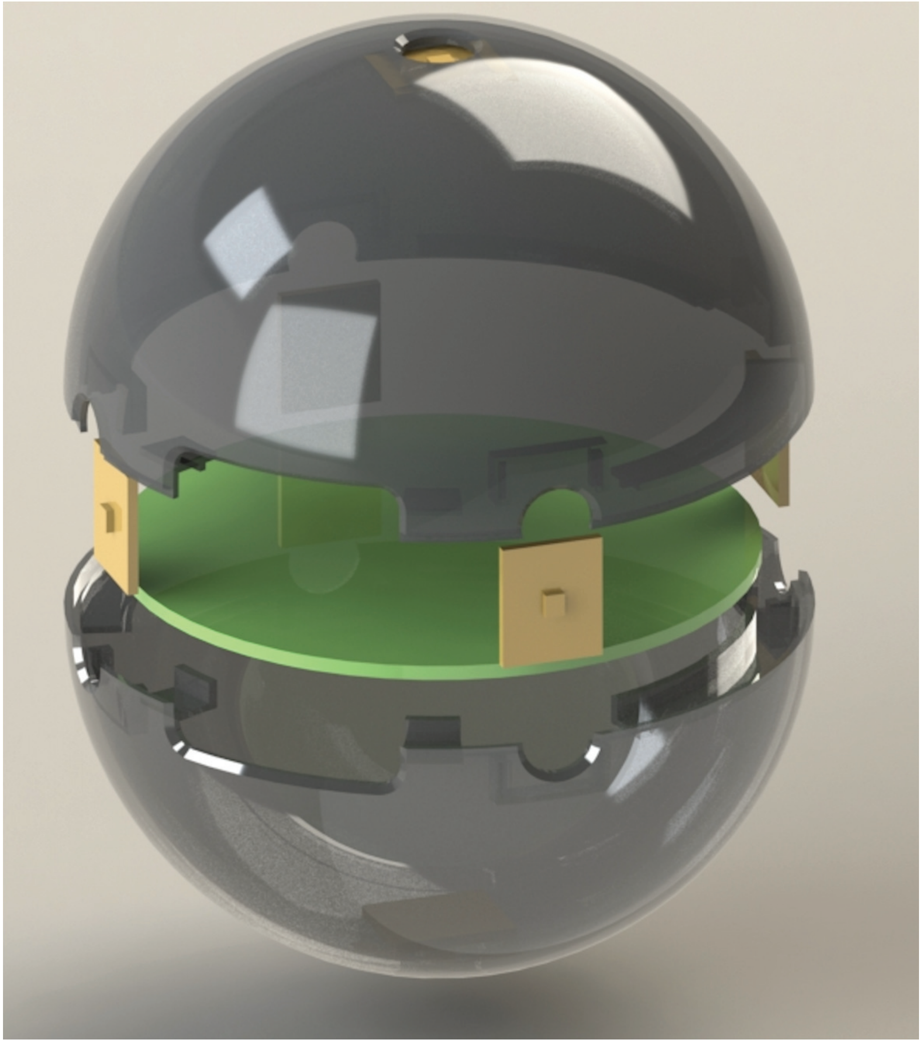


FIGURE B.1: Schematic of Macaron's case: it was designed with a piece to support the circuit board, as well as the six IR sensors.

