EINDHOVEN UNIVERSITY OF TECHNOLOGY

Perceptual crossing with artificial eyes

Citation for published version (APA):
Anas, S. A. B. (2021). *Perceptual crossing with artificial eyes: Designing bidirectional and proactive human-object interaction based on the perceptual crossing paradigm*. [Phd Thesis 1 (Research TU/e / Graduation TU/e), Industrial Design]. Technische Universiteit Eindhoven.

Document status and date:
Published: 15/04/2021

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

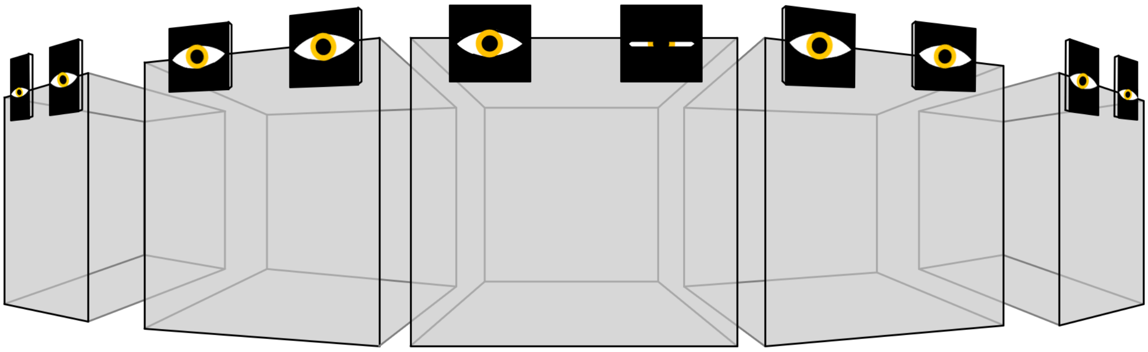
Take down policy
If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Download date: 16. Nov. 2023

PERCEPTUAL CROSSING
WITH
ARTIFICIAL EYES



SITI AISYAH BINTI ANAS

PERCEPTUAL CROSSING
WITH
ARTIFICIAL EYES

Designing bidirectional and proactive human-object interaction based on the perceptual crossing paradigm

SITI AISYAH BINTI ANAS

سيتي عايشه بنت اناس

A catalogue record is available from the Eindhoven University of Technology Library

ISBN: 978-90-386-5246-7

Cover design: Siti Aisyah binti Anas

Printed by Gildeprint.nl, Enschede

© Siti Aisyah binti Anas, 2021 All Rights Reserved.

Perceptual Crossing With Artificial Eyes

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de
Technische Universiteit Eindhoven, op gezag van de
rector magnificus prof.dr.ir. F.P.T. Baaijens, voor een
commissie aangewezen door het College voor
Promoties, in het openbaar te verdedigen
op donderdag 15 april 2021 om 16:00 uur

door

Siti Aisyah binti Anas

geboren te Selangor, Maleisië

Dit proefschrift is goedgekeurd door de promotoren en de samenstelling van de promotiecommissie is als volgt:

voorzitter:	prof.dr. L. Chen
1e promotor:	prof.dr. G.W.M. Rauterberg
2e promotor:	dr. J. Hu PDEng MEng
copromotor(en):	dr. R.-H. Liang
leden:	prof.dr.ir. L.M.G. Feijs
	prof.dr. P. Marti
	dr. M. Li
	dr. R. Vertegaal (Human Media Lab)

Het onderzoek of ontwerp dat in dit proefschrift wordt beschreven is uitgevoerd in overeenstemming met de TU/e Gedragscode Wetenschapsbeoefening.

The work in this dissertation was sponsored by Universiti Teknikal Malaysia Melaka (UTeM) and Ministry of Higher Education (MoHE), Malaysia under the auspices of the Public Institutions of Higher Education Academic Training Scheme (SLAI) Program.

Summary

Perceptual Crossing with Artificial Eyes

Designing bidirectional and proactive human-object interaction based on the perceptual crossing paradigm

Nowadays, objects are embedded with various sensors, making the objects knowledgeable and smart, gradually reducing users' need to intervene. As a result, these intelligent objects that work quietly in the background are perceived as passive and reactive objects when interacting with users. The lack of communication between the objects and the users impedes the objects from being smarter and understanding the users' requirements. Recently, many researchers actively and continuously research to improve the users' interaction and engagement with the objects. The research objectives are to increase and improve the users' awareness when interacting and engaging with the objects. Hence, it is essential to design useful feedback or feed-forward methods to indicate the objects' internal operation state and facilitating communication with engaging users. Another different method, direct manipulation of the objects that exploits the human skills, also enhanced the interaction and engagement between the objects and the users. Looking at that, yet most human-object communication adopts one-directional communication, where the human always acts as the initiator when interacting and engaging with the objects. Thus, it is questionable and remains arguable to understand the objects' smartness that could initiate and continuously communicate with the users.

The perceptual crossing paradigm is a paradigm that can create smartness in an object that could initiate and continuously communicate with the users when it is interacting and engaging with the users. The perceptual crossing paradigm also provides the necessary conditions that allow a person to recognize the other interacting entity as an intentional or reactive entity. Furthermore, the perceptual crossing paradigms emphasize the bidirectional and proactive communication between the person and the interacting entity. This paradigm also allows the interacting person to experience the differences during an interaction with an entity capable of initiating the communication or reacting towards the persons' presence. Looking at the addressed limitations in interaction design and the perceptual crossing paradigm's advantage, designers had adopted the perceptual crossing paradigm, which improves the richness and empathy of the human-object interaction. However, the conducted work does not emphasize the bidirectional and proactive interaction between the object and the interacting person. Hence, this thesis aims to design an object capable of expressing its intention to interact in a bidirectional interaction to improve the human-object interaction. Therefore, this work implemented the perceptual crossing paradigm to investigate the object's bidirectional and proactive behavior towards the interacting and engaging person.

Three studies were conducted to investigate the bidirectional and proactive interaction design based on the perceptual crossing paradigm. Visual attention is chosen as an interaction method to allow the object to differentiate an intentional user from a potential user. For instance, a user looking at the object can be interpreted as an intentional user that wants to interact. In contrast,

a user presence nearby the object can be interpreted as a potential user. In the first study, a proactive object capable of expressing its intention to communicate bi-directionally with a user is designed and developed using a simple abstract motion. An eye tracker is used to detect the user's visual attention and is integrated with the proactive object, allowing the object to initiate an interaction when a user's visual attention is detected. An exploratory user study involving 30 participants is conducted to confirm the developed proactive prototype's viability. The results show that the users did not achieve bidirectional interaction and unable to realize the prototype's proactive behavior.

Based on the perceptual crossing paradigm, even if the communication channel is reduced to a bare minimum (i.e., touch perception as the perceptual quality), two intentional entities can still recognize each other as long as they experienced the same perceptual environment. These perceptual activity outcomes suggest that people can interact with each other by depending on a single interaction method if they are involved in the same perceptual environment. Therefore, with visual attention as the only perceptual quality used to interact with the object (i.e., eyes), the object should also show the same perceptual quality to enable the interaction to occur in the same perceptual environment. Therefore, in the first study, the absence of visible expressive perceptual quality makes the users unable to experience bidirectional interaction and understand the object's proactive behavior.

In the second study, an improvised prototype mounted with a visible expressive perceptual quality similar to that of human eye contact (i.e., artificial eyes) is introduced. A camera module is used to detect the user's visual attention and is mounted together with artificial eyes. Hence, the proposed artificial eyes allow the prototype to create back-and-forth eye-to-eye contact interaction with the intentional user, and therefore, bidirectional interaction can be achieved. An exploratory user study experiment is conducted to validate the human-object bidirectional interaction. The results show that the artificial eyes staring with blinking expression help the user experience bidirectional interaction and engagement with the prototype. However, to maintain a continuous bidirectional interaction, the artificial eyes need to express proactive behavior besides staring and blinking. Therefore, the artificial eyes are proposed with proactive expressions such as winking and pupil dilation to allow the user to maintain continuous bidirectional interaction with the prototype.

In the third study, a conceptual model called Session Initiation for Proactive Object (SIPO) based on the perceptual crossing paradigm is proposed. The SIPO conceptual is referred to achieve bidirectional and proactive interaction between the object and the user. To pre-evaluate the SIPO conceptual model in single- and multi-user scenarios, two pilot studies which are, 1) visibility of the perceptual quality and 2) expression of intentions, are conducted. The SIPO conceptual model is implemented into a prototype mounted with artificial eyes and a camera module on top. Abstract motion is integrated to allow the prototype to orient its body towards the users. The pilot studies in single- and multi-user scenarios are conducted to validate the SIPO conceptual model. The pre-evaluation results show that the prototype mounted with artificial eyes and abstract motion successfully achieved the bidirectional interaction and engagement with the users in single- and multi-user scenarios. The achieved results only show bidirectional aspects instead of bidirectional and proactive interaction and engagement. The study was then expanded, involving 28 participants using a real-environment user study. The real-environment user study introduces a primitive physical object mounted with artificial eyes that express staring, winking, blinking, and turning behaviors. These four different expressive behaviors are analyzed to validate the object's proactive behavior towards the users in single-and multi-user scenarios. The results were taken into account for implementing a virtual object mounted with artificial eyes. A crowd-sourced video-based user study involving 240 participants with the virtual object expresses staring, winking, blinking, and turning behaviors is conducted. The results show and validate that the winking expression successfully revealed the object's proactive behavior towards the interacting and engaging user and encourages reciprocal input. Therefore, in reflections of the achieved results, this research successfully implemented, evaluated, and validated the bidirectional and proactive interaction and engagement between the object and the user based on the perceptual crossing paradigm.

In conclusion, this thesis presents a perceptual crossing design with artificial eyes based on the perceptual crossing paradigm to improve human-object interaction and engagement. An initiative-taking was adopted. An object was augmented with a visible perceptual quality (i.e., artificial eyes) which provides the ability to express its desire to interact and engage with the person of interest. The design and development of the object are further improved by introducing the SIPO conceptual model, which enhances the object's initiative-taking and the object's proactive communication to maintain human-object interaction and engagement. The results show that the artificial eyes winking expression is proactive towards the interacting and engaging person. This thesis's presented work could be a starting point for designers to develop a practical yet straightforward bidirectional and proactive interaction design based on the perceptual crossing paradigm to improve the human-object interaction and engagement.

Table of Contents

List of Tables

List of Figures

Chapter 1.

Introduction

The ongoing trend of incorporating computing capabilities on everyday objects has prompted the increase of smart objects in the market. While standard practices in developing smart objects are technology-focused, there has been little discussion on improving smart objects-users interaction (Petrov *et al.*, 2017). It is seen that the lack of interaction between the objects and the users impedes the objects from being smarter and understanding the users' requirements. Nonetheless, with their sensing, computing, actuating, and communication capabilities, these smart objects can function on their own, gradually reducing the need for human intervention (Fortino *et al.*, 2018). As a result, these smart objects that work quietly and *disappear* into the background (Atzori *et al.*, 2017) are perceived as passive-reactive objects (Zualkernan *et al.*, 2020). The quality and manner of interaction between smart object and the user are parameters not many explore (Petrov *et al.*, 2017). Therefore, there are not many discussions that focus on the matter. Solutions such as providing sufficient feedback and feed-forward (Chuang *et al.*, 2018) to indicate the objects' internal operation state and facilitating interaction with engaging users would enhance and improve the interaction between the smart object and the user. Nevertheless, most human-object interaction design adopts one-directional communication, where the human acts as the manipulator or initiator when interacting and engaging with the objects. Thus, it is questionable and remains arguable to understand the objects' smartness that could initiate and continuously interact with the users.

1.1. The Perceptual Crossing Paradigm

In studying social interaction dynamics, an intentional subject's interpretation is built upon how a person judges and predicts another person's behavior. From the observed behavior, the person decides whether the behavior was animated by intentional motives. The person who judges and predicts the behavior is only observant and does not interact with the subject matter. However, in real-life, under what conditions does a person recognize the presence of an intentional entity? Is a person able to distinguish *interaction with an intentional entity* and *interaction with a reactive or an autonomous entity*?

Triggered by these questions, Auvray *et al.* (2009; 2019) conducted a minimalist perceptual crossing paradigm experiment to investigate social interaction dynamics between two participants in one-dimensional space. In this experiment, two blindfolded participants were placed in separate rooms and were asked to explore a one-dimensional virtual space using a computer mouse. In the virtual space, a participant can interact with three different entities; an intentional entity (the other participant), a reactive entity (non-moving object), and an autonomous entity (an object that moves on its own). During interaction with these entities, a participant will receive a tactile stimulus on his/her index finger. For instance, when a participant bumps into a reactive entity while exploring the one-dimensional space, the participant will feel a vibration on the index finger. Both the participants in this experiment were asked to click the mouse button when they think they interact with each other in the one-dimensional space (i.e., both are intentional entities). When a participant interacted with a reactive entity, it creates a stable, unidirectional reactive interaction. Since it is a non-moving object, it gave the same tactile stimulus when the participants oscillate around the reactive entity (i.e., it reacts based on the participant's presence). When a participant interacted with an autonomous entity, it creates an unstable and unpredictable interaction. Since it is an object that moves on its own, the perceived tactile stimulus disappears even if the participants did not move their mouse. However, when a participant interacted with the other participant (i.e., intentional entity), it creates an unpredictable bidirectional interaction. The results of the experiment showed that participants clicked more often when they encountered each other. Since they received continuous back-and-forth tactile stimulation at each encounter, they manage to prolong the interaction and move together along the one-dimensional virtual space. This behavior shows that the participants' ability to distinguish intentional entity from reactive and autonomous entity was not because of individual strategy but from active-proactive perceptual activities that influenced the participants to coordinate their behavior, thus creating a stable bidirectional interaction. The outcome of this experiment demonstrates the different experiences of interactions between a person and an entity that able to communicates (i.e., intentional entity), react (i.e., reactive entity), or autonomously function (i.e., autonomous entity) during the person-entity engagement. The outcome of this paradigm can be used to improve the human-object interaction. With that, the next primary question that should be asked is,

how can a person recognize the presence of an intentional object that wants to interact and distinguish it from a reactive or autonomous object?

1.2. Designing for Perceptual Crossing

The perceptual crossing paradigm experiment has provoked a great deal of resonance in various fields of research, including experimental psychology (e.g., Barone *et al.*, 2020; Rini and Ochoa, 2020), philosophy (e.g., (Schönherr and Westra, 2017; Abramova and Slors, 2019)), psychopathology (e.g., Zapata-Fonseca *et al.*, 2019,2018), and not least of all computer/robot modeling (e.g., Saitoh *et al.*, 2017; Lenay, 2017). The perceptual crossing paradigm has also influenced researchers in interaction design to look further into involvement and empathy of human-object interaction. Deckers *et al*. (2011; 2013) designed an artifact embedded with different perceptive behaviors in the form of dynamic light movements when a subject's presence is detected. The artifact was tested under the theory that if perceptual crossing occurs between subject and object, the subject's sense of involvement increases. The results showed that experiencing perceptual crossing strongly affects the subject's feeling of being involved with the artifact. Marti (2012; 2020) developed a companion robot capable of experiencing perceptual crossing with a child. The objective was to stimulate the child's reflection during playtime and learn social competence with a companion robot. (i.e., reactive object). Findings from this experiment showed that the interaction with reactive objects such as the companion robot played a significant role in mediating the children's social skills. The children were observed to be engaged in exploring the features of the robot through their physical skills. Following the framework by Deckers *et al*. (2011; 2013) and Marti (2012; 2020), several other researchers developed an interactive object for improving the human-object relationship. For example, Liu *et al.* (2017) created expressive single-point light patterns to understand possible communications between humans and smart objects. Alexander *et al*. (2018) formulate feed-forward loops between a person's action and an object's response that influences the person's behavior. Chung *et al*. (2018) explore perceptive qualities that enable a person to have an enriched social interaction with the object through different sensory outputs, and Marti *et al.* (2014; 2020) explored the qualities of the perceptual crossing by developing three interaction designs to create a meaningful relationship between a robot and a person.

The framework of Deckers *et al*. (2011; 2013a) and Marti *et al.* (2012; 2020) followed by Liu *et al.* (2017), Alexander *et al*. (2018), Chung *et al*. (2018), Marti *et al.* (2020, 2014) was focused mainly on the human presence and engagement. This framework only illustrated the unidirectional activity of the reactive human-object interaction and not the intentional object's intention. Nonetheless, studies by Froese *et al.* (2009; 2020), Lenay *et al.* (2012), Lizuka *et al.* (2012), Deschamps *et al.* (2016), and Hermans *et al.* (2020) were variations of Auvray *et al.* (2009; 2019) perceptual crossing paradigm experiment showed that to distinguish the interaction between an intentional subject and a reactive object, the communication between two entities must be bidirectional and proactive. Therefore, to experience perceptual crossing with an object, the object has to proactively initiate communication with a person by expressing its intention to communicate. Therefore, to exploit the perceptual crossing paradigm into the human-object interaction design practice, the following questions also need to be addressed:

How can an object proactively initiate interaction and maintain a stable and continuous bidirectional engagement with an intentional person?

What is the signaling protocol that can be implemented for an object to proactively initiate interaction and maintain a continuous bidirectional engagement with an intentional person?

These two questions are essential in perceptual crossing interaction design practice to analyze the bidirectional and proactive human-object interaction. In this thesis's work, these questions were answered, and the necessary parameters to improve human-object interaction design practice based on the perceptual crossing paradigm were investigated. Proactive bidirectional behavior and reactive unidirectional behavior interactions were also addressed.

1.3. Aim and Objective

This thesis aims to create an effective bidirectional and proactive interaction design based on the perceptual crossing paradigm to improve human-object interaction and engagement. The objectives of this thesis are to:

1) Develop a proactive object prototype that can detect a user's visual attention and initiate bidirectional and proactive interaction with the user. Visual attention as an interaction method is chosen to allow the object to differentiate an intentional user from a potential user. An exploratory user-experience experiment will be conducted to collect preliminary feedback of user-experience on human-object interaction.

2) Improve the developed object prototype by integrating visible expressive perceptual quality to enhance the object's bidirectional interaction and engagement with a person. An exploratory user-experience experiment will investigate the user's experience of human-object bidirectional interaction and engagement.

3) Develop a conceptual model to conceptualize the bidirectional and proactive human-object interaction and engagement. The conceptual model will be adapted into an object prototype. A real-environment user study and crowd-sourced video-based user study will be conducted to

validate the bidirectional and proactive human-object interaction and engagement based on the developed conceptual model.

Each objective has to be individually achieved and analyzed before executing the next objective. Each objective's results will be explored through the collective inputs based on the participants' received feedback, which would validate the bidirectional and proactive interaction design based on the perceptual crossing paradigm.

1.4. Thesis Outline

This thesis comprises six chapters, including the current chapter. The outline of this thesis is presented as follows:

- Chapter 2 describes the literature research study on the perceptual crossing paradigm and previous human-object interaction designs related to the perceptual crossing paradigm. A discussion on the overview of mixed-initiative human-computer interaction and attentive user interface (AUI) are presented to develop the human-object perceptual crossing interaction design. A Session Initiation Protocol (SIP) is studied and adapted to create the human-object perceptual crossing interaction design model. The chosen interaction modality that allows a person to experience perceptual crossing with an object was also discussed in this chapter.

- Chapter 3 presents the study addressing the first objective: developing a proactive object prototype with bidirectional interaction and proactive behavior. The conducted exploratory user study and the analyzed feedback contribute to the improved proactive object prototype described in Chapter 4.

- Chapter 4 presents the study addressing the second objective: to improve the proactive object prototype with visible expressive perceptual quality similar to human perceptual quality (i.e., artificial eyes). The prototype is tested to confirm its viability and the conducted exploratory user-experience shows positive improvements in the user and the object's bidirectional interaction. A further enhancement of the object prototype is proposed to allow the object to interact with the user proactively.

- Chapter 5 presents the study addressing the third objective, focusing on investigating and exploring the proactive object prototype's perceptual crossing design practice. The work is expanded into multi-user scenarios, and a conceptual model of Session Initiation for Proactive Object (SIPO) is developed. This chapter also features a 28-participant real-environment user study and a 240-participant crowd-sourced video-based user study to investigate useful perceptual activity that makes a person feel the object is proactive.

- Chapter 6 concludes the research work that has been conducted to achieve the bidirectional and proactive human-object interaction design based on the perceptual crossing paradigm. This chapter also presents the future work possibilities and summarise the contributions of the work presented in this thesis.

Chapter 2.

Literature Review

This chapter[1] describes bidirectional and proactive interaction approaches and designs. Section 2.1 gives an overview of the perceptual crossing paradigm concerning human-human interaction, which later translates to human-object interaction. Section 2.2 and Section 2.3 present another interaction approach: the flexible mixed-initiative human-computer interaction and attentive user interface, which uses turn-taking as the interaction strategy. Section 2.4 elaborates previous human-object interaction designs related to the perceptual crossing paradigm. A Session Initiation Protocol (SIP) was used to initiate, manage, and terminate real-time interaction sessions. The protocol is explained in Section 2.5. Section 2.6 mentions previous interaction modalities of human-object perceptual crossing interaction design.

2.1. The Perceptual Crossing Paradigm

The perceptual crossing paradigm is a paradigm that offers the most straightforward paradigm wherein the most basic conditions; a person manages to differentiate an intentional subject from a reactive object in a one-dimensional virtual space. The paradigm was first described by Auvray *et al.* (2009; 2019). This paradigm is based on a real-time interaction study in which two participants were blindfolded and seated in separate rooms. They sit at a desk with the right hand handling the mouse and the left hand placed on a tactile stimulator, as shown in Figure 1a. Participants were instructed to explore a one-dimensional virtual space using a computer mouse. During the exploration, there were three objects that a participant will encounter, which are 1) the other participant's body object, 2) a fixed object, and 3) a shadow image, which movement is identical to the other participant's movement. As illustrated in Figure 1a, body objects represented the participants' mouse movement; Participant 1 (P1) is the green-colored body object. Participant 2 (P2) is the blue-colored body object, and the opposite side of each body object is the receptor fields. The tactile stimuli trigger when one of the three objects is in the receptor field. When the participants move the mouse, their respective body objects and receptor fields move too (Figure

[1] This chapter is written partly based on the following publications (Anas et al., 2016; Anas et al., 2017; Anas et al., 2020).

1b). Both participants would receive tactile stimuli if they encountered the three objects in the one-dimensional virtual space. The only difference between the body object and the shadow image is when both participants cross with each other, both receiving tactile stimulation. Although the participants were told to explore a one-dimensional space, the explored space is a circle embedded with the three objects. Only the hand movement handling the mouse is one-dimensional. The actual virtual space is circular to allow a smooth continuous exploration movement, as shown in Figure 1c. The perceptual crossing paradigm's objective is to identify whether the perceived tactile stimuli in the one-dimensional virtual space represent the participant's body object, the fixed object, or the shadow image.

Figure 2 shows the participant's movement behavior when encountered a non-moving, fixed-object. Figure 2a shows P1 and P2 exploring the one-dimensional space. Figure 2b shows P1 and P2 encountered static objects. Figure 2c-f shows P1 and P2 oscillates around the static object. The perceived tactile stimuli pattern is always identical when the participant interacts with the static-object. Hence, the participant can easily predict the interaction with the fixed-object.

Figure 3 illustrated the participants interacting with a shadow image. A shadow image is an object that follows the movement of the body object at a constant distance. Figure 3a shows P1 exploring the one-dimensional space. In Figure 3b, as P1 moves to the right of the space, S1 moves into the receptor field of P2, and P2 receives a tactile stimulation. When P1 moves further to the right, S1 moves further out of the receptor field of P2, as shown in Figure 3b-d. The perceived tactile stimuli pattern is always unstable when the participant interacts with the shadow image because the perceived tactile stimuli kept moving away. Hence, the participant experiences inconsistent interaction with the shadow image.