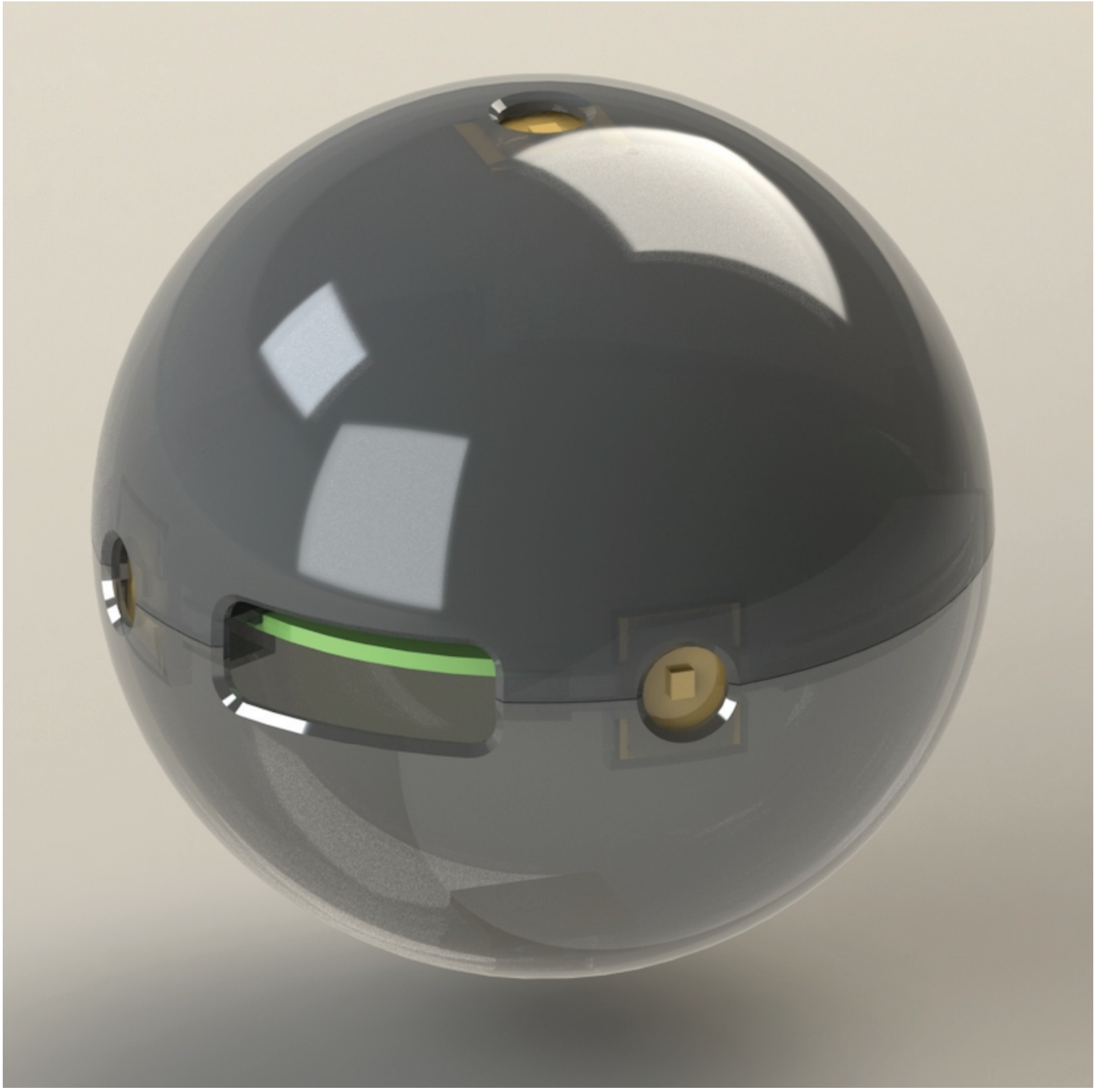
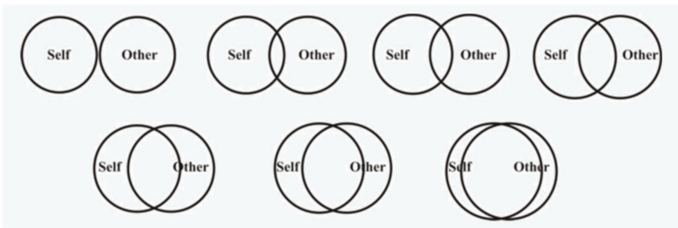


FIGURE B.2: Schematic of Macaron's case: when it is closed, the biggest hole is used to access to the switch and the charging port.