The perceptual crossing paradigm's valuable outcome is illustrated in Figure 4, where the participants encountered the other participant's body object. A body object is an object that represents the participant. Figure 2a shows P1 and P2 exploring the one-dimensional space. When P1 and P2 encountered each other, as shown in Figure 2b, Figure 2d, and Figure 2f, both P1 and P2 perceived the tactile stimulation. P1 and P2 experienced continuous back-and-forth tactile stimuli when encountering the body object. The perceived tactile stimuli pattern is always different when the participant interacts with the body object. Hence, the participant experiences an unpredictable, bidirectional interaction with the body object.

In conclusion, the participant's interaction with the three objects, i.e., fixed object, shadow image, and body object, yield different experiences. In the interaction with a fixed object, the interaction is stable as the fixed object's location is quickly established. The perceived stimuli were always the same when the body object oscillates around it. In contrast, the interaction with the shadow image produces unstable interaction. Often after a participant encountered the other participant's shadow image, the participant tends to go back to search for the shadow image. However, the perceived tactile stimuli keep moving, creating an unstable interaction; hence the shadow image can seldom be found. For the body object interaction, the participant distinguished the perceived tactile stimuli produced by another participant from a static/moving object. The participants experienced bidirectional back-and-forth interactions, which produced different tactile stimuli patterns.

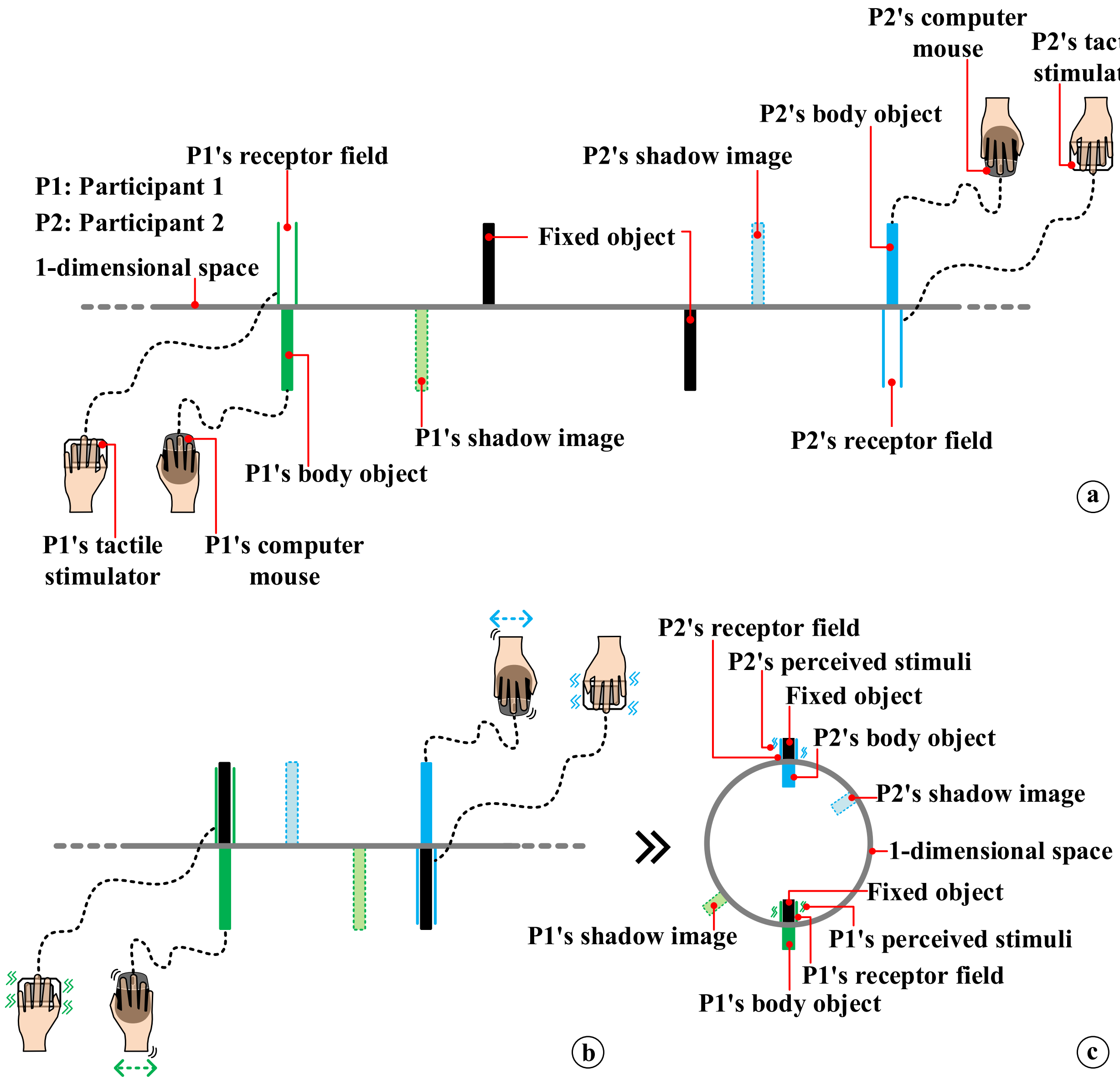


Figure 1 Tactile perceptual crossing experimental setup (Auvray et al., 2009; Auvray 2019): (a) Illustration of the perceptual crossing paradigm virtual environment. (b) Participants encountered a fixed object. (c) A Simplified version of one-dimensional virtual space the perceptual crossing paradigm experiment.

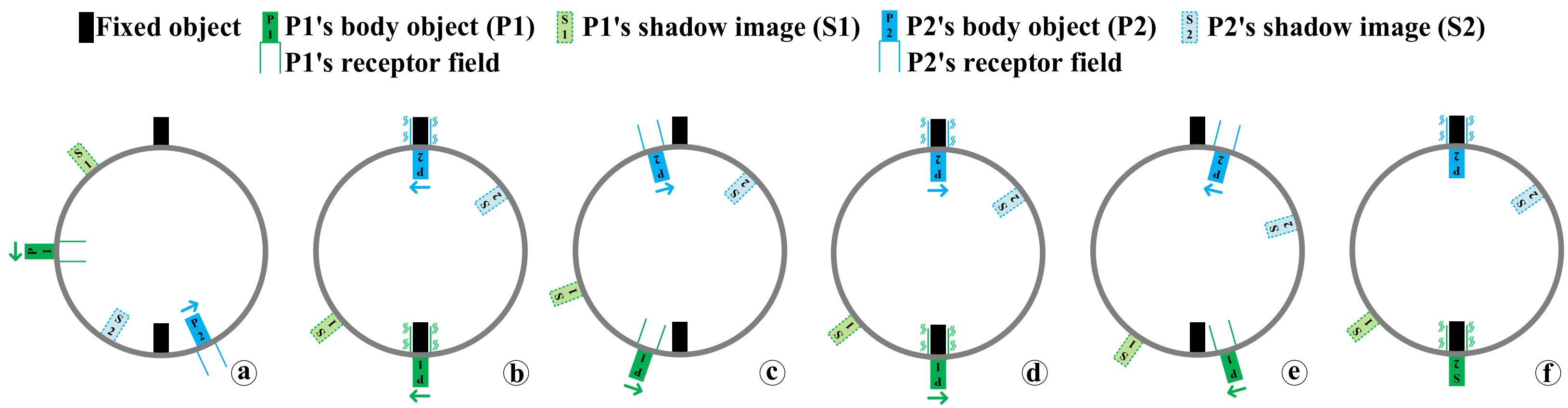


Figure 2 Both P1 and P2 encountered a fixed object in the virtual space (Auvray et al., 2009; Auvray 2019).

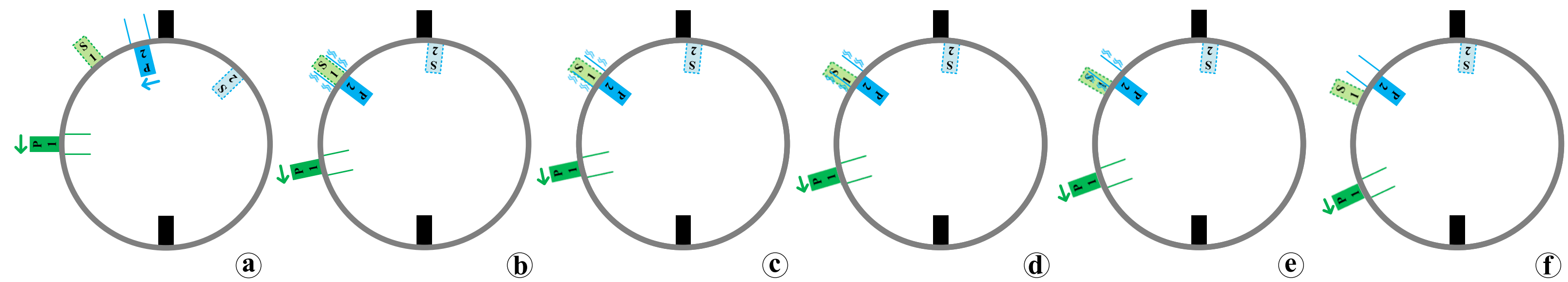


Figure 3 P2 encounter P1's shadow image in the virtual space (Auvray et al., 2009; Auvray 2019).

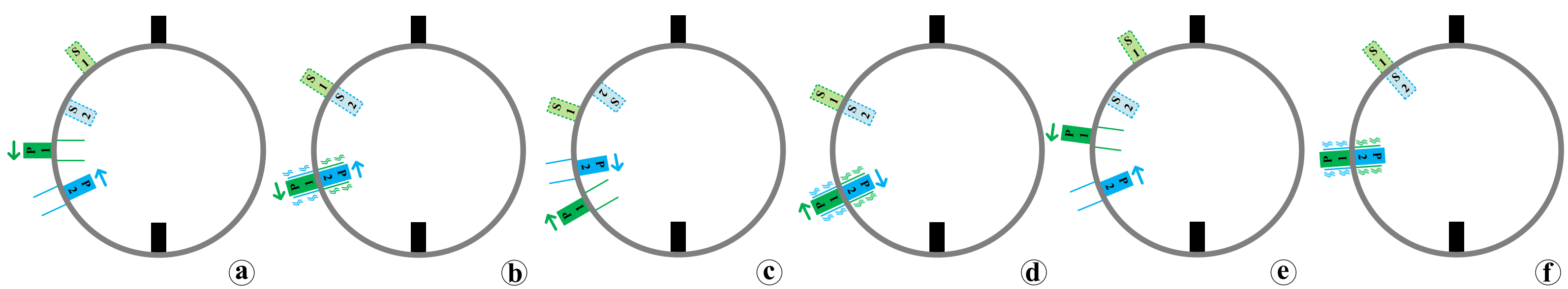


Figure 4 Both P1 and P2 encountered each other in the virtual space (Auvray et al., 2009; Auvray 2019).

The perceptual crossing paradigm was further investigated in a different experiment conducted by Froese *et al.* (2009; 2020). The experiment was based on 1) participant's body object – participant's body object and 2) participant's body object – participant's shadow image scenarios. Instead of using real subjects, this experiment uses artificial agents to perform both scenarios. In Froese *et al.* first scenario experiment, both artificial agents need to locate each other and maintain a continuous interaction until the end of the trial session (Figure 5a). The result obtained for the artificial agent – artificial agent scenario experiment shows both agents can locate each other and maintain a continuous back-and-forth interaction until the end of the trial session (Figure 5b). However, scholars have argued that both agents seemed to develop their strategies to maintain the stimulation and engagement duration to differentiate the perceived tactile stimuli. Hence, the application of the agent's strategy does not concretely define bidirectional perceptual crossing paradigm interaction. Thus, a new scenario of artificial agent – artificial agent's shadow image (Figure 5c) was conducted to justify the bidirectional perceptual crossing paradigm interaction claimed in the previously conducted artificial agent – artificial agent scenario experiment. However, it turned out that the artificial agent in the artificial agent – artificial agent's shadow image experiment singly responded towards the artificial agent's shadow image, without any proactive response from the artificial agent's shadow image. This result shows that the artificial agent communicated unidirectionally with the artificial agent's shadow image. Due to this, the artificial agent failed to stay in contact with the artificial agent's shadow image.

Lenay *et al.* (2012; 2017) conducted a similar perceptual crossing paradigm experiment to allow the participants to engage with each other in different trajectories (space and time). The experiment results show that the *act of anticipation* was only formed when the perceived tactile stimuli occur at uniform intervals during the interaction with a fixed object. Therefore, the participants' consistent tactile stimulation will be perceived during active interaction with a fixed object. In contrast to the *act of anticipation*, the *act of surprise* result shows different tactile stimuli were perceived during interaction with the encountered body object and shadow image. Thus, this result shows the participants are in contact with either the encountered body object or shadow image. Therefore, to recognize interaction with body objects, the participant must retain their bidirectional engagement by maintaining a proactive and uniform tactile stimuli engagement. Otherwise, the perceived tactile stimuli engagement is with the shadow image. Hence, different tactile stimuli were perceived during the *act of anticipation* and the *act of surprise*.

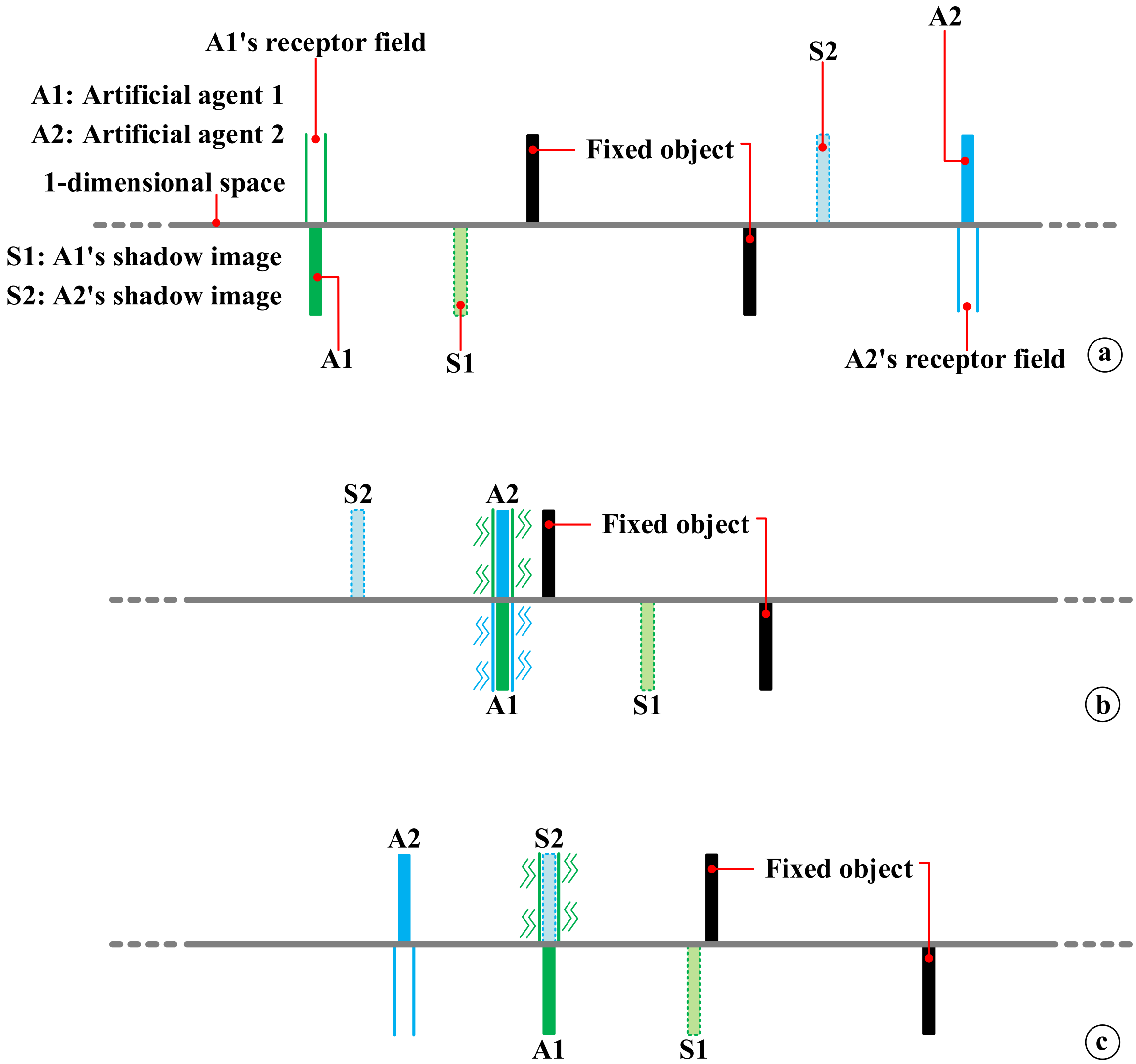


Figure 5 Illustration of Froese et al. (2009;2020) the perceptual crossing paradigm virtual environment: (a) Artificial agents performing the perceptual crossing paradigm experiment. (b) Artificial agents encountered each other in the virtual space. (c) An artificial agent encounters the artificial agent's shadow image in the virtual space.

In another experiment conducted by Lizuka *et al.* (2012; 2016), the static object and the shadow image were removed. This experiment adopted the perceptual crossing paradigm and was conducted using real-time body object and recorded body object simulation. These two body objects condition was used to let the participant distinguish between real-time perceive tactile stimuli or recorded tactile stimuli. P1's body object oscillates at P2's body object during real-time interaction condition, while P2's body object stays stagnant, recognizes, and responds to the P1's body object oscillatory behavior. This situation is known as turn-taking interaction. P1's and P2's repetitive turn-taking behavior determines whether turn-taking interaction occurs in real-time or vice versa. This experiment was further evaluated using different shapes such as Sharp (#) and

Square (□) to understand the turn-taking interaction. The different shapes represent different oscillation patterns; the sharp shape represents frequent and fast oscillation; meanwhile, the square shape represents slow oscillation. There are four interaction conditions: a) Condition A: # versus #, b) Condition B: □ versus □, c) Condition C: □ versus #, and d) Condition D: # versus □. When condition A occurs, participant P1 and P2 observe the # shape, and both individually produces slow oscillation to each other to indicate turn-taking interaction. When condition B occurs, participants, P1 and P2, observe the □ shape and produce frequent and fast oscillation to indicate turn-taking interaction. When condition C occurs, participant P1 observes the □ shape which produces slow oscillation, and participant P2 observes the # shape, which produces frequent and fast oscillation to each other to indicate turn-taking interaction. When condition D occurs, participant P1 observes the # shape, which produces frequent and fast oscillation, and participant P2 observes the □ shape, which produces slow oscillation to each other to indicate turn-taking interaction.

Deschamps *et al.* (2016) and Hermans *et al.* (2020) define the perceptual crossing paradigms within a triadic situation since dyadic situations do not always appear in real-life scenarios. A triadic situation is known as triadic interaction in which two participants coordinate their attention towards a reference point which can be an object, an event, or a third participant (Siposova and Carpenter, 2019). Therefore, to investigate the interpersonal coordination element in perceptual crossing activities during triadic interaction, the body-objects shown in Figure 1a were removed and replaced with one fixed-shared object, as shown in Figure 6. This experiment was to differentiate if space's encountered objects was a fixed object or a shared object. The presented configuration in Figure 6 also shows that both participants can simultaneously interact with each other and the shared object. The experiment result shows that the joint activities such as back-and-forth actions at the shared object through collaboration and coordination enable them to recognize the perceived object.

This paragraph concludes the conducted background studies in Section 2.1. The perceptual crossing paradigm experiments are designed to instigate a tactile perceptual interaction where an involved person differentiates the perceived tactile stimuli produced by the other person from those produced by the distractor objects placed in the space (fixed object and shadow image). The experiment results show when the partner and partner's body-object encounter each other, both develop and establish a particular strategy to achieve bidirectional and proactive interaction to communicate with each other. Thus, to experience perceptual crossing and maintain engagement based on the perceptual crossing paradigm, the interaction between two entities must be bidirectional and proactive.

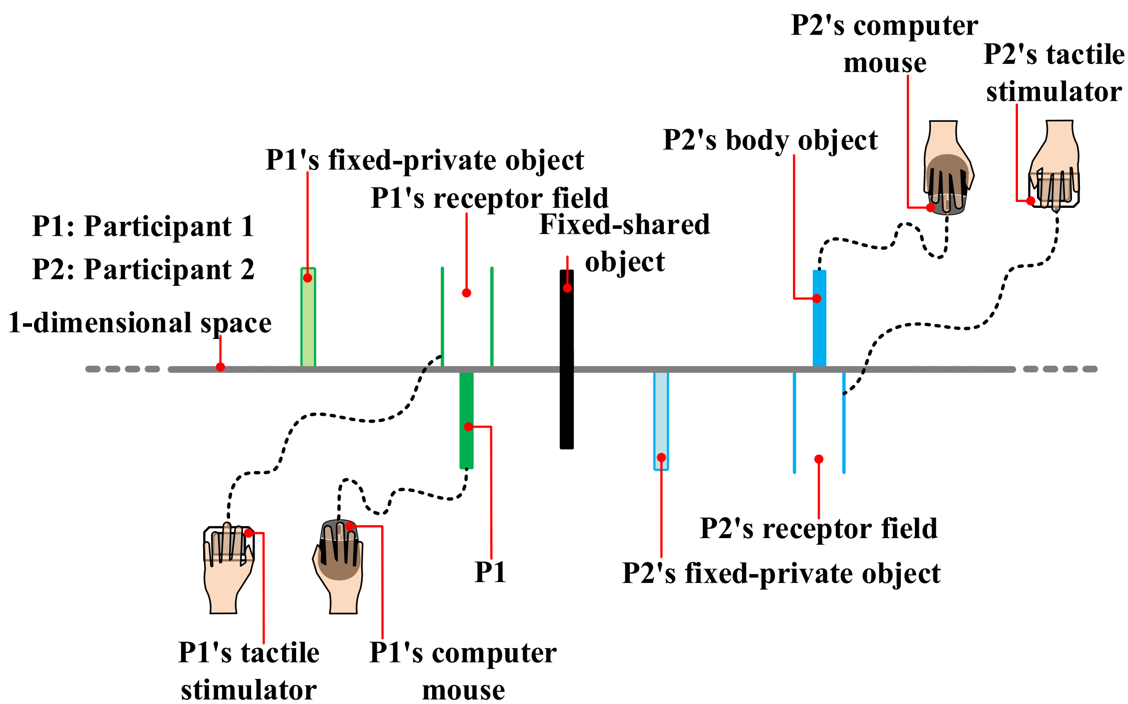


Figure 6 Illustration of the perceptual crossing paradigm virtual environment uses one fixed-shared object to investigate triadic interaction (Deschamps et al., 2016; Hermans et al., 2020).

2.2. Mixed-Initiative Interaction

Similarly, mixed-initiative interaction adopts a bidirectional and proactive turn-taking interaction strategy (Buck *et al.*, 2018). The mixed-initiative interaction approach allows either human or machine to initiate an interaction and turn-taking proactively to achieve a common goal. Figure 7 shows the occurrence of flexible bidirectional interaction (initiate and respond) between the object and the subject and vice versa. During mixed-initiative interaction, either the object initiates, and the subject responds, or the subject initiates, and the object responds. Thus, a bidirectional turn-taking interaction is initiated between the object and the subject or vice versa. However, for mixed-initiative interaction to be successful, the system requires various initiative strategies to cater to various individuals with different preferences, which is the most challenging task when implementing mixed-initiative interaction (Galitsky and Ilvovsky, 2020). Whereas in the perceptual crossing paradigm, two entities can communicate effectively using a minimal channel of interaction as long as the interaction between the two entities is bidirectional and proactive.

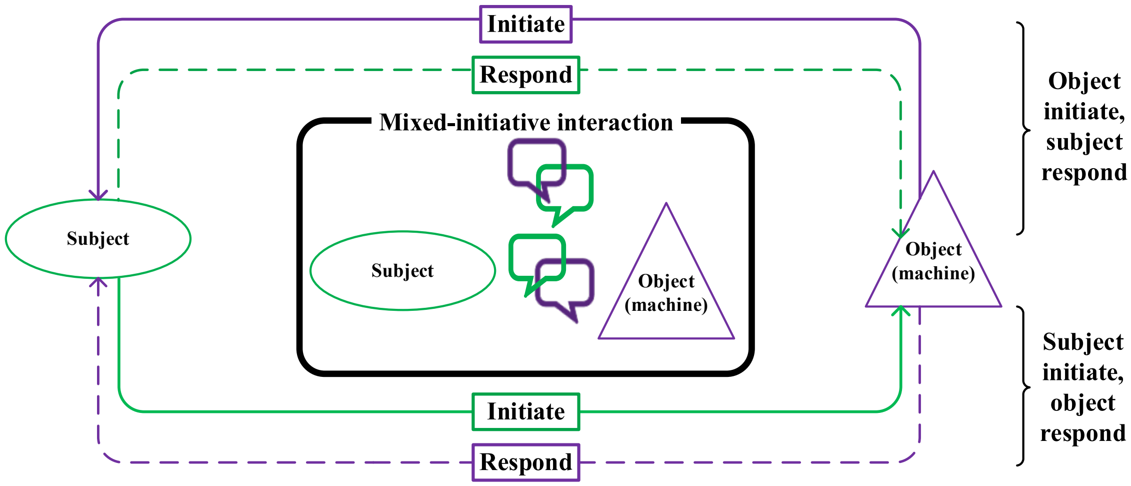


Figure 7 Schematic illustration of mixed-initiative interaction.

2.3. Attentive User Interface

Attentive User Interface (AUI) is a user interface paradigm that proactively attends to user actions (Vertegaal, 2003). Fundamentally, AUI starts by monitoring user behavior. Then, the AUI computational system model the user's goal and interest based on the monitored behavior. AUI can proactively anticipate the user needs through the developed model and provide the required information before the user explicitly requests it. AUI can also structure the information priority before proactively initiating interaction with the user. In this way, AUI can independently distribute the required information based on its attentional capacity to avoid demanding frequent attention from the user.

Like the perceptual crossing paradigm, AUI also adopts a bidirectional and proactive interaction strategy to communicate with the user. However, the computational system of AUI requires a high-level understanding of user behavior (Amershi *et al.*, 2019). It would be beneficial to adopt AUI into the human-object interaction to improve the communication strategy between the object and the user. Nevertheless, due to AUI's advanced interaction design, the perceptual crossing paradigm, which offers the most straightforward paradigm, is considered. Based on the perceptual crossing paradigm experiment, by reducing the communication channel to a bare minimum, a person can effectively communicate with another person as long as they experience a bidirectional and proactive interaction.

2.4. Perceptual Crossing Interaction Design

Several studies have been implemented in interaction design to add richness to human-object interactions by making the perceptual crossing paradigm their central concept.

2.4.1. Perceptual crossing for human-object interaction

Deckers *et al.* (2011) started their work by implementing theories on the perceptual crossing paradigm into the design of an interactive object called the Perception Pillar (PeP). The square PeP consists of four light sensors attached to each of the pillar's four sides. The light sensors are attached to allow the PeP to sense an intentional subject. The top of the PeP is also fitted with 13 evenly arranged and spaced LEDs. These LEDs will light up, searching for any intentional subject. They will follow the detected subject in its vicinity, thus creating an impression of PeP exploring and interacting with the subject. The objective is to create active bidirectional interaction between the PeP and the intentional subject. The user study results show PeP positively influenced the subject's involvement. After that, Deckers *et al.* continued to improve the design of PeP by adding extra sensors and LEDs (recognize as PeP+) to gain more detailed information about the intentional subject near the pillar (Deckers *et al.*, 2013). Three design behaviors of PeP+ were being implemented and investigated: 1) the following behavior, PeP followed the moving subject, 2) the dynamic behavior, PeP+ actively explores a different side of its pillar until it found and followed the subject, and 3) the explorative behavior, PeP+ displayed micromovement from left to right to anticipate the movement of the focused subject. The outcome of this work shows that to design for perceptual crossing, the object needs to recognize the existence of an intentional subject and engage with the subject in active reciprocal interaction.

2.4.2. Perceptual crossing for human-robot interaction

The perceptual crossing paradigm also inspires Marti (2012) and Marti *et al.* (2020) to develop an interactive robot capable of co-regulate its behavior with a person. This work's main focus was to initiate behavior change through several design scenarios by making a robot coordinate and mirror a child's movements with a mild cognitive disability. The results show the children's ability to maintain their attention and behave appropriately, suggesting that their capacity to focus on the task was better than usual. Marti *et al.* (2013; 2020) then continued exploring how to achieve an empathic relationship between a robot and the older people during the perceptual crossing. For example, a robot will react to an older person's presence by bowing forward to create a subtle reciprocal interplay. In their recent work (Marti *et al.*, 2014; Marti *et al.*, 2020), they explored the perceptual crossing paradigm's qualities by developing three interaction designs to create a meaningful relationship between a robot and a person. Comprehensively, Marti *et al.*'s work demonstrated that perceptual crossing could positively influence an object's behavior to act appropriately according to the situation when interacting with a person.

2.4.3. Perceptual crossing for the Internet of Things (IoT)

The perceptual crossing paradigm has also been implemented for the Internet of Things (IoT) to improve smart things design structure and architecture. By observing concept videos, Chuang *et al.* (2018) managed to extract one vocabulary item concerning *I see you seeing me* in which machines were socializing with people or other machines related to the perceptual crossing. The authors then conclude that the mentioned vocabulary can help designers add social aspects when designing for IoT and improve user engagement and involvement. Another work proposed by Chung *et al.* (2018) concerning issues involving social things and the agency of things inspired them to further improved Deckers *et al.*'s work for things-to-things interaction (Deckers *et al.*, 2013). By redesigning the perceptual crossing model (Deckers *et al.*, 2013), the authors suggest and address the interplay within two social objects to explore and provoke discussion in the topics of IoT design and agency of things. However, Chuang *et al.* (2018) and Chung *et al.* (2018) proposals were not thoroughly evaluated with users in designing smart things.

2.5. Session Initiation Protocol (SIP)

Session Initiation Protocol (SIP) is a bidirectional general-purpose request/response live communication protocol for creating, modifying, and terminating sessions regardless of the type of session established (Ahson and Ilyas, 2018). Initiating a SIP requires an invitation to the recipient via various methods. Each method requires a sender's request and the receiver's acknowledgment. Analogically, SIP operates based on the example shown in Figure 8. According to Figure 8, a sender will initiate a SIP session with an expression of intention to communicate with the recipient via an *INVITE* session. The receiver will send a *Ringing* tone signal to the sender to acknowledge the *INVITE* session, also known as the *OK* signal. Upon receiving the *OK* signal from the recipient, the *INVITE* session sender will acknowledge (*ACK*) the *recipient's OK* signal. Upon the *INVITE-ACK* session's success, the *media session* will be commenced for media and data exchanging. The active

Media session is terminated after the recipient sends a *BYE* signal to the sender and the sender sends an *OK* signal to the recipient.

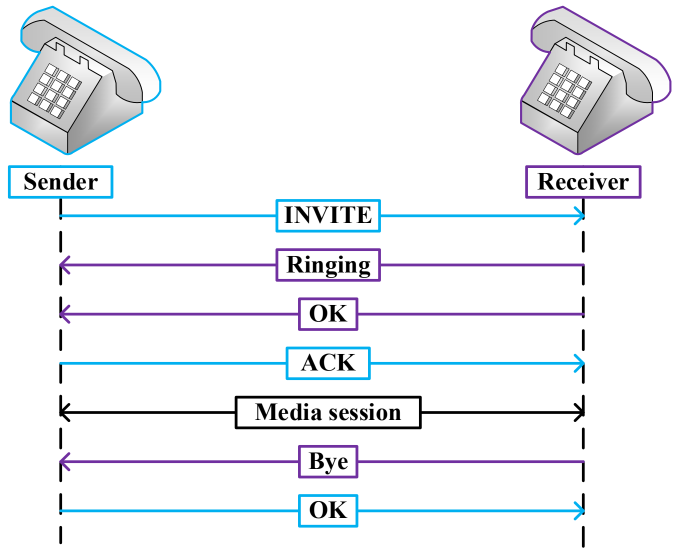


Figure 8 An example of a SIP session establishment (Ahson and Ilyas, 2018).

Therefore, to implement the perceptual crossing paradigm into the human-object interaction design, the SIP model analogy is adapted to create a protocol that allows a person to establish a perceptual crossing with an object. This protocol is proposed because it implements a bidirectional and proactive interaction session, which is the core element of experiencing perceptual crossing. Chapter 5 will further discuss the implementation of SIP into the interaction design.

2.6. Methods of Interaction for Perceptual Crossing Interaction Design

Based on the conducted literature review related to the perceptual crossing interaction design, methods such as distance measuring, motion detection, human presence, gesture detection have been used to create a bidirectional human-object interaction. However, these methods were also prone to detect objects besides humans (i.e., unintentional entity). Therefore, the used methods could not differentiate an intentional entity from an unintentional entity. Thus, detecting any presence would initiate a reaction from the object.

The perceptual crossing paradigm demonstrates the interaction with an unintentional entity produces a unidirectional-reactive interaction, whereas engaging with an intentional entity produces bidirectional-proactive interaction. Therefore, to implement the perceptual crossing paradigm to improve the human-object interaction, the object needs to recognize and differentiate an intentional entity from the unintentional entity. Referring to AUI (Vertegaal, 2003), the system understands the user's intention by continuously sense and analyze the user's visual attention. The system can also differentiate intentional requests for information from unintentional behavior by sensing the user's visual attention. Visual attention can also be used to negotiate the initiation of communications with an object (Jabarin *et al.*, 2003). Such communication can be realized using an eye tracker mounted on the object. To initiate the interaction, a person needs to direct his/her visual attention to the object to express the intention to interact or *look-away* to terminate or ignore the interaction. The mounted eye tracker allows the object to recognize an intentional subject's existence, thus avoiding initiating a reactive interaction towards the unintentional subject. Therefore, detecting visual presence with attention is an effective interaction method for allowing an object to recognize and differentiate an intentional subject from an unintentional subject.

2.7. Summary

The perceptual crossing paradigm is a paradigm that shows if a person can recognize the interaction between an intentional entity and interaction with a reactive or autonomous entity. The paradigm emphasized people's different experiences when interacting with an entity that intends to communicate and an entity solely reacting or functioning to their presence. According to the paradigm, to experience perceptual crossing, the interaction between two intentional entities must be bidirectional and proactive. In this way, the interacting person can differentiate interaction with an intentional entity from interaction with a reactive or autonomous entity.

However, the previously mentioned work (Deckers *et al.*, 2011; Deckers *et al*., 2013; Marti, 2012; Marti *et al.*, 2014) has only been limited to interacting with a reactive object where the user initiates the interaction while the object responds to the user. Even though the authors (Deckers *et al*., 2013; Marti *et al.*, 2014) mentioned that reciprocal interaction between an object and a person is crucial in designing for perceptual crossing, their designed objects only involved interaction initiated by a person. Therefore, the objects only reacted towards the person's presence. Without a systematic user evaluation, it remains unclear how an object could engage in perceptual crossing in the context of human-object interaction.

Chapter 3.

Study 1: Exploring Perceptual Crossing Interaction Design

This chapter[2] presents a user-experience object prototype development based on Auvray *et al.*'s perceptual crossing paradigm (2009; 2019). According to the paradigm, to experience perceptual crossing, the interaction between two entities must be bidirectional and proactive. Therefore, to investigate the perceptual crossing paradigm for improving human-object interaction, an object prototype capable of expressing its intention using abstract motion was developed. Visual attention was used as the interaction method. The prototype consisted of a microcontroller integrated with an eye tracker to allow the prototype to differentiate interaction with an intentional user from a potential user. When detecting a potential user's visual presence, the prototype sought the user's visual attention to initiate the interaction. An exploratory user experience experiment was conducted to gain quantitative and qualitative feedback about the user's experience engaging with the prototype. The experiment involved two groups of 15 participants. One group of participants had no prior knowledge regarding the prototype's capability of expressing its intention to interact, whereas the other group of participants has prior knowledge. The quantitative and qualitative results showed that the participants did not achieve a bidirectional and proactive interaction with the prototype through their visual attention. The absence of perceptual quality to represent the object interaction method makes the participant unable to realize and recognize the prototype's intention. Therefore, an expressive and recognizable visual feedback mechanism is introduced to improve the bidirectional interaction between the object and the participant, as explained in Chapter 4.

3.1. Object Prototype Design and Development

This section explains the object prototype's design and development based on the perceptual crossing paradigm. The prototype is proactive and capable of expressing its intention to interact with the user through their visual attention. An eye tracker sensor is used to detect the user's visual presence with attention. Apart from detecting the user's visual presence with attention, the prototype must show its intention to interact proactively. Therefore, abstract motion is integrated into the prototype to allow the prototype to interact with the user proactively. The prototype's abstract motion uses a combination of mechanical linear and rotary actuators to self-actuate itself

[2] This chapter is written partly based on Anas et al. (2016)

to express its intention to interact with the user. The following section describes the eye tracker and abstract motion. The eye tracker and abstract motion are combined to create a proactive behavior in the object prototype towards the user's visual attention.

3.1.1. Visual attention detection in a physical space

Eye-tracking is a technology that enables the user to be in control of their device by using their eyes. For example, a computer equipped with a screen-based eye tracker can see what the user is looking at, allowing them to interact with the screen-based content naturally. Therefore, to create an object that is aware of the user's visual attention, the screen-based eye tracker's ability is extended and implemented in the physical space.

The Eye Tribe Tracker (Johansen, 2015) is a device that can detect and determine the point of gaze defined by a pair of (x,y) coordinates. It comes with software that allows client applications to access the underlying eye tracker's server to obtain a real-time stream of gaze data in smoothed forms. A Java (Oracle Corporation, 2020) program using Eclipse (Eclipse Foundation, 2020) is developed to reveal the user's visual attention on the computer screen. The computer screen area is evenly divided into a 3×3 grid and mapped onto a 3×3 grid on the physical surface (Figure 9). To confirm the viability of detecting the user's visual attention on the physical surface, the collision detection test is conducted. This test is required to identify and verify the area visited by the user's eye gaze on the 3×3 grid on the computer screen matches with the 3×3 grid on the physical surface, respectively. As illustrated in Figure 10a, a user is looking at the B2 grid on the computer screen; therefore, it matches the B2 grid on the physical surface. The screen is later removed, and an object prototype can be placed onto one of these grids, as shown in Figure 10b. Thus, the placed object prototype can now detect the user's visual presence with attention.

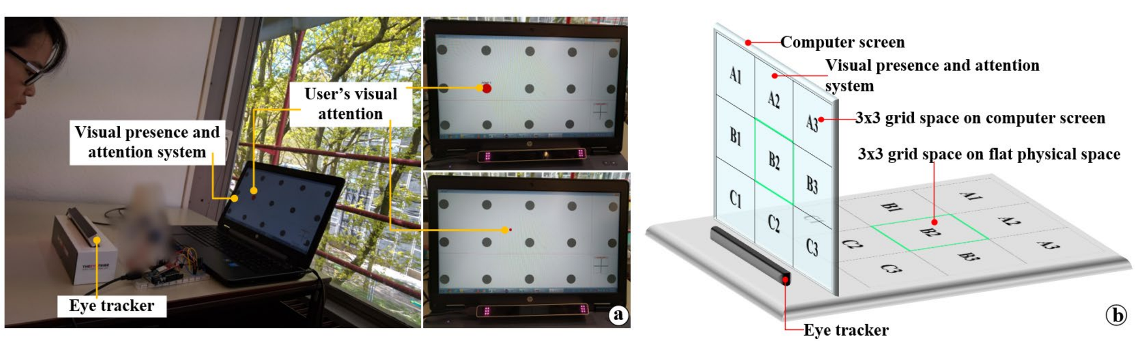


Figure 9 (a) Visual presence and attention system. (b) Setup and configuration of the on-screen eye tracker in physical space.

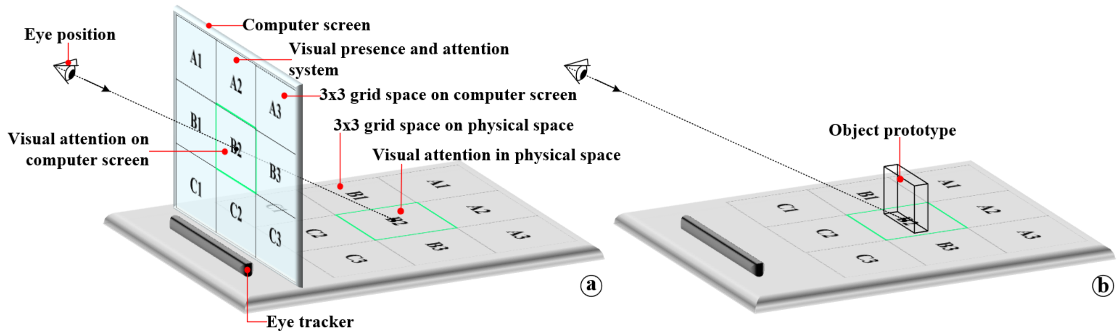


Figure 10 (a) 3×3 on-screen grid mapped onto a 3×3 flat physical surface grid. Visual attention on the computer screen grid B2 *matches grid* B2 *in physical space (b) The eye is looking at an object prototype in physical space.*

3.1.2. Expressing intention with abstract motion

Motion is one of the ways to express an intention to people. A whole-body motion cue was proven to be a reliable and easily perceived source of information about intentions (Hetherington, 2020). The prototype shown in Figure 22 is designed with the actuating mechanism to create a whole-body abstract motion. Figure 12a shows the mechanical design. It consists of racks and pinions attached to two micro servo motors for vertical movements (up or down) and a single horizontal gear attached to the third micro servo motor to produce horizontal circular movements (left or right rotations). The vertical and horizontal movements are used to express the object's intention. The low-torque micro servo motor is chosen to control the gears as it produced less noise than the noisy high-torque servo motor, which might create an attention-drawing feature to the structure. Figure 12b shows different horizontal and vertical whole-body movements being captured to indicate the object produced intention by turning its whole-body towards an engaged person.

3.1.3. A proactive object prototype implementation

The visual presence with attention and the whole-body abstract motion was used to develop the proactive object prototype. The prototype is designed to be proactive towards the user's visual attention, which initiates the object's whole-body abstract motion to express an intention to interact. For example, Figure 11 shows a coffee cup is used as an object to interact and engage with a person's visual attention. Grid B2, shown in Figure 11, shows a user is visually looking at the coffee cup to interact and engage with it.

Figure 13 shows a) the idle state of the object, b) horizontal movements when the user's visual attention is present, and c) vertical rotation movements after continuous visual attention from the user. When the object does not detect the conditions (b) and (c), it will go into the condition (a). The prototype shown in Figure 13 was used to conduct the exploratory user-experience experiment.

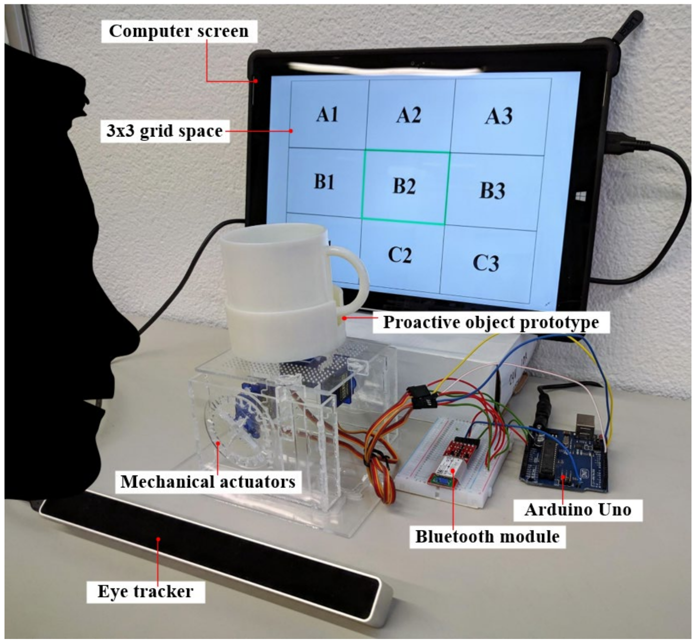


Figure 11 The proof-of-concept implementation of proactive object prototype design.

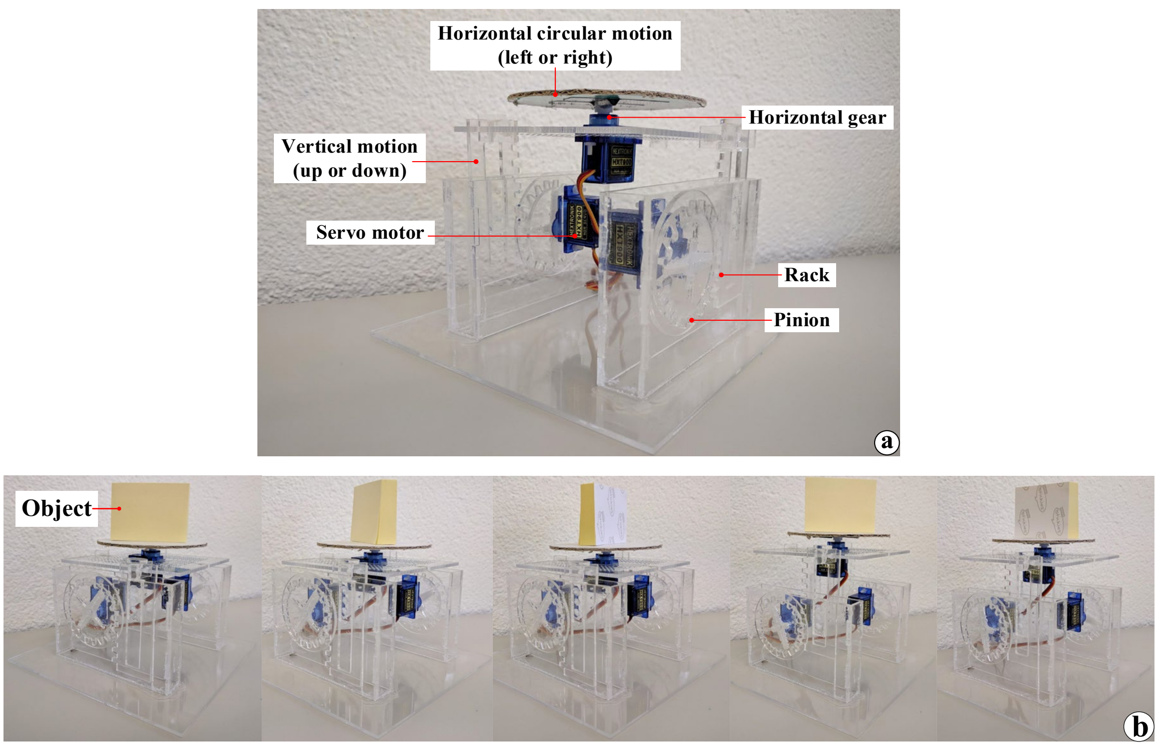


Figure 12 Mechanism of actuation: (a) Mechanical design of linear and rotary actuators. (b) Example of abstract motion to show the object's intention to create communication towards the user.