Appendix C

Questionnaires

Which diagram represents how close you felt to your partner while watching the movies? (circle your answer)



What do you think of the method of interaction?

The interaction with my partner was….

1) Unsociable 1 2 3 4 5 6 7 Sociable

2) Very cold 1 2 3 4 5 6 7 Very warm

3) Impersonal 1 2 3 4 5 6 7 Personal

While watching the movies, I felt aware of my partner

Strongly disagree 1 2 3 4 5 6 7 Strongly agree

FIGURE C.1: Form Q1: the first item is IOS ([83]), the next three evaluate aspects related to the quality of the interaction. These questions were used to evaluate an affective technology for communication [82].

1. I think that I would like to use this system frequently

Not at all 1 2 3 4 5 6 7 Very much

2. I would imagine that most people would learn to use this system very quickly

Not at all 1 2 3 4 5 6 7 Very much

3. I think this system would help me to know how my friends and family are feeling

Not at all 1 2 3 4 5 6 7 Very much

4. I think this system would help me to feel more connected with people that I am geographically separated from

Not at all 1 2 3 4 5 6 7 Very much

5. I think this system is fun to use

Not at all 1 2 3 4 5 6 7 Very much

6. I think this system is useful

Not at all 1 2 3 4 5 6 7 Very much

7. I think this system would be easy to operate

Not at all 1 2 3 4 5 6 7 Very much

8. I think this system would help me to communicate my emotions to friends and family

Not at all 1 2 3 4 5 6 7 Very much

9. I think this system would give me a sense of ambient presence from people that are geographically distanced

Not at all 1 2 3 4 5 6 7 Very much

10. I would recommend this system to a friend

Not at all 1 2 3 4 5 6 7 Very much

11. I could see myself use this system in my everyday life

Not at all 1 2 3 4 5 6 7 Very much

12. I think it is easy to understand this system

Not at all 1 2 3 4 5 6 7 Very much

13. I think using this system would affect my mood

Not at all 1 2 3 4 5 6 7 Very much

14. I think this system would contribute to a richer social experience

Not at all 1 2 3 4 5 6 7 Very much

15. I think this system is pleasant to use

Not at all 1 2 3 4 5 6 7 Very much

FIGURE C.2: Form Q1: the first item is IOS ([83]), the next three evaluate aspects related to the quality of the interaction. These questions were used to evaluate an affective technology for communication [82].

Appendix D

Movies for elicitation of emotions

Happiness		Marie Antoniette (2006)	Jenkins and Andrewes (2012)
Fear		Road kill (2001)	Jenkins and Andrewes (2012)
Sadness		Sylvia (2003)	Jenkins and Andrewes (2012)
Amusement		Mr. goes to town episode (1990)	Jenkins and Andrewes (2012)

FIGURE D.1: Films selected from libraries. From left to right: 1) Targeted emotion, 2) Screenshot of the material, 3) Name of the film from which the material was taken from, 4) Library that contains details of the films extract.

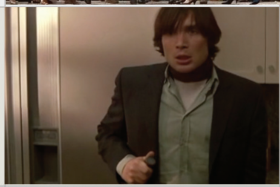
Happiness		500 days of summer (2009)	Bednarski (2012)
Fear		Red eye (2005)	Jenkins and Andrewes (2012)
Sadness		The Shawshank Redemption (1994)	Bartolini (2011)
Amusement		The Hangover (2009)	Bartolini (2011)

FIGURE D.2: Films selected from libraries. From left to right: 1) Targeted emotion, 2) Screenshot of the material, 3) Name of the film from which the material was taken from, 4) Library that contains details of the films extract.

Happiness		Deep Blue (2003)	Jenkins and Andrewes (2012)
Fear		Silence of the lambs (1991)	Gross and Levenson (1995)
Sadness		My girl (1991)	Jenkins and Andrewes (2012)
Amusement		When Harry met Sally (1989)	Gross and Levenson (1995)

FIGURE D.3: Films selected from libraries. From left to right: 1) Targeted emotion, 2) Screenshot of the material, 3) Name of the film from which the material was taken from, 4) Library that contains details of the films extract.