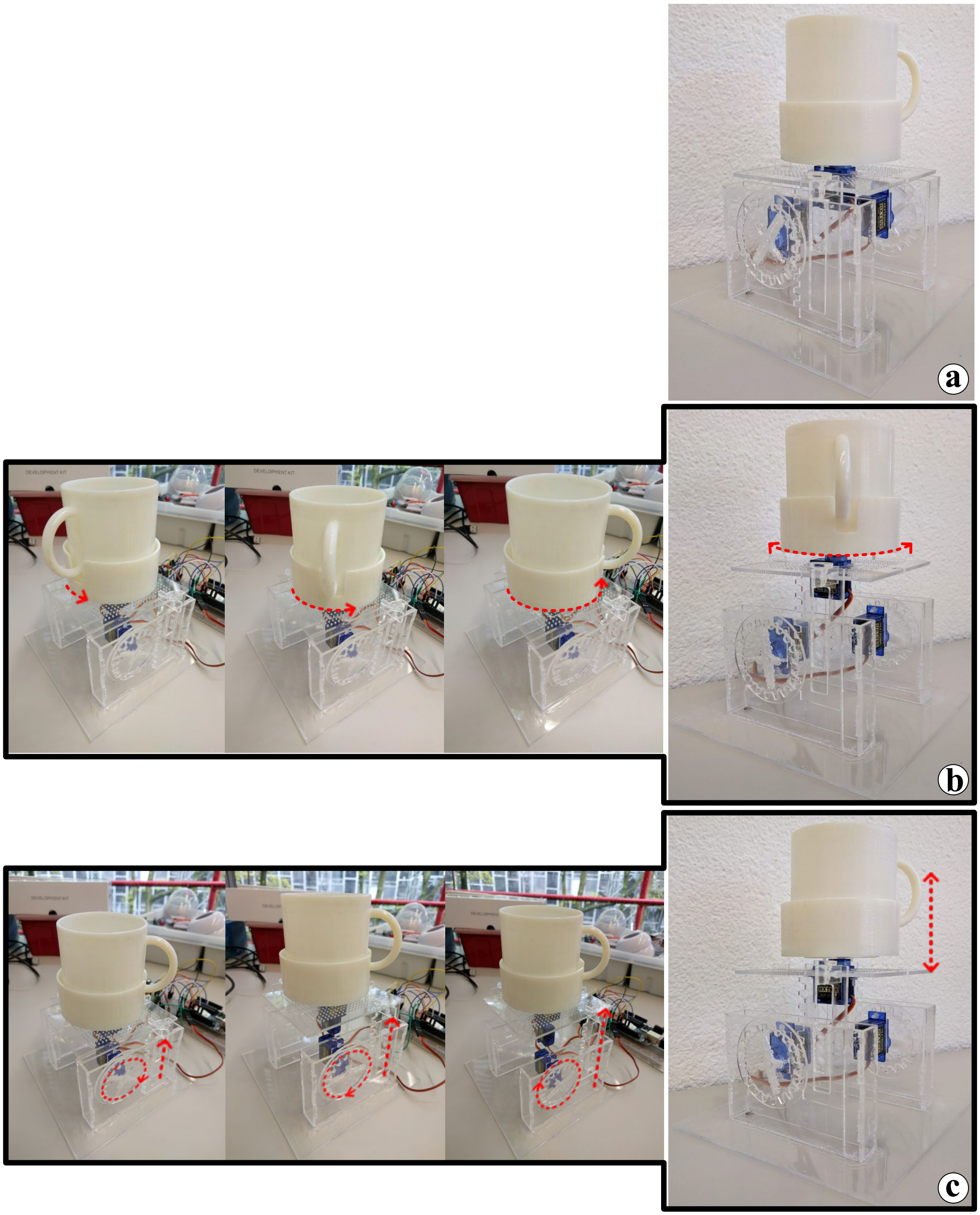


Figure 13 Intention initiation: (a) An idle proactive coffee cup waits for a potential user, (b) horizontal movements when active user's visual attention is present, and (c) vertical rotation movements after continuous visual attention from the user.

3.2. Proof-of-Concept Implementation and Example Scenario

Like previous work, an everyday object for proof-of-concept prototype implementation is chosen (e.g., Kao and Schmandt, 2014; Angelini *et al.*, 2015; Levillain and Zibetti, 2017; Barreiros *et al.*, 2017). As productivity methods suggest, taking a proper break periodically, such as 52-minute work, 17-minute break (Park *et al.*, 2019), is good for the worker's health and productivity. Therefore, a proactive coffee cup that can invite the workers to take a short break and have a coffee is implemented.

Figure 14 shows a proactive coffee cup that invites a user to take a short break. Figure 26 shows a proactive coffee cup that terminates the interaction when the user ignores the invitation. As shown in Figure 14a, the coffee cup is in idle condition waiting for a potential user. The coffee cup handle is facing away from the user during the idle condition. After some time, the coffee cup identified a potential user looking at it. The coffee cup horizontally moves to express its awareness towards the potential user's visual attention, as shown in Figure 14b. The coffee cup vertically orientates its body and stops when the handle faces the user as a friendly invitation to invite the user to take a short break, as shown in Figure 14c and Figure 14d. The user accepts the invitation by lifting the cup and take a sip of coffee, as shown in Figure 14e and Figure 14f. Notably, when the coffee cup finds the user ignore the invitation by looking away, as shown in Figure 15a and Figure 15b, it orientates itself and switches to idle condition, as shown in Figure 15c and Figure 15e.

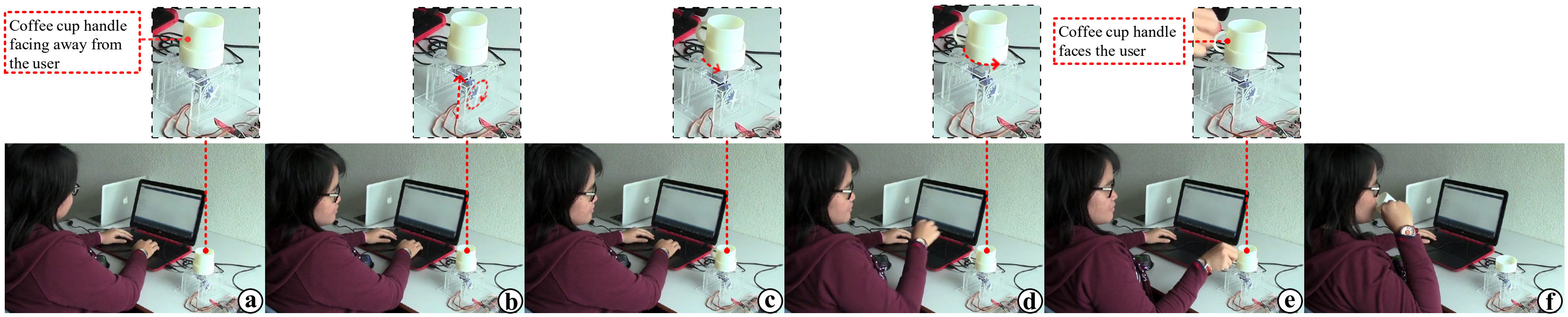


Figure 14 User accepting an invitation: (a) a user busy with work, (b) looks at the coffee cup. The coffee cup horizontally moves to express its awareness, (c) vertically orientates its body, and (d) stops when the handle faces the user as an invitation to take a short break. (e) The user accepts the invitation by lifting the cup and (f) take a sip of coffee.

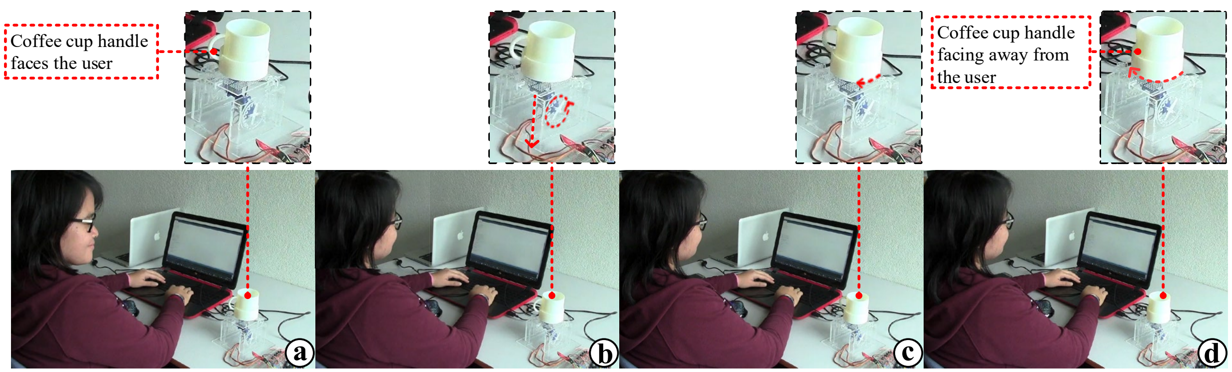


Figure 15 User ignoring an invitation: (a)Coffee cup invite a user to take a short break, (b) the user looks away to ignore the coffee cup invitation (c) the coffee cup orientates itself, and (d) goes into the idle condition.

3.3. User-Experience Experimental Design

3.3.1. Experimental setup

As shown in Figure 16, an experimental setup was developed to study the user's experience interacting with the proactive coffee cup. The eye tracker used in the experiment was positioned in parallel with the sitting participant at a distance of 60 cm to ensure a reliable visual presence with attention is detected. The distance of 60 cm to provide reliable eye-gaze detection was recommended and referred from The Eye Tribe developer site (Johansen *et al.*, 2013). Two identical coffee cups, a) non-proactive coffee cup was placed at grid B1 and b) proactive coffee cup was placed on grid B2. The non-proactive coffee cup was not integrated with an embedded system, and the proactive coffee cup was integrated with an embedded system. According to Auvray *et al.*'s perceptual crossing paradigm (20019;2019), the interaction between two entities must be bidirectional and proactive to experience perceptual crossing. Therefore, to implement the perceptual crossing paradigm into the human-object interaction design, a person must realize that they were experiencing bidirectional and proactive interaction. Hence, this experimental setup's motivation was to investigate the participants' alertness to discover the bidirectional and proactive interaction with the proactive coffee cup when they looked at it compared to when they shifted their visual attention to the non-proactive coffee cup. Laptop and electronic components used in the experiment were covered to avoid distracting the participants' attention during interaction with the coffee cups. The prototype's mechanical design was protected using a small box to hide the mechanical structure and reduce the servo motor noise that might distract the participants during the interaction.

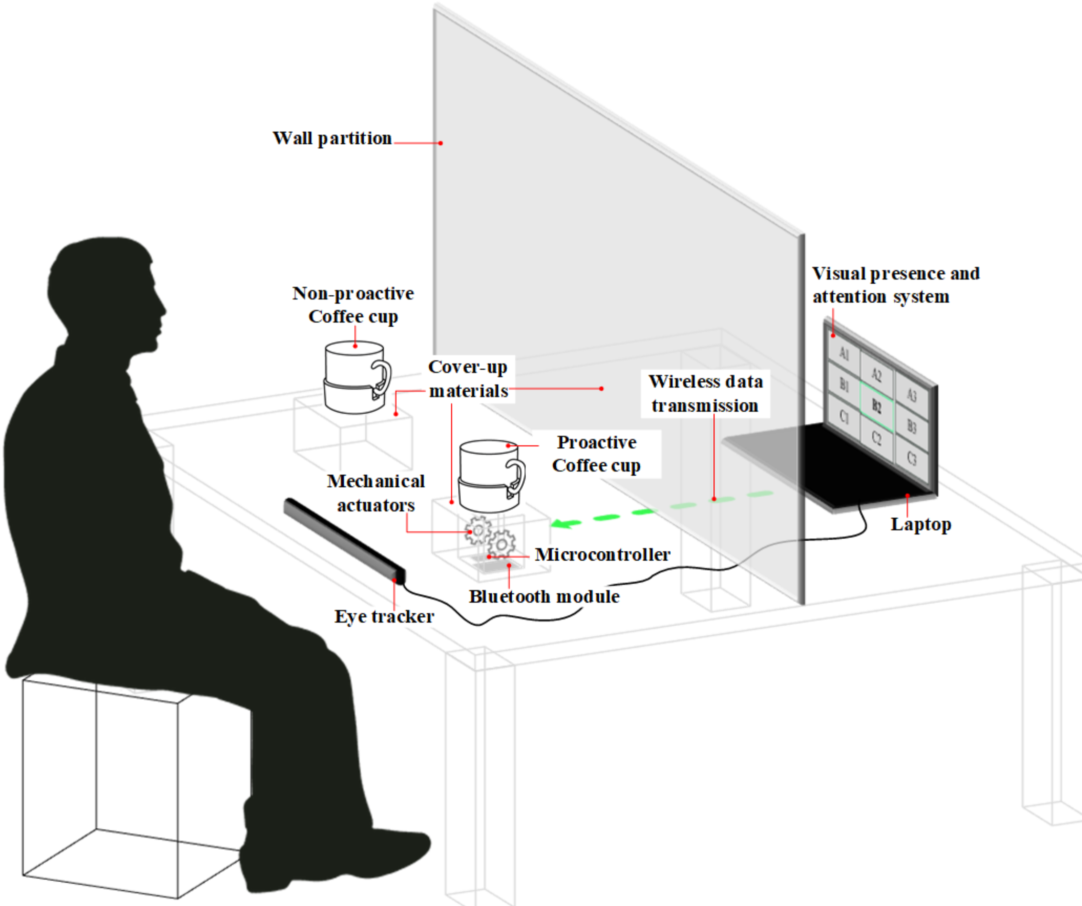


Figure 16 Overview of the experimental setup. a) Non-proactive coffee cup and b) proactive coffee cup is placed on the table to evaluate the interaction between participants and the object prototype.

3.3.2. Participants

A between-subjects experiment was designed with 30 subjects (17 males, 13 females, age range: 25 to 38). All the recruited participants had no prior experience interacting with an eye tracker device. The participants were evenly separated into two groups of 15 participants, participants with no prior knowledge ($P_{\overline{pk}}$) and participants with prior knowledge (P_{pk}). $P_{\overline{pk}}$ participants had no prior knowledge regarding the prototype's capability of expressing its intention to interact when they looked at it, whereas P_{pk} participants had prior knowledge. All participants signed written, informed consent to participate in the experiment.

3.3.3. Procedures

As mentioned in section 3.2.2, participants were divided into two groups: $P_{\overline{pk}}$ and P_{pk}. Each participant from both groups conducted an individual eye tracker calibration process. The eye tracker calibration process was required to estimate the user's visual presence with attention accurately. Each participant was asked to minimize their head movements during active interaction with the coffee cups because the eye tracker had a very low tolerance towards the participant's head movement. Next, each participant of $P_{\overline{pk}}$ and P_{pk} was required to observe and explore the a) non-proactive coffee cup and b) proactive coffee cup for five minutes. A brief instruction was given: *"while sitting down, you are free to observe, explore and interact with the coffee cups placed in front of you."* After five minutes of observing, exploring, and interacting with the proactive coffee cup, the participants must fill out a set of questionnaires and participate in a post-evaluation interview. The post-evaluation interview was conducted to understand the participant's perspective towards the proactive coffee cup.

3.3.1. Measurement

The User Experience Questionnaire (UEQ) (Laugwitz *et al.*, 2008) was used to measure the participant's experiences interacting with the proactive coffee cup (see *Appendix A*). The UEQ questionnaire consists of 26 items divided into six individual scales, which were 1) *Attractiveness* indicated the overall impression of the proactive coffee cup, 2) *Perspicuity* measured the difficulty of participants to get familiar and learned how to interact with the proactive coffee cup, 3) *Efficiency* was the ability of participants to realize the interaction modality required to interact with the proactive coffee cup without unnecessary effort, 4) *Dependability* indicated if participants felt in control while interacted with the proactive coffee cup 5) *Stimulation* shows participants' excitement and motivation to use the product and 6) *Novelty* indicated whether the proactive coffee cup was innovative, creative and able to catch the participant's interest. Participants must express their experience with the proactive coffee cup by marking each item of the UEQ questionnaire with a seven-point Likert scale and briefly describing their experience during interaction with the coffee cup.

3.4. Results

The user experience results are presented in Figure 17, Figure 18, Table 1, and Figure 19. Figure 17 and Figure 18 show the diverging stacked bar chart that summarized the participants' responses while interacted with the proactive coffee cup for $P_{\overline{pk}}$ and P_{pk} groups. The participants' feedbacks were further analyzed as tabulated in Table 1. Firstly, the UEQ questionnaire used the scale from -3 (negative) to 3 (positive), where the mean values (M) of the participants' feedback were achieved and categorized into neutral evaluation between values -0.8 to 0.8, positive evaluation value greater than 0.8, and a negative evaluation value of less than -0.8. Next, the variance values were required to obtain standard deviation values to analyze the participants' feedback experience reliability while interacting with the proactive coffee cup. Based on the standard deviation definition, if an item shows a significant deviation between the same scales, it means the participants have misinterpreted the evaluating item (Hinderks *et al.*, 2018). Hence, to better understand the participants' responses, all the measured items tabulated in Table 1 have been summarized in Figure 19. The information presented in Figure 19 is representing the groups of $P_{\overline{pk}}$ and P_{pk} which has been group into six scales. It can be interpreted the $P_{\overline{pk}}$ group with no prior knowledge seems not to understand how to interact with the proactive coffee cup.

The next section will present the quantitative and qualitative findings to elaborate on the results shown in Figure 26, Table 1, and Figure 27. The quantitative findings will discuss the neutral evaluation, especially from the group with no prior knowledge, $P_{\overline{pk}}$. Lastly, an overall summary of both groups' obtained results will be summarized at the end of the quantitative section.

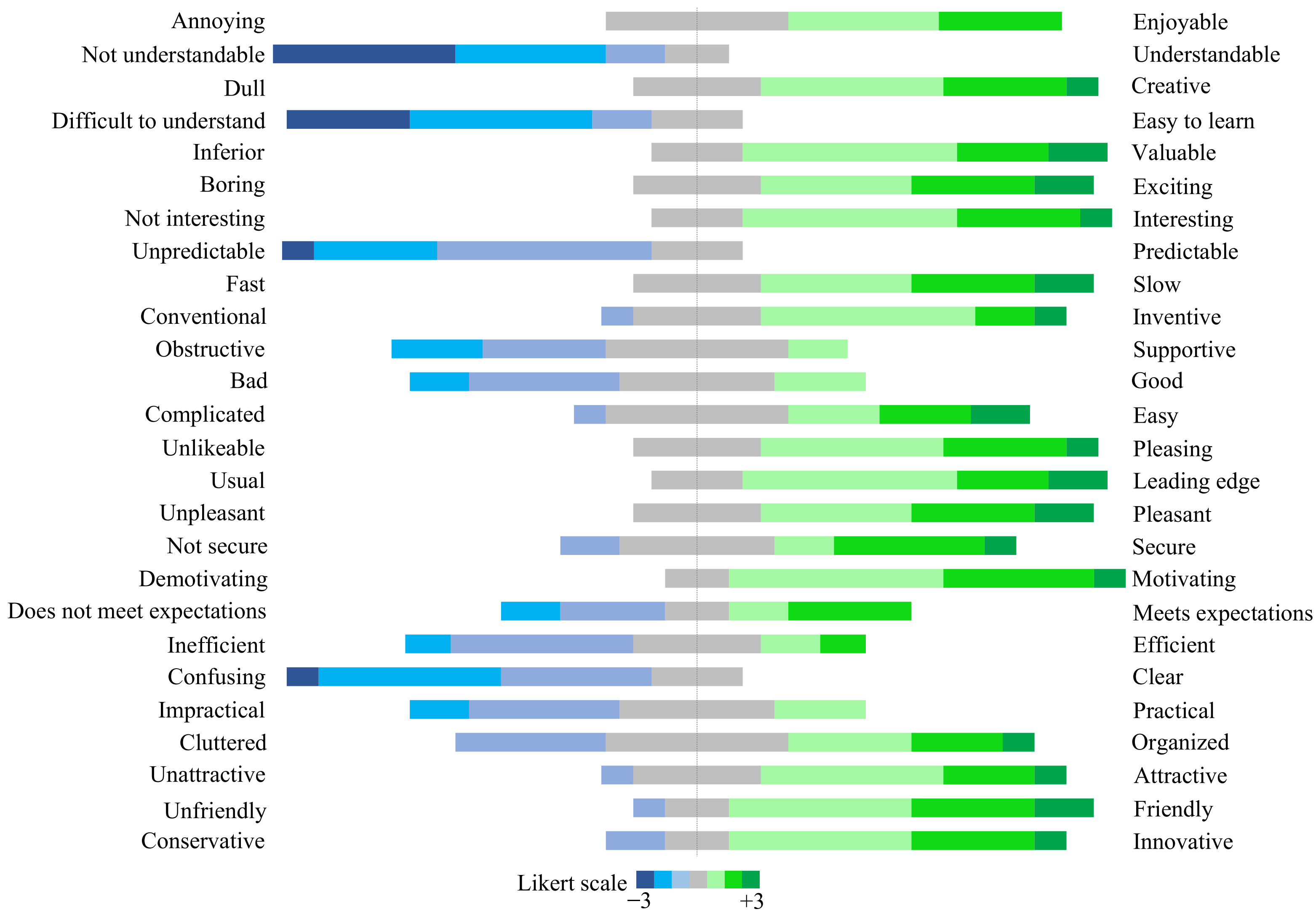


Figure 17 Diverging stacked bar chart of participants' responses to the UEQ seven-point Likert scale for each item from the $P_{\overline{pk}}$ group (without prior knowledge).