Appendix E

Targeted segments of data

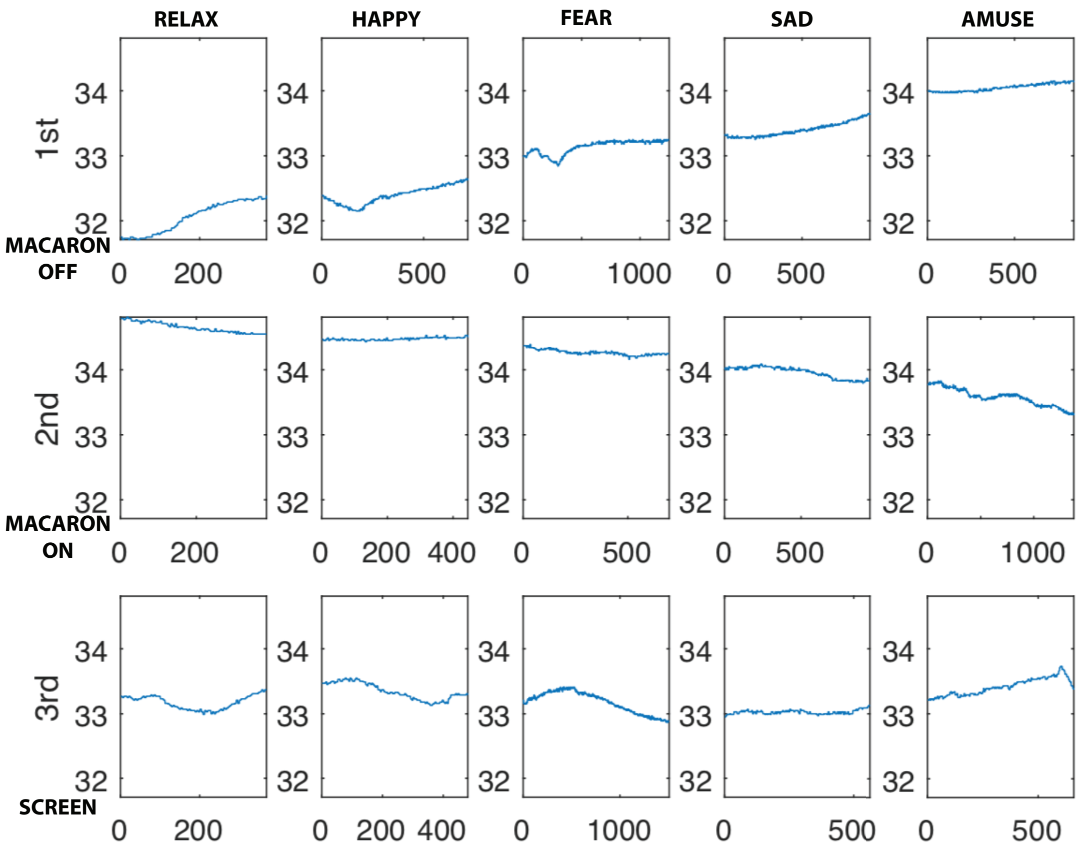


FIGURE E.1: Skin temperature data from the targeted segments of the experiment.

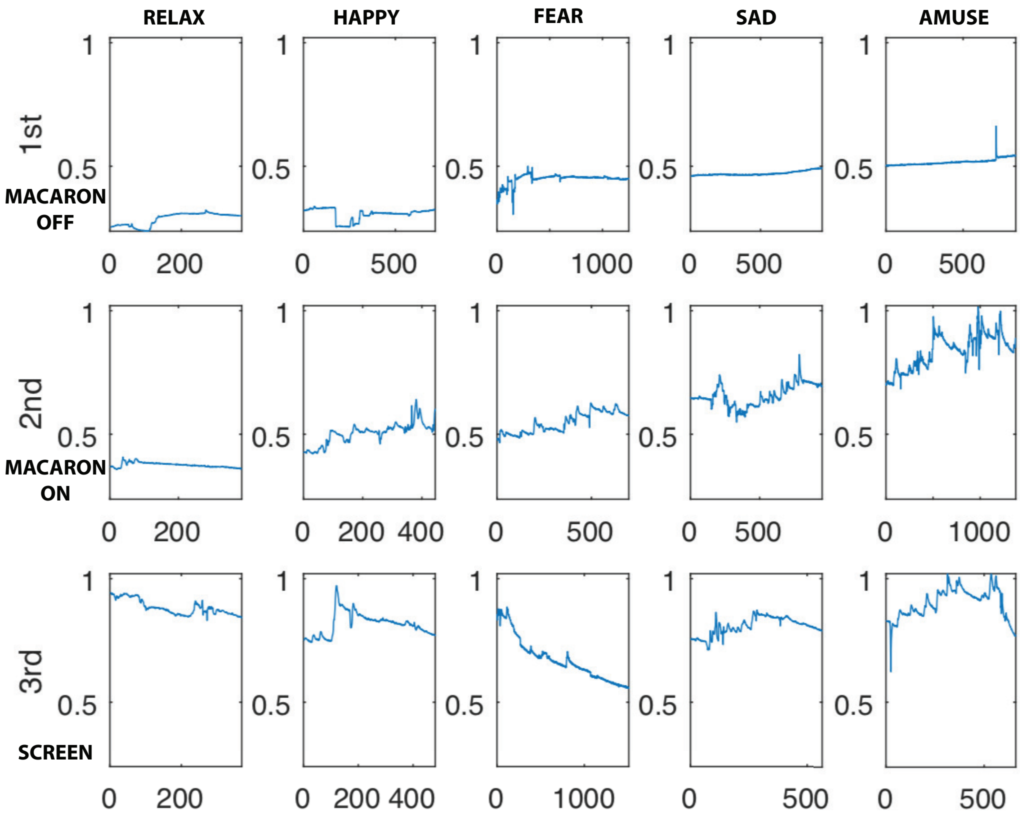


FIGURE E.2: Electrodermal activity data from the targeted segments of the experiment.