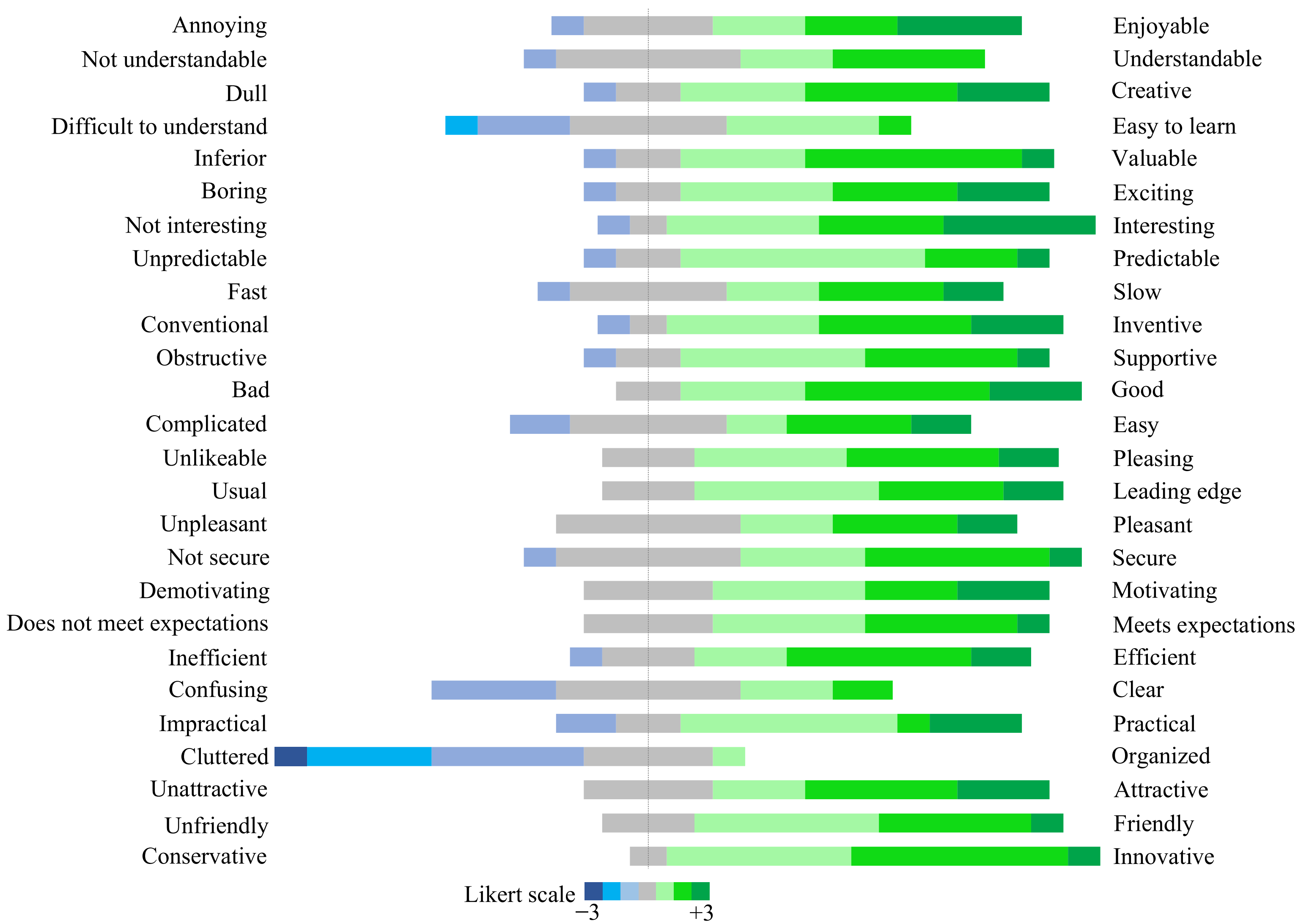


Figure 18 Diverging stacked bar chart of participants' responses to the UEQ seven-point Likert scale for each item from the P_{pk} group (with prior knowledge).

Table 1 Results from the user experience questionnaire (UEQ) (n = 15). Shaded rows indicate collected results from participants with no prior knowledge, $P_{\overline{pk}}$. Unshaded rows indicate collected results from participants with prior knowledge, P_{pk}. Significant results (p<.05) are highlighted and marked by a red asterisk (*). The arrows indicate whether the collected value per item is positive (↑), negative (↓), or neutral (→).

Items	Mean, M	Variance	Standard Deviation	Left	Right	Scales
1	↑ 0.9	0.7	0.8	annoying	enjoyable	*Attractiveness*
	↑ 1.4	1.5	1.2			
2	↓ **−1.5**	**0.8**	**0.9**	**not understandable**	**understandable**	***Perspicuity****
	↑ **0.8**	**1.0**	**1.0**			
3	↑ 1.6	0.4	0.6	creative	dull	*Novelty*
	↑ 1.9	0.5	0.7			
4	↓ **−1.7**	**0.8**	**0.9**	**easy to learn**	**difficult to learn**	***Perspicuity****
	→ **0.2**	**0.9**	**0.9**			
5	↑ 1.1	0.3	0.5	valuable	inferior	*Stimulation*
	↑ 1.5	0.7	0.8			
6	↑ 1.2	0.6	0.8	boring	exciting	*Stimulation*
	↑ 1.5	1.0	1.0			
7	↑ 1.2	0.5	0.7	not interesting	interesting	*Stimulation*
	↑ 1.9	1.0	1.0			
8	↓ **−1.2**	**0.7**	**0.9**	**unpredictable**	**predictable**	***Dependability****
	↑ **1.0**	**0.4**	**0.7**			
9	↑ 0.9	0.7	0.8	fast	slow	*Efficiency*
	↑ 1.1	1.3	1.1			
10	↑ 1.0	0.3	0.5	inventive	conventional	*Novelty*
	↑ 1.7	0.8	0.9			
11	→ **−0.2**	**0.3**	**0.6**	**obstructive**	**supportive**	***Dependability****
	↑ **1.2**	**0.5**	**0.7**			
12	→ 0.6	0.5	0.7	good	bad	*Attractiveness*
	↑ 1.7	1.0	1.0			
13	→ **−0.3**	**0.8**	**0.9**	**complicated**	**easy**	***Perspicuity****
	↑ **0.9**	**1.8**	**1.3**			
14	↑ 0.9	0.6	0.8	unlikable	pleasing	*Attractiveness*
	↑ 1.2	0.6	0.8			
15	↑ 1.0	0.4	0.7	usual	leading edge	*Novelty*
	↑ 1.1	0.5	0.7			
16	↑ 0.9	0.6	0.8	unpleasant	pleasant	*Attractiveness*
	↑ 1.1	1.3	1.1			
17	↑ 0.8	1.3	1.1	secure	not secure	*Dependability*
	↑ 1.2	1.2	1.1			

Items	Mean, M	Variance	Standard Deviation	Left	Right	Scales
18	↑ 1.2	0.5	0.7	motivating	demotivating	*Stimulation*
	↑ 1.4	1.1	1.1			
19	→ **−0.7**	**0.8**	**0.9**	**meets expectations**	**does not meet expectations**	***Dependability****
	↑ **1.2**	**0.9**	**0.9**			
20	→ **−0.7**	**0.2**	**0.5**	**inefficient**	**efficient**	***Efficiency****
	↑ **1.3**	**1.2**	**1.1**			
21	↓ **−1.7**	**0.4**	**0.6**	**clear**	**confusing**	***Perspicuity****
	→ **0.2**	**1.0**	**1.0**			
22	→ **−0.3**	**0.5**	**0.7**	**impractical**	**practical**	***Efficiency****
	↑ **1.1**	**1.6**	**1.3**			
23	→ **−0.1**	**0.5**	**0.7**	**organized**	**cluttered**	***Efficiency****
	↑ **0.8**	**0.7**	**0.9**			
24	↑ 1.1	0.8	0.9	attractive	unattractive	*Attractiveness*
	↑ 1.1	1.3	1.1			
25	↑ 1.2	0.5	0.7	friendly	unfriendly	*Attractiveness*
	↑ 1.3	0.8	0.9			
26	↑ 1.3	0.6	0.8	conservative	innovative	*Novelty*
	↑ 1.5	0.6	0.7			

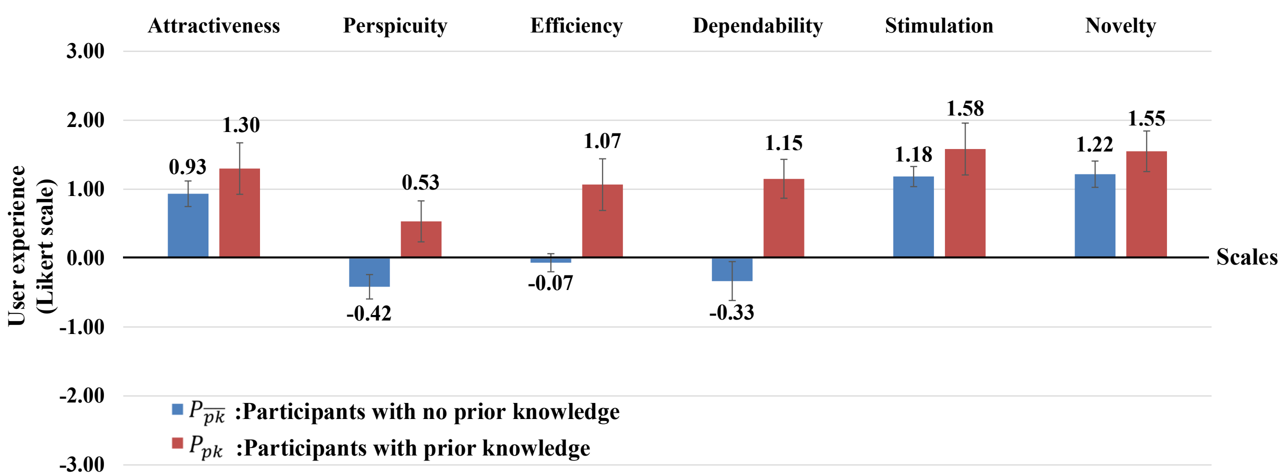


Figure 19 Overall results for UEQ items, per scale.

3.4.1. Quantitative results

Perspicuity: Is it easy to engage with the proactive coffee cup?

As shown in Figure 19, the $P_{\overline{pk}}$ group gave *Perspicuity* the most negatively evaluated scale. As shown in Table 1, item 4:M=−1.7(↓), item 21:M=−1.7(↓), and item 2:M=−1.5(↓) has been poorly evaluated due to inability to interact with the proactive coffee cup, which means the participants had difficulty realizing, feel confused, and not able to understand the interaction. To support the participants' feedback experience, the collected data during the participant's engagement with the proactive coffee cup shows 86% of the participants from the $P_{\overline{pk}}$ group rated item 4 and item 2 as difficult to realize and not able to understand the interaction, as shown in Figure 17. Whereas for item 21, 80% of the $P_{\overline{pk}}$ group participants felt confused while interacting with the proactive coffee cup as shown in Figure 17. This can be concluded that engaging and interacting participants from the $P_{\overline{pk}}$ group felt clueless on the behavior shown by the proactive coffee cup. Meanwhile, item 13 falls under the neutral evaluation based on the calculated mean value shown in Table 1. The result suggests that the participants' experience interacting with the prototype was neither complicated nor easy.

According to Figure 19, the P_{pk} group gave *Perspicuity* the lowest positively evaluated scale. As shown in Table 1, item 4:M=0.2(→) and item 21:M=0.2(→) have been neutrally evaluated due to participants' inability to understand the purpose of observing the proactive coffee cup. To support the participants' feedback experience, the collected data during the participant's engagement with the proactive coffee cup shows 47% of the participants from the P_{pk} group rated item 4 and item 21 as difficult to learn and confusing, as shown in Figure 18. It can be interpreted even though P_{pk} group had prior knowledge; the participants yet felt complicated interacting with the proactive coffee cup.

Dependability: Are the participants feel in control during interaction with the proactive coffee cup?

As shown in Figure 19, the $P_{\overline{pk}}$ group gave *Dependability* negative evaluated scale. The *Dependability* scale has been negatively evaluated because item 8 (predictable/unpredictable) is rated poorly due to the proactive coffee cup's unpredictability behavior. To support the participants' feedback experience, the collected data during the participant's engagement with the proactive coffee cup shows 81% of the participants from the $P_{\overline{pk}}$ group rated item 8 as unpredictable as shown in Figure 17. The probability of item 8 is negatively evaluated because of the proactive coffee cup's unpredictable behavior. To further support the Dependability scale's negative evaluation and referring to Table 1, item 19: M=−0.7(→) and item 11:M=−0.2(→) has been neutrally evaluated, which insignificantly contributed towards negative evaluation. Overall, the summary of the participants' feedback experience identified that participants in $P_{\overline{pk}}$ were not able to be in control while engaging with the proactive coffee cup.

According to Figure 19, the P_{pk} group gave *Dependability* positive evaluated scale. As shown in Table 1, item 11:M=1.2(↑), item 19:M=1.2(↑), and item 8: M=1.0(↑) has been positively evaluated. To support the participants' feedback experience, the collected data during the participant's engagement with the proactive coffee cup shows 80% of the participants from the P_{pk} group rated item 11 and item 19 as supportive and meeting their predicted expectations, whereas 73% of the participants from the P_{pk} group rated item 8 as predictable, as shown in Figure 18. Overall, it can be summarized the participants felt they could control and predict the behavior while interacting with the proactive coffee cup.

Efficiency: Is unnecessary effort required to engage with the proactive coffee cup?

As shown in Figure 19, the $P_{\overline{pk}}$ group gave *Efficiency* almost neutral evaluated scale. Item 20:M=−0.7(→), item 22:M=−0.3(→) and item 23:M=−0.1(→) has been neutrally evaluated due to participants in $P_{\overline{pk}}$ were undecided whether the amount of effort they have to put in to engage with the proactive coffee cup was neither inefficient nor efficient and neither impractical nor practical. To support the participants' feedback experience, the collected data during the participant's engagement with the proactive coffee cup shows 50% of the participants from the $P_{\overline{pk}}$ group rated item 20 as inefficient, 46% rated item 22 as impractical, and 33% rated item 23 as cluttered, as shown in Figure 17. Nonetheless, by referring to Table 1, item 9 has been positively evaluated with a mean value, M=0.9(↑). By referring to item 9 in Figure 17, 83% of the participants in $P_{\overline{pk}}$ considered the reaction of the proactive coffee cup's behavior was fast. Overall, the summary of the participants' feedback experience shows that participants in $P_{\overline{pk}}$ were not able to identify the strategies required to engage with the proactive coffee cup.

According to Figure 19, the P_{pk} group gave *Efficiency* positive evaluated scale. As shown in Table 1, item 20:M=1.3(↑), item 9:M=1.1(↑), item 22: M=1.1(↑), and item 23:M=0.8(↑) has been positively evaluated. As shown in Figure 17, to support the participants' feedback experience from the P_{pk} group, the collected data during the participant's engagement with the proactive coffee cup shows 73% of the participants rated item 20 as efficient, 60% rated item 9 as fast, 74% rated item 22 as practical, and 54% rated item 23 as organized, as shown in Figure 18. Overall, it can be summarized the participants felt through visual attention, the interaction with the proactive coffee cup was quick and effective.

Overall impression

Analyzing the results according to the scales which has been presented in Figure 19, the participants in $P_{\overline{pk}}$ group showed *Attractiveness* (M=0.93), *Stimulation* (M=1.18), and *Novelty* (M=1.22) scales which indicates their satisfaction while positively evaluating the proactive coffee cup. However, the *Perspicuity* (M=−0.42), *Efficiency* (M=−0.07), and *Dependability* (M=−0.33) scales are rated as neutral. The participants evaluated these scales as neutral because they had difficulty learning and required some time and self-effort to understand the input modality while interacting with the proactive coffee cup. Whereas, the P_{pk} group participants showed *Attractiveness* (M=1.3), *Efficiency* (M=1.07), *Dependability* (M=1.15), *Stimulation* (M=1.58), and *Novelty* (M=1.55) scales were positively evaluated, which indicates that the proactive coffee cup is well-received by the participants. Despite that, *Perspicuity* (0.53) scale is neutrally evaluated. Nonetheless, even though P_{pk} group participants had prior knowledge about it, yet the participants felt uneasiness to understand the behavior shown by the proactive coffee cup.

3.4.2. Qualitative results

Impression during interaction with the proactive coffee cup

Impression during the interaction has been qualitatively captured for the group of $P_{\overline{pk}}$ and P_{pk}. Thirteen participants from the $P_{\overline{pk}}$ group found it difficult to understand the responsiveness of the proactive coffee cup. Therefore, several reflections from the participants are presented in the following:

"I am clueless regarding what triggered the coffee cup to turn itself." (P1)
"I am confused why one of the coffee cups turned its body while the other one was not moving." (P3)
"One of the coffee cups turned itself while the other one was not moving, and it confused me." (P7)
"I did not understand and confused why one of the coffee cups moved, and the coffee cup next to it maintained its position." (P9)

"I mostly ignored the non-moving coffee cup and focused on the interactive one but failed to understand what influenced the coffee cup's reaction." (P4)
"I struggled to find out why at one time the coffee cup handle turned towards me, and some other time it turned away from me." (P5)
"The coffee cup handle turned toward me and then turned the other way around. It was difficult for me to comprehend why it behaved like that." (P8)
"I understand that the coffee cup turned its handle toward me, but I am baffled to find the reason behind it." (P14)

While several participants had difficulties in understanding the method of responsiveness of the proactive coffee cup, two participants from the $P_{\overline{pk}}$ group mentioned that the proactive coffee cup could express its behavior when the participants engaged through their visual attention. Therefore, the reflections from the two participants are presented in the following:

"I observed both coffee cups and realized one was static, and the other one turned itself toward me, and it happened when I looked at it." (P2)
"Through several investigations, I concluded that the device (eye-tracker) detected my eyes and therefore I knew the coffee cup moved because I looked at it." (P15)

Seven participants from the P_{pk} managed to understand the proactive coffee cup's capability and appreciated its responsiveness due to the eye tracker's initial introduction to them before the experiment was conducted. Therefore, several reflections from the participants are presented in the following:

"I did not realize that the coffee cup was reactive because of my visual attention if the experimenter did not mention it to me first." (P2)
"I never engaged with an eye tracker before. Luckily, the experimenter was willing to explain about it at the beginning of the experiment." (P6)
"It was hard to comprehend the situation if the experimenter was not there with me." (P7)
"This interaction was new to me. I would be struggling without the help from the experimenter." (P10)
"It was very thoughtful of the experimenter to provide me with useful information, especially about the eye tracker." (P15)

Even though the participants from the P_{pk} group has prior knowledge of the proactive coffee cup's capability, yet four participants expressed the difficulty of continuously staying engaged during interaction with the proactive coffee cup. Therefore, the reflections from the four participants are presented in the following:

"It was hard to focus on the coffee cup when it started to move." (P1)
"I did not know where I should point my eyes while the coffee cup turned its handle towards me." (P3)
"I lost my visual attention the instanced it turned to me." (P8)
"Maintained my eyes at the coffee cup while it moved was very inconvenienced for me." (P12)

Whereas three participants from the P_{pk} group failed to acknowledge the proactive coffee cup's responsiveness when the participants engaged through their visual attention. Therefore, the reflections from the three participants are presented in the following:

"I was focused on the motion, and it made me forgot that it moved because I stared at it." (P4)
"I understand the object's intention, but I am not convinced that it moved because of my eyes." (P5)
"There was no indication that my attention influenced the coffee cup's motion." (P8)

Establishing engagement with the proactive coffee cup

Establishing engagement for interaction with the proactive coffee cup has been qualitatively captured for the group of $P_{\overline{pk}}$ and P_{pk}. Six participants from the $P_{\overline{pk}}$ group tried and investigated the input modality required for the proactive coffee cup to respond. Therefore, six reflections from the participants are presented in the following:

"I waved at the coffee cup out of my curiosity, but it did not make sensed at all." (P4)
"I snapped my finger to get the cup to move, but nothing happened." (P5)
"I thought there was a sensor on my chair to detect my presence (when sitting down), but my assumption was wrong." (P6)
"When I sat in front of the coffee cup, it suddenly rotated. I assumed the chair had a sensor underneath it to detect my presence, but there was nothing attached to that chair." (P13)
"The device (eye-tracker) seemed detached to the object, so it had no relation with the coffee cup responsiveness." (P11)
"I thought the coffee cup detected my presence, but there was no clear indication of it." (P15)

Five participants from the $P_{\overline{pk}}$ group speculated the proactive coffee cup's physicality or the eye tracker to understand the interaction channel to engage with the proactive coffee cup. Therefore, several reflections from the participants are presented in the following:

"Perhaps touched the coffee cup handle was the possible input modality." (P2)
"The coffee cup turned and stopped when the handle faced toward me. I thought that if I wanted to interact with the coffee cup, I needed to hold the coffee cup by the handle" (P10)
"It seemed that the device (eye-tracker) supposed to detect my hand movement." (P7)
"The device (eye-tracker) detected my presence. Therefore the coffee cup turned itself toward me." (P9)

Seven participants from the P_{pk} group suggested adopting explicit design feedback would further improve engagement with the proactive coffee cup. Therefore, seven reflections from the participants are presented in the following:

"The compact plastic bar (eye-tracker) did not reflect its functionality. Using a camera will be more straightforward." (P2)
"The interaction was uncleared to me. Added visual display onto the object to confirm the engagement would be helpful." (P3)
"Embedded the eye tracker with a pan/tilted face tracking mechanism to make the user felt more connected with the object." (P5)
"I did not know where to point my eyes. It is easier for me to focus and stayed engaged if there was a visual point on it to hold my eyes." (P6)
"It is difficult to reassure that I engaged with the object. Illuminated the area when I looked at it would make a difference." (P8)
"An eye icon pasted onto the object might alert the user that it was an object that required a person to look at it." (P13)
"Covered the outer layer of the cup with LED strip and turned it on when the user looked at it." (P14)

The impression of the proactive coffee cup's abstract motion for expressing an intention

The impression of proactive coffee cup's abstract motion for expressing intention has been qualitatively captured for the group of $P_{\overline{pk}}$ and P_{pk}. Five participants explained the proactive coffee cup's abstract motion attracted their attention and engagement. Therefore, five reflections from the participants are presented in the following:

"When the coffee cup handle turned towards me, I felt that it wanted me to grasp it." (P1)
"The moving coffee cup was competing with the other coffee cup to get my attention." (P5)
"The coffee cup's motion was trying to lure me into using it." (P7)
"I moved away from the installation (experiment setup) and saw the coffee cup turned its handle to the other side. That was how I knew the coffee cup was expressing interest toward me." (P8)
"I ignored the motionless coffee cup and paid more attention to the interactive one." (P15)

Next, five participants from the $P_{\overline{pk}}$ group defined the abstract motion by the proactive coffee cup as a behavior to express its intention to interact. Therefore, five reflections from the participants are presented in the following:

"When the coffee cup turned toward me, I realized it was an invitation to interact." (P2)
"The coffee cup expressed its usefulness through motion." (P6)
"The direction of the coffee cup handle was a clue that it wanted to interact with me." (P10)
"The coffee cup's motion demonstrated its capability of interactivity." (P12)
"The coffee cup tried to evoke positive emotion out of me by playfully reoriented itself back and forth." (P14)

Even though several participants understand the abstract motion attracted the participant to stay engaged with the proactive coffee cup, yet five participants from the $P_{\overline{pk}}$ group mentioned that the proactive coffee cup's abstract motion did not catch their attention. Therefore, four reflections from the participants are presented in the following:

"I saw the coffee cup turned toward me, but I doubted it was because of me" (P4)
"The coffee cup turned on its own will without required any action from me" (P9)
"The coffee cup was programmed to behave like that" (P3)
"The coffee cup's motion was based on its internal state" (P11)
"The coffee cup automatically turned because it wanted to interact with me" (P13)

Nonetheless, five participants from the P_{pk} group commented that the abstract motion distracted them from understanding the coffee cup's proactive behavior. Therefore, five reflections from the participants are presented in the following:

"The coffee cup's motion diverted my focus of attention." (P4)

"I was holding my eyes at one point on the coffee cup, but it annoyed me when the point started to move away." (P7)
"It was difficult to keep my eyes on it while tried to understand the motion." (P8)
"It took some times for me to familiarize the meaning behind the motion while holding my eyes on the coffee cup." (P11)
"I briefly distracted when the coffee cup oriented itself." (P13)

Also, seven participants from the P_{pk} group was uncertain about their visual attention that could initiate an engagement with the proactive coffee cup. Therefore, six reflections from the participants are presented in the following:

"I looked at both coffee cups, and it does nothing. It was very confusing." (P3)
"There was no indication of how long should I looked at the coffee cup before it started to react to me." (P7)
"I unsure if the coffee cup's motion was because of my visual attention." (P8)
"I failed to grasp that the coffee cup's behavior was because of my eyes." (P5)
"It was uncleared whether the coffee cup's motion was a reaction because I stared at it." (P11)
"The coffee cup unexpectedly turned before I could comprehend the situation." (P14)
"The coffee cup's motion was noticeable, but I am curious if the reaction was because I looked at it." (P15)

Consequently, eight participants from the P_{pk} group described the coffee cup's proactive behavior as engaging. Therefore, six reflections from the participants are presented in the following:

"Such an effort coming from the coffee cup was very unusual, but it motivated me to engage with it." (P1)
"I appreciated the coffee cup that was able to show its usefulness." (P2)
"Although the coffee cup's action was tough to understand, I acknowledged and appreciated the coffee cup's intention to interact with me." (P4)
"The interactive coffee cup offered a unique experience that made me more aware of its presence." (P6)
"A simple object enhanced with the capability to approach people would be practical, especially when it tried to remind us of doing something." (P9)
"I felt more involved with the interactive coffee cup when it started to interact with me." (P10)
"The effort that the coffee cup took to engage with me was very much appreciated." (P12)

"Although it took some time to get used to it, I favored the idea of an object that able to initiate the interaction." (P13)

3.5. Discussion

This section discusses the overall findings for the groups of $P_{\overline{pk}}$ (without prior knowledge) and P_{pk} (with prior knowledge) after observing, exploring, and interacting with the proactive coffee cup. The participants from the group $P_{\overline{pk}}$ found it complicated to engage with the proactive coffee cup when the modality of interaction is unknown. The participants also took several unnecessary efforts to figure out the input modality to interact and engage with the proactive coffee cup. Consequently, the participants in the group $P_{\overline{pk}}$ failed to understand and realize that they were experiencing bidirectional interaction with the proactive coffee cup through their visual attention. Although with prior knowledge, the participants from the group P_{pk} found it difficult to stay engaged during interaction with the proactive coffee cup. Suggestions such as adopting expressive design feedback would improve the engagement with the proactive coffee cup. According to Cha *et al.* (2018) and Rosen *et al.* (2020), implementing expressive and recognizable design feedback helps the user experience bidirectional interaction with the object. The process of initiative-taking from the object-to-user would be easily perceived, too (Amershi *et al.*, 2019).