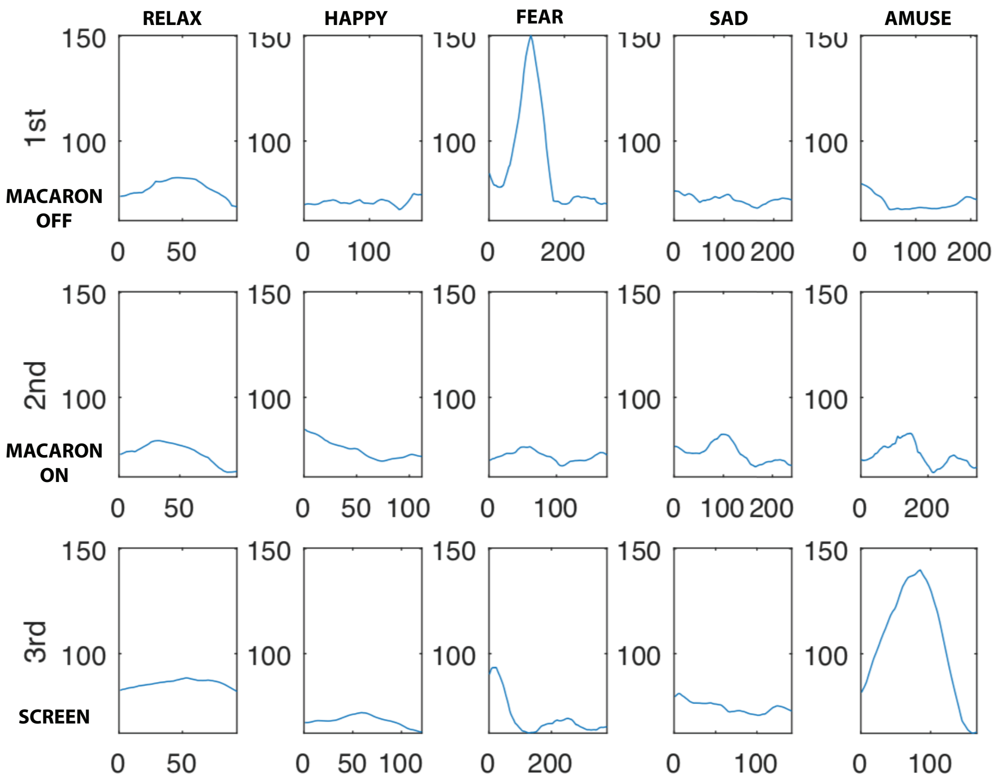


FIGURE E.3: Heart rate data from the targeted segments of the experiment.

Appendix F

Exploring pepita

F.1 Questionnaire for Exploring the Design of Pepita

- Page 1: The questionnaire consists of different photos and videos, and you will be asked to give your impressions. Before starting please be aware of: (1) In case you do not understand an English word, please refer to a dictionary to be sure of the meaning before answering, (2) Look at the picture and answer based on your first impressions, (3) Read all the sentences and instructions, (4) The entire questionnaire will take approx. 15min.
- Page 2: Consent to participate in the research.
- Page 3: Nationality, age, gender.
- Page 4: Part 1: This section will introduce different huggable robots. In other words, a robot that can sense hugging actions. These robots differ in size, appearance, and shape. You will be asked to rate them using a scale.
- Page 5: Please rate the following statements for each presented picture. (Photo of 1 of the 4 huggable robots, counterbalanced order) The robot looks huggable (5 points scale, from strongly disagree to “strongly agree”). The robot looks easy to hug (5 points scale, from “strongly disagree” to “strongly agree”). The robot looks appealing to hug (5 points scale, from “strongly disagree” to “strongly agree”).
- Page 6: (Photo of 1 of the 4 huggable robots, counterbalanced order) The robot looks huggable (5 points scale, from strongly disagree to “strongly agree”). The robot looks easy to hug (5 points scale, from strongly disagree to “strongly agree”). The robot looks appealing to hug (5 points scale, from strongly disagree to “strongly agree”).
- Page 7: (Photo of 1 of the 4 huggable robots, counterbalanced order) The robot looks huggable (5 points scale, from “strongly disagree” to “strongly agree”). The robot looks easy to hug (5

points scale, from "strongly disagree" to "strongly agree"). The robot looks appealing to hug (5 points scale, from "strongly disagree" to "strongly agree").

- Page 8: (Photo of 1 of the 4 huggable robots, counterbalanced order) The robot looks huggable (5 points scale, from "strongly disagree" to "strongly agree"). The robot looks easy to hug (5 points scale, from "strongly disagree" to "strongly agree"). The robot looks appealing to hug (5 points scale, from "strongly disagree" to "strongly agree").
- Page 9: When you rated the different huggable robots, how relevant were these features for your answers? (5 points scale, from unimportant to extremely important) (1) The shape of the robot body, (2) Size of the robot body, (3) Weight of the robot body, (4) Texture of the skin, (5) Softness of the robot body, (6) Appearance of the robot.
- Page 10: Part 2: When interacting with people, robots need to understand and convey a representation of emotions. In this section, you will watch four different videos in succession displaying different robot's expressions. Then, using some photos as reference, you will be asked to give your general impressions about them.
- Page 11: (video of 1 of the 4 robots displaying facial expression by a display or with a mechanical face, counterbalanced order).
- Page 12: (Photo showing the previous robot's expressions) Based on your first impression, please express using the following scale how acceptable for you is the robot's expressions of emotions? (5 points scale using emoticons from sad to happy).
- Page 13: (video of 1 of the 4 robots displaying facial expression by a display or with a mechanical face, counterbalanced order).
- Page 14: (Photo showing the previous robot's expressions) Based on your first impression, express using the following scale how acceptable for you is the robot's expressions of emotions? (5 points scale using emoticons from sad to happy).
- Page 15: (video of 1 of the 4 robots displaying facial expression by a display or with a mechanical face, counterbalanced order).
- Page 16: (Photo showing the previous robot's expressions) Based on your first impression, express using the following scale how acceptable for you is the robot's expressions of emotions? (5 points scale using emoticons from sad to happy).
- Page 17: (video of 1 of the 4 robots displaying facial expression by a display or with a mechanical face, counterbalanced order).