Therefore, to let the user realize the bidirectional and proactive interaction with an intentional object through visual attention, expressive and recognizable visual feedback is proposed to allow the participant to experience communication using the same perceptual interaction (i.e., visual attention). The implementation of expressive and recognizable visual feedback would help the user spontaneously experience and maintain a bidirectional interaction and engagement with the object by only depending on visual attention.

Apart from that, during interaction with the proactive coffee cup, the participants were also required to move as little as possible due to the eye tracker's low tolerance for head movement. This restriction makes the interaction inefficient as the participants need to avoid making frequent head movements during interaction with the proactive coffee cup. Due to the eye tracker's limitations, a compact camera module with the capability to capture the participant's visual attention is proposed to improve the experience of interacting and engaging with the object.

3.6. Conclusion

The developed prototype that displays proactive behavior through abstract motion has been tested and validated. The results show the prototype's abstract motion to initiate an interaction has successfully attracted the participants to interact and engage with it. However, with visual attention as the interaction method, the participants failed to experience bidirectional interaction with the prototype. Therefore, an expressive and recognizable visual feedback mechanism is proposed to resolve the participants' inability to experience bidirectional interaction with the prototype when they look at it.

Chapter 4.

Study 2: Designing Bidirectional Interaction with Artificial Eyes

In chapter 3, the user-experience experiment results have revealed that with visual attention as the interaction method, the participants did not achieve bidirectional interaction with the prototype. Therefore, an expressive and recognizable visual feedback similar to that of human eye contact (i.e., artificial eyes) is introduced in this chapter[3]. A compact camera module that detects a person's visual presence with attention and a pair of artificial eyes are mounted together on top of an object to allow the person to establish bidirectional interaction with the object. The proposed design is used in a 33-participant user-experience experiment. The results show that with visual attention as the interaction method, the artificial eyes effectively guide the participants to experience and maintain a bidirectional interaction with the object. However, continuous bidirectional interaction with artificial eyes makes the object act reactively. Therefore, it unable to create a proactive expression with the participant. Thus, the participants did not achieve perceptual crossing with the object. Therefore, the artificial eyes are further improved using a conceptual model called the Session Initiation for Proactive Object, as explained in Chapter 5.

4.1. Introduction

Looking into the Auvray *et al.*'s perceptual crossing paradigm, with the interaction method reduced to a bare minimum, both participants interact with the same interaction method by depending on their *touch perception* to communicate with each other. The same active-proactive perceptual activities cross each other, enabling them to coordinate their behavior, thus creating a stable bidirectional interaction. When interacting with a reactive or autonomous entity, the established communication creates a unidirectional-unstable-reactive interaction. The outcome of these perceptual activities suggests that people can still interact with each other even if the communication channel is limited to a single interaction method as long as both are involved in the same perceptual environment and experience bidirectional and proactive interaction.

[3] This chapter is written partly based on Anas et al. (2017).

As presented in Chapter 3, detecting a user's visual attention has proved to be a reliable method of interaction to allow the object to differentiate an intentional user from a potential user. It also allows the object to recognize an entity that has no intention to interact. However, with the absence of the required perceptual quality to engage in the object's perceptual activity, the participants failed to understand the object's intention to interact. Therefore, with the interaction method limited to visual attention, a person can still interact with the object if both are involved in the same perceptual environment. Thus, the object should also have eyes to allow the user to experience interaction in the same perceptual environment.

Direct eye contact signals the intention to engage in an interaction (Jarick and Bencic, 2019). In face-to-face communication, a bidirectional interaction is quickly established when making eye contact (Sharmin and Hoque, 2020). Whereas in a group conversation, to maintain bidirectional interactions, people are more likely to make eye contact with those they are talking to than others (Vertegaal *et al.*, 2001). Apart from that, face-to-face communication also allows a person to negotiate the intention to interact by making or avoiding eye contact (Vertegaal *et al.*, 2003). Therefore, eye contact provides a visible, expressive perceptual quality that can signal the intention to engage in a bidirectional interaction. Making prolonged eye contact can also maintain bidirectional interaction.

With visual attention as the interaction method (i.e., eyes as perceptual quality), the object should also visibly display the same perceptual quality. In this way, both the object and the person can experience active-proactive perceptual activities (i.e., by making eye contact), and therefore bidirectional interaction can be achieved. Hence, the object should also have eyes to show its intention to interact with the user through visual attention. Thus, a pair of artificial eyes and a compact camera module is mounted on an object to enhance the object's ability to signal the intention to engage in a bidirectional interaction by making eye contact with the user of interest, as shown in Figure 20. A coffee machine detects the presence of a potential user. The coffee machine then recognizes the potential user is an intentional user when the user makes eye contact with it. When making eye contact, the user also understands that the coffee machine expresses its intention to engage. The back-and-forth eye-to-eye contact between the coffee machine and the user creates a bidirectional interaction.

Figure 20 Example of a coffee machine mounted with a pair of artificial eyes on top of it.

4.2. Object Prototype Design and Development

In Chapter 3, the designed object prototype used an eye tracker sensor to detect the user's visual presence with attention and the abstract motion to express the object's intention to interact. However, the result of the conducted user-experience experiment revealed that the participants did not achieve bidirectional interaction with the prototype. Therefore, this section further enhances, improves, and explains the proposed design to resolve the participant's inability to experience bidirectional interaction with the prototype. The enhancement and improvement are categorized into three stages, which are 1) visible expressive perceptual quality with artificial eyes, 2) detection of a user's visual presence with attention, and 3) object prototype implementation. The artificial eyes are mounted on top to enhance and improve prototype-user engagement. A compact camera module that detects the user's visual presence with attention enables the prototype to recognize and differentiate an intentional user from a potential user. A coffee machine for proof-of-concept prototype implementation is chosen to allow the object to sense and invite an intentional user for a short coffee break.

4.2.1. Expressive perceptual quality with artificial eyes

Bidirectional interaction is needed for two entities to experience perceptual crossing based on the perceptual crossing paradigm (Auvray *et al.*, 2009; Auvray, 2019). Therefore, a pair of expressive artificial eyes is introduced to improve the experience of engaging in bidirectional interaction. With visual attention as the interaction method, users can experience bidirectional interaction by shifting their visual attention to the artificial eyes to make eye contact. Hence, bidirectional interaction is quickly established (Sharmin and Hoque, 2020). Making eye contact also signals the intention to interact (Jarick and Bencic, 2019). Therefore, mounting a pair of artificial eyes on an object enhances its ability to express its intention to engage in a bidirectional interaction with a user. Some conventional smart things, such as Furby (Hasbro, 2020), Cabbage Patch Kids (Wicked Cool Toys, 2020), and Ulo smart surveillance camera (Muller, 2019), also integrates eyes to improve the object's engagement towards the user.

Figure 21a shows artificial eyes that can enable such a feature. Each eye is displayed on a 1.5-inch full-color Organic Light-Emitting Diode (OLED) and is synchronously animated using a Teensy 3.2 microcontroller. Each eye consists of upper and lower eyelids to give a blinking effect and an iris, a pupil, and sclera to show the gaze's direction (Ruhland *et al.*, 2014). Therefore, the eyes can appear natural by programming them to move their gaze fixation point and blink periodically. The eyes should blink by default since staring without blinking may be perceived as either unfriendly or unresponsive, decreasing the level of believability of the eyes in terms of interaction (Hayashi and Mizuuchi, 2017; Zhang *et al.*, 2017). Continuous eye blinking also enhances intelligence (Michelle *et al.*, 2020) and reduces physiological reactance (Ghazali *et al.*, 2017). Uncanny valley effects (Ciechanowski *et al.*, 2019) were mitigated by changing the graphics from a realistic presentation (Figure 21b) to a cartoonish design (Figure 21c), making them look friendly and approachable. The Mona Lisa effect (Todorovic, 2019) is also reduced using a lens (acrylic cabochons) that enhances and gives a 3D effect to the flat-screen OLED display.

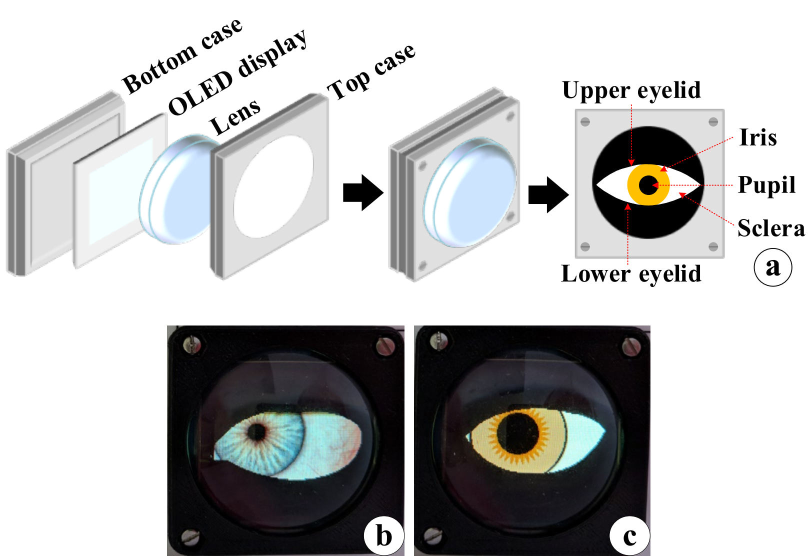


Figure 21 Artificial eyes: (a) overview. (b) realistic eyes. (c) cartoonish eyes.

Figure 22 shows several eye animation effects produced by the artificial eyes that have been embedded, programmed, and displayed via 1.5-inch full-color OLED. The captured eye animation effects shown in Figure 21a and Figure 21b validates the embedded program to produce the eye animation effects for the artificial eyes.

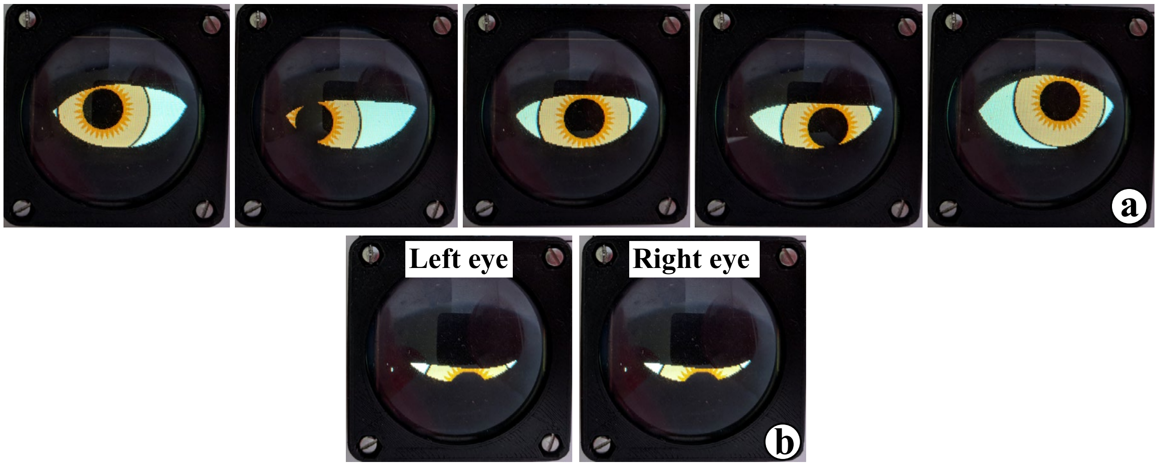


Figure 22 Artificial eyes with (a) gazing effect and (b) blinking effect.

4.2.2. Detection of a user's visual presence with attention

In Chapter 3, the Eye Tribe Tracker (Johansen, 2015) has been used to detect the user's visual presence with attention. However, it required frequent eye calibration for each user and has a low tolerance for head movements. For this reason, the used eye tracker is replaced with a compact camera module that can detect the user's gaze direction, and therefore, eye calibration is not required. This camera module also provides high tolerance for head movements.

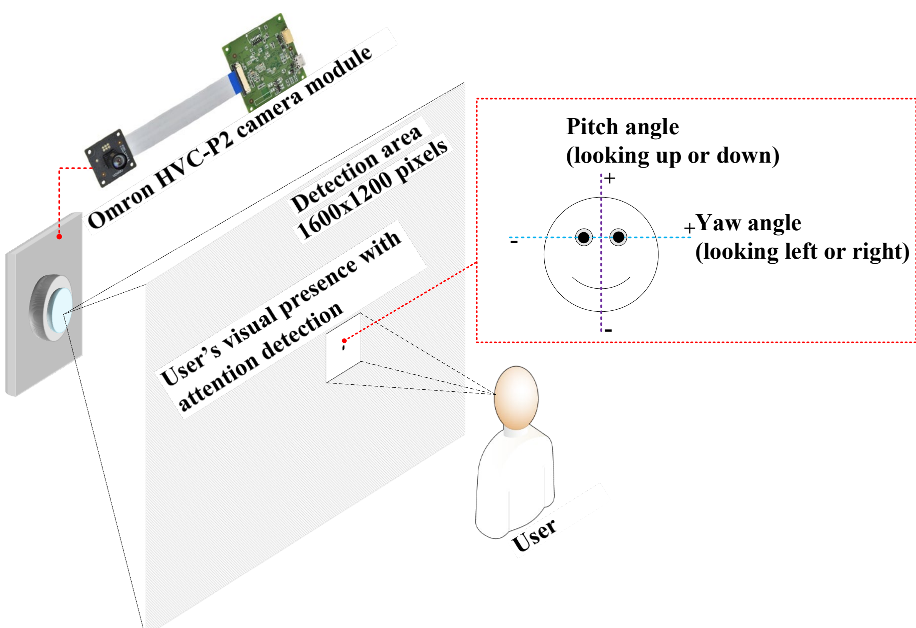


Figure 23 Detection of head movement and gaze direction.

Figure 23 shows a compact Omron HVC-P2 (Mouser Electronics, 2016) camera module with a multi-function image sensor that natively supports the gaze direction and head movements.

Based on the illustration, the yaw and pitch angle are divided into positive and negative. The yaw angle value is positive when the user is looking to the right, whereas when the yaw angle's value is negative, it means the user is currently looking to the left. Meanwhile, when the pitch angle value is positive, the user is now looking up. In contrast, when the pitch angle value is negative, it means the user is looking down. Direct visual attention to the camera module produces zero yaw and pitch angle values. The detection of yaw and pitch angle value allows the camera module to detect the user's gaze direction without calibration.

Although an exposed camera could attract the user's attention to interact, it failed to provide feedback to sustain the engagement. An exposed camera may also raise privacy concerns (Oliveira *et al.*, 2018). The design should consider the camera's placement to make it less intrusive.

4.2.3. Object prototype implementation

Similar to the previous implementation in Chapter 3, an everyday object for proof-of-concept prototype implementation is chosen. As productivity methods suggest, taking regular breaks increases workers' productivity (Park *et al.*, 2019). Therefore, a coffee machine can invite the workers for a short break by offering a cup of coffee.

Figure 24 shows an off-the-shelf coffee machine, disassembled, modified, and mounted with a pair of artificial eyes on top of it. The camera module is also mounted together with the eyes, allowing it to detect direct visual attention when making eye contact with the artificial eyes. Two individual Teensy 3.2 microcontrollers enable smooth and simultaneous information processing between the camera module, artificial eyes, and the coffee machine while the user interacts with the coffee machine. The coffee machine is covered with fabric to hide the buttons, which might tempt the users to push them. The mounted camera module is also well-hidden to avoid distracting the users' attention during interaction with the coffee machine.

Figure 25 explains the operation of the coffee machine. It starts with the artificial eyes in close-idle when there is no user's visual presence with attention detected by the camera module, as shown in Figure 25a. When the artificial eyes are blink-open, as shown in Figure 25b, it means that the camera module detects a user's visual presence with attention. After confirming the user's visual presence with attention and the user having direct eye contact for 3.5 seconds with the artificial eyes, an instruction is sent to the coffee machine to dispense hot coffee, as shown in Figure 25c. A 3.5 seconds initiation processing time between the artificial eyes and the user is required to provide sufficient time to form a bidirectional interaction. Ten users were asked to engage and provide feedback on the 3.5 seconds bidirectional interaction with the coffee machine before implementing the proposed initiation processing time. Based on the ten users' feedback, 3.5 seconds is sufficient to allow the user to feel their bidirectional interaction and engagement with the coffee machine before the coffee machine starts to dispense coffee.

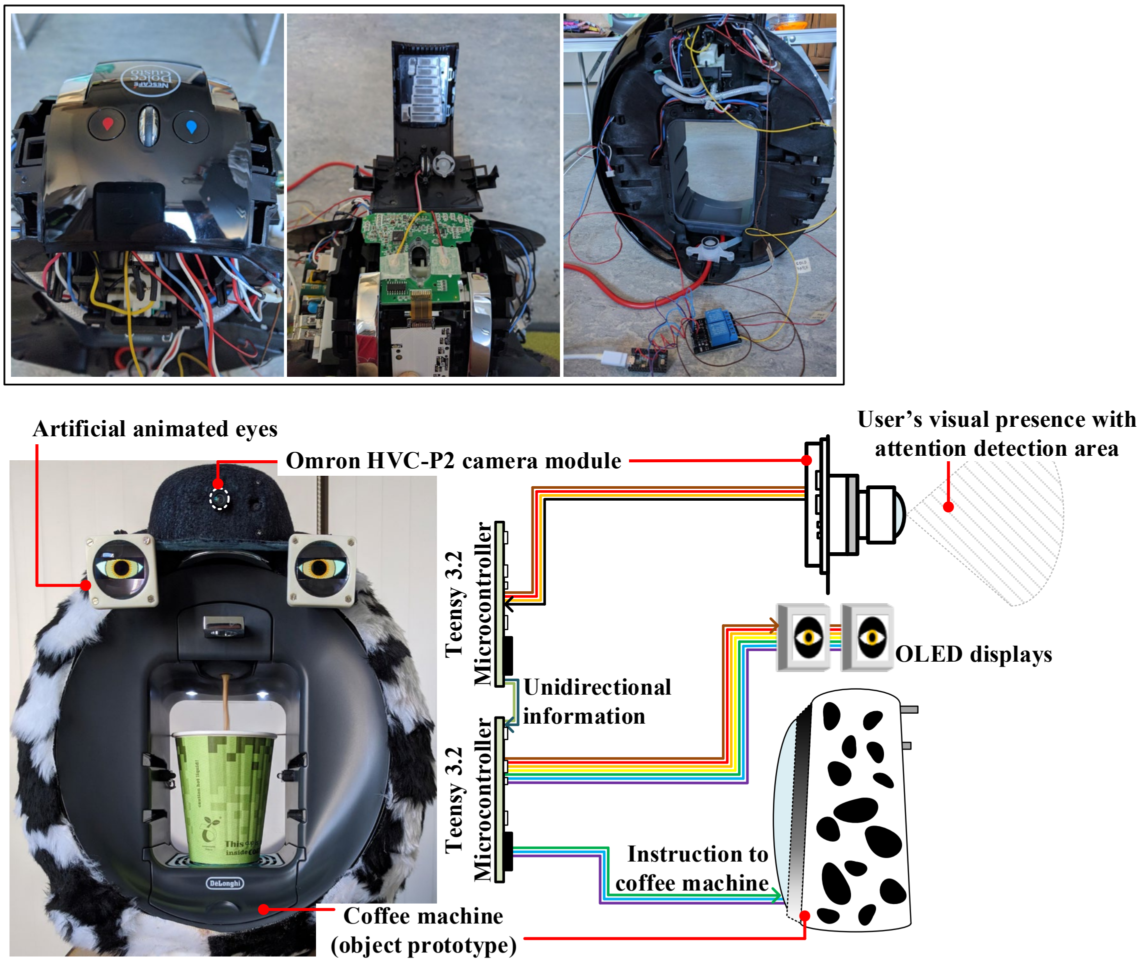


Figure 24 Coffee machine (object prototype) mounted with artificial eyes.

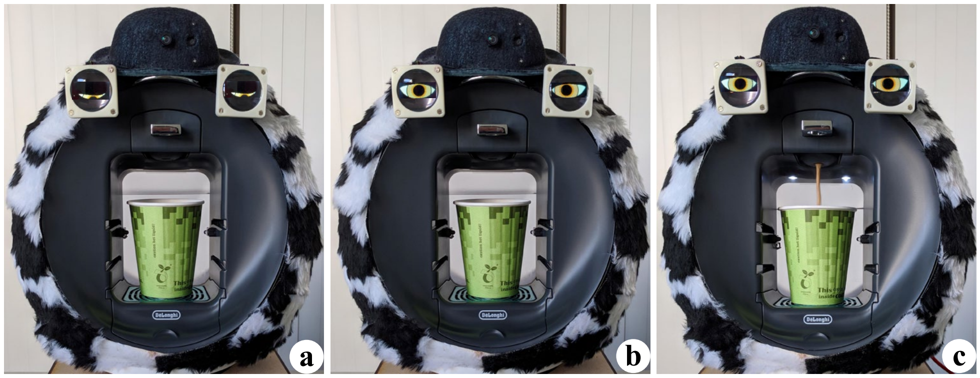


Figure 25 (a) A close-idle coffee machine waits for a potential user, (b) blink-open its eyes upon detecting the user's visual presence with attention, and (c) dispense coffee after having direct eye contact with the user for 3.5 seconds.

4.3. User-Experience Experimental Design

4.3.1. Experimental design

The experiment design was conducted based on three different conditions of eye gaze animations produced by artificial eyes, which were staring eyes (eye-stare), random eye gaze (eye-random), and gaze following (eye-follow). The eye-stare condition consists of four stages, as shown in Figure 26. Next is the eye-random condition, which consists of six stages, as shown in Figure 27. The eye-random condition also introduced a two-second eye contact, and the artificial eyes *look-away* for approximately 1.5 seconds (Andrist *et al.*, 2017). The third condition is the eye-follow, which consisted of six stages, as shown in Figure 28. The condition shown in Figure 28d introduced the artificial eyes that followed the user's eye movements. A delay of one second was given to the artificial eyes to allow the user to notice the artificial eyes' gaze direction when the user made eye contact again. These three conditions will trigger the coffee machine to dispense coffee after 3.5 seconds of eye contact is maintained.

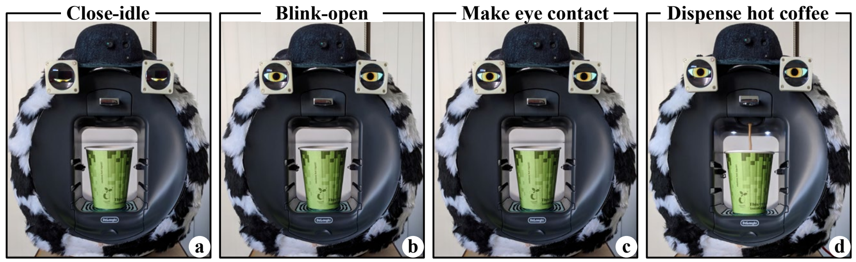


Figure 26 Eye-stare condition: (a) From close-idle, the coffee machine (b)eyes blink-open, (c) make eye contact with the user, and (d) dispense coffee if the user stays engage for 3.5 s.

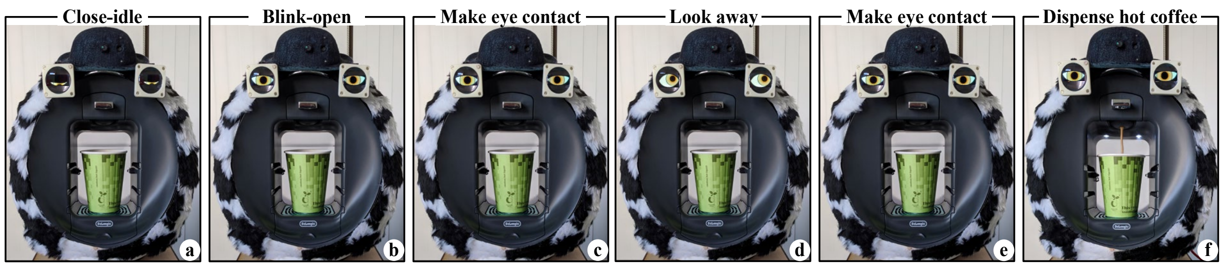


Figure 27 Eye-random condition: (a) From close-idle, the coffee machine (b) eyes blink-open, (c) make eye contact with the user for 1-2 s (d) randomly look-away for 1-1.5 s, (e) make eye contact again and (f) dispense coffee if the user stays engaged for 3.5 s.

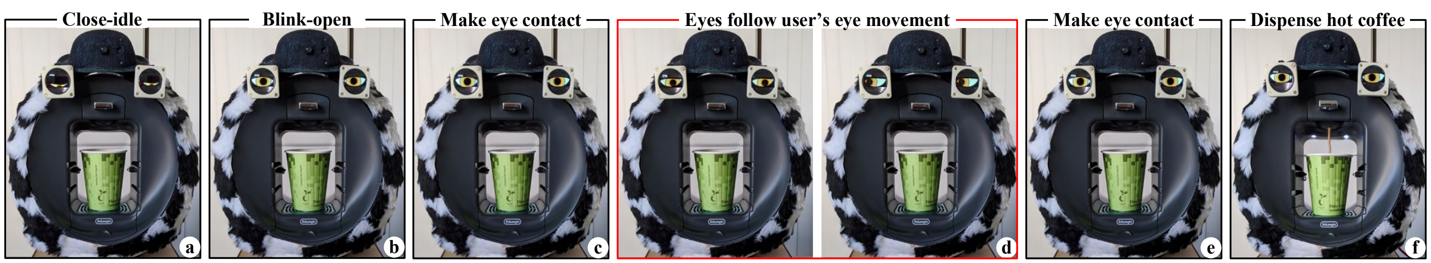


Figure 28 Eye-follow condition: (a) From close-idle, the coffee machine (b) eyes blink-open, (c) make eye contact with the user, (d) interact by following the user's eye gaze, (e) make eye contact again, and (f) dispense coffee if the user stays engage for 3.5 s.

4.3.2. Experimental setup

As shown in Figure 28, an experimental setup was developed to study the user's experience interacting with the coffee machine. Figure 29a shows the overall experiment setup, consisting of a coffee machine, Lego blocks (Lego System, 2020), and Lego instruction booklets. The coffee machine was placed on a box to ensure a parallel eye-to-eye interaction was established with the sitting participant. The distanced between the coffee machine and the participant could be up to 1.3 meters without any head, eye, and body movement restrictions. Therefore, the participant could be busy engaging with other activities (Figure 29b) while at the same time having an eye-to-eye interaction with the coffee machine (Figure 29c).

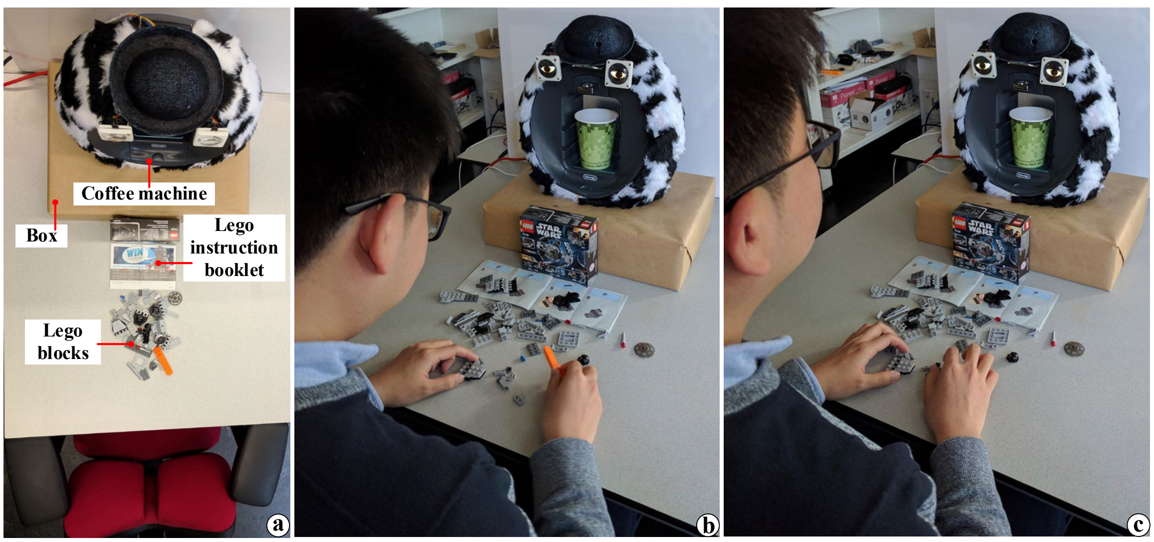


Figure 29 (a) Overview of the experimental setup. (b) A participant builds the Lego set and (c) interacts with the coffee machine while completing the task.

4.3.3. Participants

A between-subjects experiment was designed with 33 subjects (17 males, 16 females, age range: 20 to 35). All the recruited participants have no prior experience interacted with an eye tracker device. The participants were evenly separated into three groups of 11 participants. All three groups were assigned as follows. Group one was assigned to participate in the eye-stare condition. Group two was assigned to participate in the eye-random condition, and group three was assigned to participate in the eye-follow condition. All participants signed written, informed consent to participate in the experiment.

4.3.4. Procedures

As mentioned in section 4.3.3, 33 participants were divided into three groups of 11 participants. Each group must participate in one of the eye-stare, eye-random, and eye-follow conditions. Each participant was allocated six minutes to assemble and connect the Lego blocks by referring to the instruction booklet. Participants were also not informed regarding neither the coffee machine's functionality nor the artificial eyes. A brief instruction was given: *"There is a coffee machine placed in front of you. While assembling and connecting the Lego blocks, you are also free to observe, explore and interact with the coffee machine".* At the end of the experiment, the participants must fill out a set of questionnaires and participate in a post-evaluation interview. The post-evaluation interview was conducted to understand the participant's perspective towards the coffee machine.

4.3.5. Measurement

The life-like interface agent (Parise *et al.*, 1999; Aljaroodi *et al.*, 2019) questionnaire was used to measure the participant's experience interacting with the coffee machine and validate the human-object bidirectional interaction (see *Appendix B*). The life-like interface agent questionnaire consists of eight items divided into three individual scales, which were 1) *Appearance* scale that indicated the overall impression of the coffee machine*,* 2) *Partnership* scale to measure the experience of bidirectional interaction with the coffee machine, and 3) *Interaction* scale to measure the level of understandability during interaction with the coffee machine. Participants must express their experience interacting with the coffee machine marking each item with a seven-point Likert scale where a score of seven stands for *strongly agree* and one stands for *strongly disagree*. Participants were also asked to describe their experience during interaction with the coffee cup briefly. The participants' visual attention was recorded to analyze the participants' eye-to-eye interaction with the coffee machine.

4.4. Results

Figure 30 shows the diverging stacked bar chart summarizing the participants' responses to the seven-point Likert scale of the life-like interface agent questionnaire. The diverging stacked bar chart shows eight items divided into *Appearance*, *Partnership,* and *Interaction* scales. The participants' feedback to the life-like interface agent questionnaire is further analyzed using the non-parametric Friedman test analysis tabulated in Table 2. The test analysis results show statistically significant differences in the *Appearance*, *Partnership,* and *Interaction* scales depending on the eyes conditions while interacting with the coffee machine, $x^2(2)=31.714, p<0.01$. The *Appearance* scale consists of four items: *Human-like*, *Attractive*, *Sociable,* and *Intelligent*. To examine the differences between the four items, a post hoc analysis with six comparisons of *Human-like-Attractive*, *Human-like-Sociable*, *Human-like-Intelligent*, *Attractive-Sociable*, *Attractive-Intelligent,* and *Sociable-Intelligent* items were conducted using the Wilcoxon signed-rank tests. A Bonferroni correction applied, resulting in a significance level set at $p<0.0083$ (six comparisons divided with initial significance level, $p=0.05$). The analysis results show no significant differences between the *Attractive* and *Human-like* items $(Z=-1.368, p=0.174)$ or *Sociable* and *Human-like* items $(Z=-0.315, p=0.753)$. The same goes for the *Intelligent* and *Human-like* items $(Z=-1.711, p=0.087)$, *Sociable* and *Intelligent* items $(Z=-1.648, p=0.099)$, or *Sociable* and *Attractive* items $(Z=-1.858, p=0.063)$. However, there was a significant reduction for the *Appearance* scale in the *Intelligent* vs. *Attractive* items $(Z=-2.928, p=0.003)$. As for the *Partnership* scale, it consists of two items which are *Mutual-like* and *Trustworthy* items. A post hoc analysis of the *Mutual-like-Trustworthy* comparison was conducted using the Wilcoxon signed-rank test. The analysis shows a significant difference between the *Mutual-like* and *Trustworthy* items $(Z=-2.69, p=0.007)$. For the *Intelligent* scale, which consists of two items, *Difficult to understand* and *Enjoyable*, the post hoc analysis with Wilcoxon signed-rank test result also shows a significant difference between the *Difficult to understand* and *Enjoyable* items $(Z=-2.209, p=0.027)$. All the significant differences between the items are marked with a red asterisk, as shown in Table 2. Figure 31 presents the collective distribution of the participants' visual presence with attention during interaction with the coffee machine. The next section will explain the quantitative and qualitative findings to elaborate on the results shown in Figure 30 and Table 2.

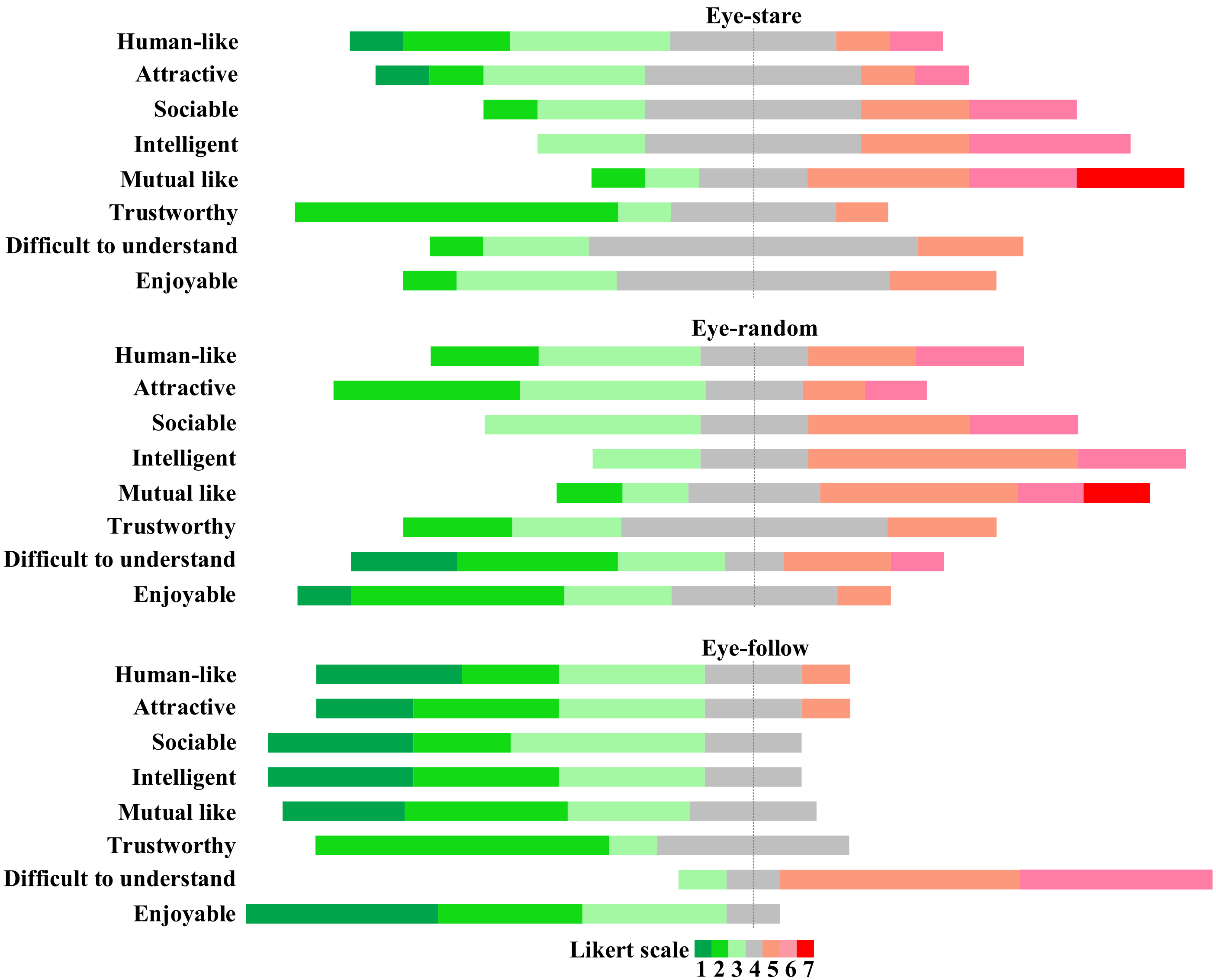


Figure 30 Diverging stacked bar chart of participants' responses to the seven-point Likert for eye-stare, eye-random, and eye-follow conditions where a score of 1 stands for strongly agree, and 7 stands for strongly disagree.

Table 2 Results from the participants towards Appearance, Partnership, and Interaction scales. Significant results ($p_{Apperance}<.0083$, $p_{Partnership}<.05$, $p_{Interaction}<.05$) are marked by a red asterisk (*).

Items	**Mean**	**Median**	**Scales**
Humanlike	4.48	5	*Appearance*
Attractive	4.88 *	5	
Sociable	4.06	4	
Intelligent	4.42	5	
Mutual like	* 3.23	4	*Partnership*
Trustworthy	4.97	5	
Difficult to understand	* 4.08	4	*Interaction*
Enjoyable	5.74	5	

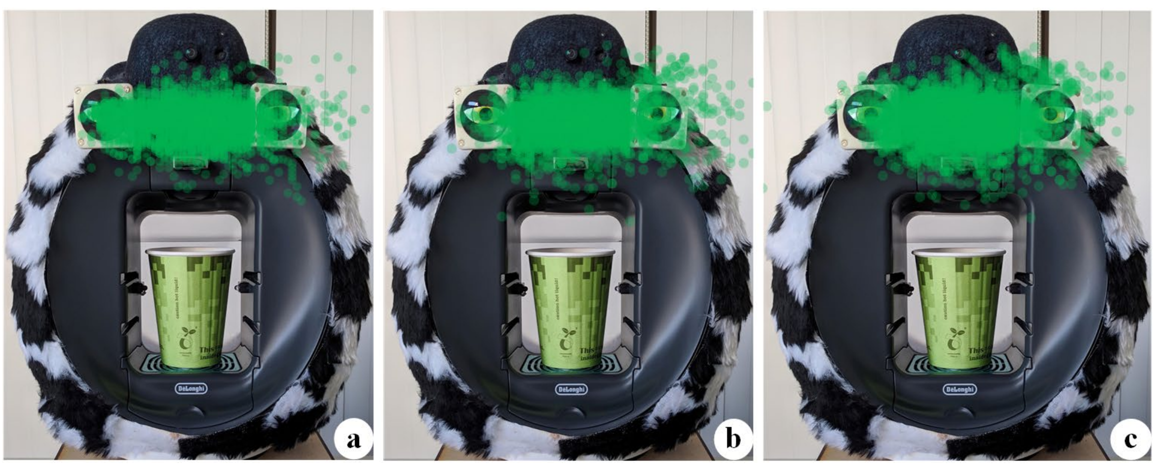


Figure 31 Collective distribution of visual presence with attention for 33 participants during interaction with (a) eye-stare (b) eye-random (c) eye-follow conditions.

4.4.1. Quantitative results

Appearance: impression of the coffee machine

As shown in Figure 30 and Table 2, the coffee machine's appearance is measured based on the *Humanlike*, *Attractive*, *Sociable,* and *Intelligent* items. Median *Appearance* scale levels for the *Humanlike*, *Attractive*, *Sociable,* and *Intelligent* items were 5 (4 to 5), 5 (4 to 5), 4(3 to 5), and 5 (3 to 5), respectively. The post hoc analysis with Wilcoxon signed-rank test result shows a significant difference between the *Intelligent* and *Attractive* items *(Z=–2.928, p=0.003)*.

As presented in Table 2, for the *Human-like* item, participants agree and rated the coffee machine appearance as *human-like*. As shown in Figure 30, 55% of participants from the eye-stare, 45% of participants from the eye-random, and 73% of participants from the eye-follow conditions agree that the artificial eyes mimic the human eye. The percentages show that the artificial eyes can perform as realistic as human eye behavior such as blinking and gazing. The result also proved that the eye blinking behavior performed by the artificial eyes in the eye-stare, eye-random, and eye-follow conditions is essential for participants to establish engagement with the coffee machine by making eye contact.

Next, the embellishment to hide the coffee machine's buttons and a sensor including the cartoonish eyes make the participants rated the coffee machine as attractive. As shown in Figure 30, 45% of participants from eye-stare, 55% of participants from eye-random, and 73% of participants from eye-follow conditions found that the coffee machine appearance was attractive. Specifically, eye-random and eye-follow interactive behavior positively improved the *Attractive* item's rating scale.

The same applied to the *Sociable* item. The artificial eyes mounted on the coffee machine makes the participant rated the coffee machine sociable. As shown in Figure 30, 27% of participants from the eye-stare condition, 36% of participants from the eye-random condition, and 82% of participants from the eye-follow conditions agree that the artificial eyes make the coffee machine sociable. However, staring eyes and the artificial eyes' *look-away* behavior makes the object less sociable.

For the *Intelligent* item, the participants were inclined to rate the coffee machine as intelligent. As shown in Figure 30, 45% of the eye-stare participants and 63% of the eye-random participants rated the coffee machine as unintelligent. The artificial eyes only stare at the participants without any explicit gesture might confuse the participants on how to further interact with the coffee machine. The eye-random *look-away* behavior also negatively impacted the coffee machine as the participants might interpret the behavior as a sign of ignorance from the coffee machine.

Therefore, based on the *Attractiveness*, *Sociable* and *Intelligent* items, participants preferred the artificial eyes that respond to their eye movements. However, starring eyes and random eyes still contribute to the coffee machine's attractiveness, sociability, and intelligence.

Partnership: the experience of bidirectional interaction with the coffee machine

As shown in Figure 30 and Table 2, the partnership with the coffee machine is measured based on *Mutual like* and *Trustworthy* items. Median *Partnership* scale levels for the *Mutual like* and *Trustworthy* items were 4(3 to 5.5) and 5(4 to 6), respectively. The post hoc analysis with Wilcoxon signed-rank test result shows a significant difference between the *Mutual like* and *Trustworthy* items *(Z=–2.69, p=0.007)*.

As presented in Table 2, the participants were undecided whether they can develop mutual like with the coffee machine. As shown in Figure 30, 91% of participants from the eye-follow experience mutual liking with the coffee machine. The percentage suggests that the participants perceived the coffee machine to show the same interest because of the eye-follow reactive behavior that imitates the participant’s eye movements, as shown in Figure 32. However, 63% of the eye-stare participants and 45% of the eye-random participants could not experience mutual liking with the coffee machine. The percentages suggest that without proactive eye gesture besides staring, the participants unable to experience mutual liking with the coffee machine.

Next, for the *Trustworthy* item, the participants rated the coffee machine as an object that can be trusted. As shown in Figure 30, 64% of participants from the eye-stare and eye-follow conditions and 36% participants from the eye-random condition trusted and depended on the artificial eyes when interacting and engaging with the coffee machine. Therefore, this shows that the participants confidently interacted with the artificial eyes in guiding them to engage with the coffee machine.

Figure 32 The artificial eyes and a participant are looking in the same direction, creating the illusion of joint attention.

Interaction: the level of understandability during interaction with the coffee machine

As shown in Figure 38 and Table 2, the experience of interacting with the coffee machine is measured based on the *Difficult to understand* and *Enjoyable* items. Median *Interaction* scale levels for the *Difficult to understand* and *Enjoyable* items were 4(2 to 5) and 5(4 to 6), respectively. The post hoc analysis with Wilcoxon signed-rank test result shows a significant difference between the *Difficult to understand* and *Enjoyable* items *(Z=–2.209, p=0.027)*.

Overall, the participants found the interaction with the coffee machine as either difficult or easy to understand. As shown in Figure 30, 27% of participants from the eye-stare, 64% of participants from the eye-random, and 9% of participants from the eye-follow conditions experienced difficulty understanding the method of interaction with the coffee machine. Hence, this shows that the participants preferred the artificial eyes that respond to their eye movements. However, this only defines the artificial eyes' reactive behavior instead of proactive behavior.

Next, the participants pleasantly experience interaction with the coffee machine as enjoyable. As shown in Figure 30, 36% of the eye-stare participants, 64% of the eye-random participants, and 91% of participants from the eye-follow conditions enjoyed interacting with the coffee machine. Therefore, the artificial eyes manage to attract and encourage the participants to engage with the coffee machine.