- Page 18: (Photo showing the previous robot's expressions) Based on your first impression, express using the following scale how acceptable for you is the robot's expressions of emotions? (5 points scale using emoticons from sad to happy).
- Page 19: Part 3: In this section, you will be asked to give your general impression about the social robot companion Pepita. This robotic device was designed to be placed at home and interact with people in everyday life. (Video of a person interacting with Pepita)
- Page 20: (Photo of Pepita) Please express your impressions of Pepita using the following scale: (7 points scale with 8 items, from Awful to Nice, from Machinelike to Humanlike, from Artificial to Lifelike, from Unpleasant to Pleasant, from Fake to Natural, from Unfriendly to Friendly, from Unconscious to Conscious, from Unkind to Kind). This question was followed by two blank spaces to collect the features of Pepita that positively and negatively impacted the answers.

F.2 Questionnaire for Exploring the Affective Feedback Using Projected Avatars

- Page 1: The questionnaire consists of two sets of two videos followed by some questions: (1) The videos display Pepita, a robotic device displaying different visual feedback; (2) Then, you will be asked about your perception and impressions, (3) The entire questionnaire will take approx. 10 min.
- Page 2: Consent to participate in the research.
- Page 3: Nationality, age, gender.
- Page 4: Task overview.
- Page 5: Case 1: In the following video, the robot is displaying light color patterns. (You can play this video multiple times) (Player showing Projector or LED condition, happy or sad. All the options are counterbalanced)
- Page 6: Case 2: In the following video, the robot is displaying light color patterns. (You can play this video multiple times) (Player showing Projector or LED condition, happy or sad. All the options are counterbalanced)
- Page 7: Please select the option that reflects your immediate response to each statement. Do not think too long about each statement. Make sure you answer every question. (Photo of

case 1) As a total impression, I consider that the perceived behavior of the robot makes reference to the following statements: Happy-like behavior (5 points scale, from “strongly disagree” to "strongly agree"). (Photo of case 2) As a total impression, I consider that the perceived behavior of the robot makes reference to the following statements: Happy-like behavior (5 points scale, from strongly disagree to strongly agree).

- Page 8: Case 3: In the following video, the robot is displaying light color patterns. (You can play this video multiple times) (Player showing Projector or LED condition, happy or sad. All the options are counterbalanced)
- Page 9: Case 4: In the following video, the robot is displaying light color patterns. (You can play this video multiple times) (Player showing Projector or LED condition, happy or sad. All the options are counterbalanced)
- Page 11: Please select the option that reflects your immediate response to each statement. Do not think too long about each statement. Make sure you answer every question. (Photo of case 3) As a total impression, I consider that the perceived behavior of the robot makes reference to the following statements: Happy-like behavior (5 points scale, from “strongly disagree” to “strongly agree”). (Photo of case 4) As a total impression, I consider that the perceived behavior of the robot makes reference to the following statements: Happy-like behavior (5 points scale, from strongly disagree to strongly agree).
- Page 12: (Photo of Pepita projecting avatars) From the following statements, choose the one that most closely reflects your perception about the robot body interface: (1) I perceive the robot body interface as two entities: an avatar and a robot. (2) I perceive the robot body interface as one entity: the robot and its avatar.