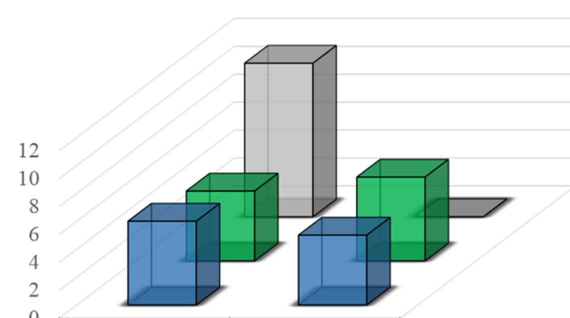
Collective distribution of participants' visual presence with attention

Figure 31 shows the collective distribution of the participants' visual presence with attention during interaction with the coffee machine. The participant's visual presence with attention is captured to analyze the participant's eye-to-eye interaction while engaging in the eye-stare, eye-random, and eye-follow conditions. The participant's visual presence with attention is captured based on the total fixation count during their bidirectional interaction with artificial eyes. Based on the participants' fixation dispersion patterns, the staring eyes shown in Figure 30a help the participants stay engaged with the coffee machine without averting their eye gazes. Figure 31b shows the participants' fixation pattern slightly dispersed due to eye-random *look-away* behavior. Perhaps the participants interpreted the coffee machine's intention to interact as not intended for them. Figure 31c depicts the participants' fixation pattern more widespread due to eye-follow reactive behavior towards the participants' eye movements.

Therefore, based on the dispersion patterns, eye-stare managed to make the participants continuously maintain the bidirectional interaction with the coffee machine compared to eye-random and eye-follow.

The bidirectional interaction success rate

Figure 33 shows the participants' success rate of maintaining a continuous 3.5 seconds bidirectional interaction with the coffee machine. The eye-follow condition shows 11 participants successfully maintained a continuous 3.5 seconds bidirectional interaction. The result suggests that the artificial eyes' reactive behavior towards the participants' eye movement encouraged them to prolong their bidirectional engagement with the coffee machine. The eye-stare condition shows six participants successfully maintained a continuous 3.5 seconds bidirectional interaction with the coffee machine. In contrast, the remaining five participants were unable to prolong their interaction, probably due to the absence of proactive behavior from the artificial eyes during the active engagement. Therefore, the participants failed to maintain the bidirectional interaction with the coffee machine. For eye-random condition, five participants successfully maintained a continuous 3.5 seconds bidirectional interaction with the coffee machine. Initially, the artificial eyes kept constant eye contact with the participants but will perform *look-away* expression after two seconds. Due to the *look-away* behavior, it might tempt the participants to *look-away* too. The result could also suggest that the participants misinterpret *look-away* behavior as a sign of terminating active engagement. Hence, the required 3.5 seconds bidirectional interaction with the coffee machine gets disconnected due to the artificial eyes *look-away* behavior.



Coffee machine dispensed coffee	Yes	No
eye_{stare}	6	5
eye_{random}	5	6
eye_{follow}	11	0

Figure 33 The number of participants' bidirectional interaction success rate.

4.4.2. Qualitative results

Participants experiences interacting with the artificial eyes

The artificial eyes mounted on the top of the coffee machine produce three different eye gaze animations: eye-stare, eye-random, and eye-follow.

Five participants inferred that the staring eyes help them maintain bidirectional interaction and engagement for the eye-stare condition. The following statements show the participants' reflection.

"When I made eye contact, I am engaged with the coffee machine." (P1)
"Although the eyes were unresponsive, it able to grab my attention and enable myself to stay engaged with the coffee machine throughout the interaction." (P2)
"Made eye contact was necessary to engage with the coffee machine, and because the eyes stared at me without looking elsewhere, it encouraged me to do the same too." (P5)
"The staring eyes indicated the coffee machine's readiness to communicate with me, and I could easily engage with it by making eye contact." (P7)
"When I looked at the eyes, I entered into an interaction with the coffee machine." (P9).

Whereas three participants mentioned that the staring eyes indicate the object's intention to interact, hence encouraging them to stay engaged with the coffee machine. Therefore, three reflections from the participants are presented in the following:

"I made and hold eye contact with the coffee machine for too long because I thought that it wanted to communicate something which made me forgotten about the Lego blocks." (P3)
"I made eye contact with it, but nothing happened, and it kept staring at me, which was very uncomfortable. If the coffee machine closed its eyes, I could assume that it does not want to interact with me, and I could pay more attention to assembling the Lego blocks." (P6)
"I felt that the coffee machine was waiting for me to react appropriately, and it was toughed for me to focus on assembled the Lego blocks." (P11)

For the eye-random condition, six participants emphasized the artificial eyes' *look-away* behavior as a sign of reluctance to interact with them. Therefore, six reflections from the participants are presented in the following:

"The coffee machine wanted nothing further to do with me." (P3)
"It was an obvious indication that the coffee machine refused to acknowledge my presence." (P4)

"The coffee machine tried to ignore me, but suddenly, I got coffee out of it, which was very confusing." (P6)
"I felt the behavior unpredictable and therefore made me ignored the coffee machine's presence." (P8)
"The coffee machine was internally busy with something and was not willing to interact with me." (P9)
"The eyes were used to enable the coffee machine to engage with me, but I understand the opposite of it (refused to interact)." (P11).

Whereas three participants mentioned, they ignored the artificial eyes *look-away* behavior and focused on assembling the Lego blocks. Therefore, three reflections from the participants are presented in the following:

"The coffee machine's annoying behavior made me move a bit sideways so that I did not have to interact with it anymore and focused on assembling the Lego blocks." (P1)
"I rather played with the Lego blocks than tried to understand the look-away behavior." (P2)
"I pay no attention to the coffee machine after failed to maintain a good eye contact." (P5)

For the eye-follow condition, five participants claimed that they could influence the artificial eyes' behavior. Therefore, five reflections from the participants are presented in the following:

"The eyes were responsive, and I could persuade the coffee machine to focus on the Lego blocks." (P1)
"I able to influence the coffee machine's eyes to follow my eyes." (P4)
"I could capture the coffee machine attentions, and my gazed influenced the coffee machine to respond accordingly" (P5)
"The coffee machine mimicked my eye movements." (P7)
"The coffee machine waited for me to make any sudden eye movements, and it mirrored it." (P8)
"The coffee machine's eyes were interactive and predictable because it followed my eye movements." (P11).

Two participants voluntarily constrained their working space to maintain continuous engagement with the coffee machine. Therefore, two reflections from the participants are presented in the following:

"I realized that the eyes were the only interactive part of the coffee machine. Therefore I made sure that I assembled the Lego blocks within a small space to let the coffee machine observed what I am currently doing." (P2)
"The coffee machine could not move its body. That was why I decided to complete the task in front of it so that I could show and shared which Lego blocks I picked up." (P6)

Artificial eyes as the enabler for establishing interaction and engagement

This section explained the participants' experiences with artificial eyes to establish interaction and engagement with the coffee machine.

For the eye-stare condition, seven participants explained the artificial eyes enabled the participants to engage with the coffee machine. Therefore, seven reflections from the participants are presented in the following:

"The instant I saw the coffee machine had eyes; I knew that I should make eye contact with it." (P2)
"The blinking eyes made me believed that the coffee machine was capable of looking around, and therefore, I should make eye contact to interact with it." (P3)
"It was easy to understand the modality of the interaction to communicate with the coffee machine." (P5)
"The eyes were useful to confirm that I am engaged with the coffee machine." (P8)
"I am more focused on making eye contact rather than to find any other ways to interact with the coffee machine." (P9)
"The eyes were an obvious indication of how I should be engaged with the coffee machine." (P10)
"It always looked at me, and I felt that it wanted to engage and communicated something to me, and for that, I should look back." (P11)

Six participants mentioned they failed to engage with the artificial eyes for the eye-random condition because of the *look-away* behavior. Therefore, six reflections from the participants are presented in the following:

"I already engaged with the eyes, but it look-away, which was very confusing." (P1)
"I am unable to attract its attention to me because it always looked away from me." (P2)
"I could not engage at all with the eyes because it tried to ignore me." (P5)

"The coffee machine did not have any intention to interact with me because it always looked away whenever I made eye contact with it." (P9)
"The coffee machine looked bored of me, and because of the jumpy eyes, it made me wanted to focus more on assembled the Lego blocks instead of focused on interacted with the coffee machine." (P10)
"I attempted to make eye contact several times, but it always averted its gazed, which was very rude." (P11).

For the eye-follow condition, nine participants pointed out that the artificial eyes' responsiveness confirmed the coffee machine's intention to establish a continuous bidirectional interaction. Therefore, nine reflections from the participants are presented in the following:

"I realized that the coffee machine wanted to interact when it looked to where I looked." (P1)
"*The coffee machine tried to persuade me into paying attention to it, and when I did, it gave me coffee."* (P2)
"I am surprised that the coffee machine could reciprocate my eyes. It made me understand that it wanted to engage with me." (P4)
"After several eyes interaction, I could comprehend the coffee machine's intention when it gave me coffee when I made a stabled eye contact." (P6)
"When I looked at the Lego blocks, the coffee machine did the same too. It showed that the coffee machine wanted to be involved." (P7)
"It responded to my eye movements and dispensed hot coffee when we made eye contact." (P8)
"I picked up the Lego blocks, and the coffee machine looked at it, and the moment I looked at the coffee machine, it looked back at me." (P9)
"It was an obvious indication that the coffee machine wanted to interact with me." (P10)
"It was very engaging of the coffee machine to divide its attention to me and to focus on the Lego blocks." (P11).

Maintaining bidirectional interaction with artificial eyes

Six participants considered the lack of responsive behaviors besides blinking for the eye-stare condition made the interaction a non-interactive session. Therefore, six reflections from the participants are presented in the following:

"There were no obvious behaviors that made me felt that the coffee machine wanted to start communicated something to me (non-proactive)." (P4)

"I felt engaged with the coffee machine when we made eye contact, but the eyes only stared at me." (P7)
"I looked back, but nothing happened, which made me assumed the eyes was used as a decorative accessory." (P3, P8)
"The eyes were very well designed, but I am disappointed when it just looked at me." (P6)
"The coffee machine wanted to interact with me by staring and blinking. However, that was just normal eye behavior, which was unusual to be conveyed as an intention to interact. Normally people made eye contact and then spoke to initiate a conversation." (P9).

For the eye-random condition, five participants implied the artificial eyes *look-away* behavior expressed the coffee machine's intention to either engage or disengage with them. Therefore, five reflections from the participants are presented in the following:

"We both made eye contact, and I thought that something engaging would happen next, but it looked away from me, which was an obvious signed of avoided the interaction." (P1)
"I assumed we had a conflict of interest. I am willing to interact with the coffee machine, but the coffee machine undecided whether to proceed with the interaction or not." (P5)
"The coffee machine gave mixed signals. We made eye contact, then it looked away and looked back, acted like it was interested or tried to ignore me." (P8)
"It was impolite of the coffee machine to ignore my request (eye contact). It gave a negative impression, and I decided to withdraw myself from being involved with the coffee machine and focused on assembled the Lego blocks." (P10)
"The interaction was unstable and very confusing. It was either the coffee machine wanted to interact or a signed of rejection." (P11)

Five participants expressed being overwhelmed with the artificial eyes' responsiveness during interaction and engagement with the eye-follow condition. Therefore, five reflections from the participants are presented in the following:

"I am always curious the whole time, looking at the coffee machine's eyes, checked its current eye movement just to make sure where the coffee machine looked at." (P3)
"I felt obligated to interact with the object, especially after it gave me coffee. Moreover, it was tough to divide my attention assembled the Lego blocks and to keep on interacting with the coffee machine." (P4)
"The eyes were too interactive, and I could not get out of the interaction." (P7)

"I thought the coffee machine would be in a sleep mode after dispensed the coffee, but it still interacted with me. Should I do something else to satisfy its needs?" (P8)
"There was no collaboration happened between us, and I am the one who controlled and guided the coffee machine's behavior, which was very tiring." (P11).

4.5. Discussion

This section discusses the artificial eyes' overall findings for three different conditions of eye gaze animations: eye-stare, eye-random, and eye-follow.

A neutral impression was shown towards the *Appearance*, *Partnership,* and *Interaction* scales for eye-stare. Even though eye-stare lack responsive behavior besides blinking and staring, the participants maintained an active, stable, and continuous bidirectional interaction with the coffee machine. As shown in Figure 31a, participants' visual presence with attention stays engaged with the coffee machine.

For eye-follow, participants showed a positive impression towards the *Appearance*, *Partnership,* and *Interaction* scales. The eye-follow responsiveness to imitate and react towards the participant's eye movements attracts the participants to interact and engage with the coffee machine. However, the eye-follow condition interprets reactive behavior instead of proactive behavior, which leads to overwhelming the participants to continuously initiate the interaction to maintain engagement with the coffee machine. As shown in Figure 31c, the participants' visual presence with attention dispersed due to eye-follow reactive behavior. Consequently, the participants could not maintain a stable and continuous bidirectional interaction and engagement with the coffee machine.

A positive impression was shown toward the Appearance scale for eye-random, whereas a neutral impression was shown toward the *Partnership* and *Interaction* scales. The responsiveness of eye-random attracts the participants to maintained bidirectional interaction and engagement with the coffee machine. However, once the coffee machine *look-away*, the active bidirectional interaction and engagement between the coffee machine and the participant were disconnected. As shown in Figure 31b, the participants' visual presence with attention is slightly dispersed due to eye-random *look-away* behavior. Consequently, the participants were only momentarily able to maintain the bidirectional interaction and engagement with the coffee machine.

Nonetheless, based on eye-random and eye-stare results, the artificial eyes that stare at the participant are seen as useful visual feedback that helps them notice the coffee machine's intention to interact. Artificial eyes that stare at the participant also enable them to maintain a stable bidirectional interaction and engagement with the coffee machine. However, to retain a continuous bidirectional interaction and engagement, the artificial eyes need to display proactive expression besides staring and blinking. Therefore, artificial eye expressions such as winking and pupil dilation can improve the artificial eyes' proactive behavior. Thus, with eye-to-eye interaction that creates bidirectional interaction and winking as proactive expression, this implementation could allow the participants to experience perceptual crossing based on the perceptual crossing paradigm.

4.6. Conclusion

The artificial eyes (i.e., visible expressive perceptual quality) mounted on an object, and the participants experienced bidirectional interaction with the object has been tested and validated. The results show the artificial eyes that animates eye-stare, eye-random, and eye-follow have successfully attracted the participants to achieve bidirectional interaction with the object. Nevertheless, eye-follow only interprets reactive behavior, where the eyes react by mimicking the participant's eye movements. For eye-stare and eye-random, staring and blinking expression helps the participants to experience bidirectional interaction with the object. However, to maintain a continuous bidirectional interaction, proactive expression besides staring and blinking from the artificial eyes needs to be implemented to allow the participant to experience perceptual crossing. Based on the perceptual crossing paradigm, the interaction between two entities must be bidirectional and proactive to experience perceptual crossing. Therefore, artificial eyes with proactive expression are proposed. Hence, the next chapter aims to implement artificial eyes with proactive expressions such as winking and pupil dilation to allow the participant to maintain continuous bidirectional interaction with the object.

Chapter 5.

Study 3: Designing Proactive Interaction with Artificial Eyes

In Chapter 4, the user-experience experiment results have revealed that with artificial eyes that stare and blink, the participants can experience bidirectional interaction with the prototype. However, the participants could not experience perceptual crossing with the prototype due to the absence of proactive expression from the artificial eyes. Therefore, proactive expressions from artificial eyes such as winking are introduced in this chapter[4]. Before validating the artificial eyes' proactive expression, a Session Initiation for Proactive Object (SIPO) conceptual model is proposed based on the Session Initiation Protocol (SIP) and the perceptual crossing paradigm. The SIPO conceptual model depicts how an object can maintain bidirectional engagement between the object and the people in its surroundings and proactively express its intention towards these people. A prototype mounted with artificial eyes and placed on a rotating base is designed using proof-of-concept implementation to validate the SIPO conceptual model. The prototype is tested in single- and multi-user scenarios. Based on the pilot study results, the prototype design's viability is confirmed and can show its intention clearly through artificial eyes and simple abstract motion. The pilot studies prototype is simplified to let the user focus on the expression rather than the object itself. The simplified prototype is used in the real-environment user study and the crowd-sourced video-based user study for further testing. Both user studies show that winking can be a useful expression that makes the user view the object as proactive and encourages reciprocal input.

5.1. Introduction

According to Auray *et al.*'s perceptual crossing paradigm, perceptual crossing happened when both entities experience bidirectional and proactive interaction. Even if the communication channel is reduced to a bare minimum, such as eye-to-eye contact communication, people can still experience perceptual crossing as long as they are involved in a bidirectional and proactive interaction. A study conducted in Chapter 4 confirmed the bidirectional interaction occurs during eye-to-eye contact

---

[4] This chapter is written partly based Anas et al. (2020).

communication. However, the participants cannot maintain the bidirectional interaction during the eye-to-eye contact communication; therefore, the perceptual crossing is not achieved.

Winking is a minimalist expression that voluntarily shows an intention to the person winked at (Admoni and Scassellati, 2017; Goldstein, 2019); it retains a natural level of eye contact and engagement and works without eyebrows (Lopez *et al.*, 2017). This eye-based gesture does not introduce distractions and requires no extra mechanism, unlike implementing head/hand/body gestures and voices. Hence, winking is a suitable gesture that indicates proactive expression.

To investigate artificial eyes' design with proactive expression, a conceptual model called Session Initiation for Proactive Object (SIPO) is proposed based on the perceptual crossing paradigm (Figure 34a). This model, adapted from the INVITE method of the Session Initiation Protocol (Ahson and Ilyas, 2018), depicts how an object can maintain engagement between the object and the people in its surroundings and express intention. A proactive object searches for a user of interest using its sensors (i.e., computer vision). Once found, it expresses its intention by turning its body orientation towards the user and then starts winking at the user. The user who receives the signal understands that the object intended to interact with them and, therefore, engages with the interaction by looking back. Meanwhile, the perceptual crossing is established, and the communication session starts. The user also can ignore the signal or terminate the session of their own will using established communication, and the object will return to continue exploring the environment, if appropriate.

Based on this model, a minimalist prototype design (Figure 34b and Figure 34c) is proposed to examine the conceptual model. As mentioned by Dove *et al.* (2017), imposing gender on an object can affect the way users respond to it, as it may influence both the human-object engagement (Iacobelli and Cassell, 2007) and the perceived task (Forlizzi *et al.*, 2007). Therefore, a pair of expressive gender-neutral cartoonish eyes are placed on the top of a primitive shape to avoid the impact of gender stereotyping (Bryant *et al.*, 2020). A servo motor is also placed at the bottom of the object, allowing its body to be oriented (Figure 35). The minimal expression is deliberately used to keep the user's focus on the expression rather than the object itself; thus, a more generalizable understanding of the interplay between the artificial eyes and the mechanical movements can be obtained.

The prototype design is used to conduct a series of studies with users, including two pilot studies, a 33-participant real-environment user-study, and a 240-participant crowd-sourced user study using Amazon Mechanical Turk. The first pilot study results suggest that expressive perceptual quality (i.e., artificial eyes) needs to be visualized in this context. The second suggests that combining mechanical movements with artificial eyes can signify the object's intention to initiate interaction more effectively than eye movements alone. Based on the results, the interaction model is evaluated by designing a real-environment user study and ten videos involving a proactive object interacting in single or multiple-user scenarios. The results show that people perceive winking as a useful perceptual quality of an eyes-embedded object, making the users feel that the object is proactive.

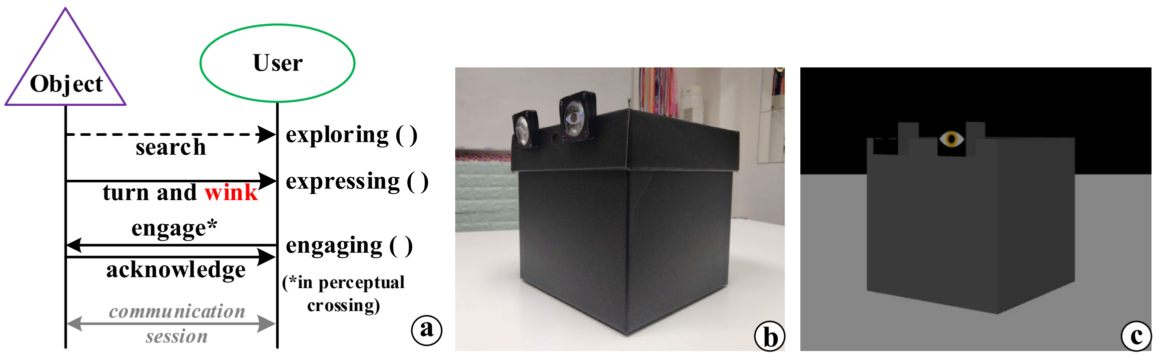


Figure 34. SIPO (Session Initiation for Proactive Object) model based on perceptual crossing: (a) a proactive object finds a user of interest. It expresses its intention by turning its orientation and winking at the user. The user receives the intention and engages in the communication session; (b) a proactive physical object that turns and winks; (c) a proactive virtual object that turns and winks.

Figure 35 Abstract motion is introduced to allow the object to search while identifying a potential user's interaction.

5.2. The SIPO Conceptual Model

Figure 36 shows the state diagram of the SIPO model, which is the interaction diagram's counterpart shown in Figure 34a. A proactive object in its idle state moves to the 1) exploring state by searching for a user of interest using its embedded sensors (e.g., camera, microphone). Once an intended user is found, it moves to the 2) expressing state by showing its intention via some expression (e.g., visual animation, audio messages, mechanical movement). It then moves to the 3) engaging state if the user gives reciprocal input (e.g., looks back, vocally responds to it, turns toward the object) within a time threshold T. Perceptual crossing is then established, and the communication session starts. Object 4) terminates the communication if the user ignores the object's signal, yielding the response time $t_r > T$ or terminating the communication. After a session is terminated, the object will continue to explore the environment, if appropriate.

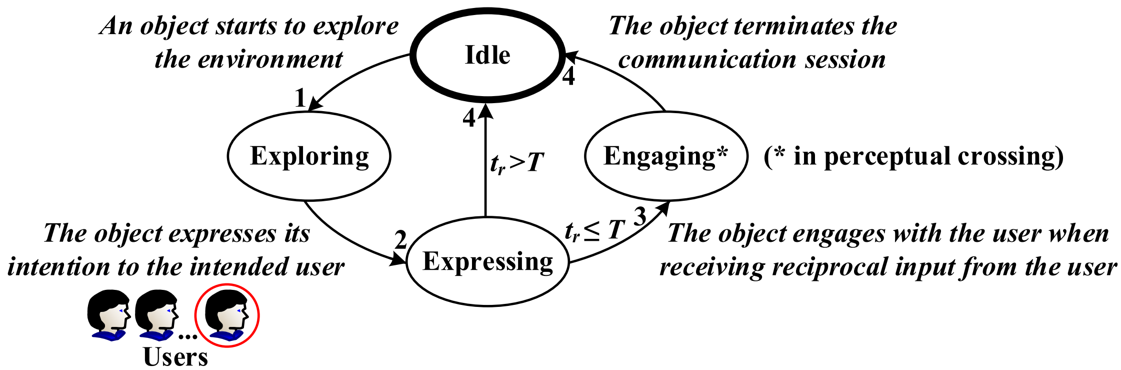


Figure 36 State diagram of the SIPO conceptual model (t_r: response time of the user).

5.3. Expressing Intentions

The object's expressive intentions can be enhanced using artificial eyes mounted on top of it. The previously introduced artificial eyes in Chapter 4 are implemented with a winking expression. A single wink is a standard signal that could mean a silent agreement between two people and is usually a friendly gesture implying a degree of intimacy (Admoni and Scassellati, 2017; Goldstein, 2019). The simplest way to realize a wink is through a more extended blink in one eye. The winking duration must be noticeably longer than normal blinking so that the engaging user can recognize the wink signal, as shown in Figure 37a. After the engaging user realized the winking and stays connected with the object for some time, the artificial eyes will dilate its pupil, as shown in Figure 37b as a subtle acknowledgment to maintain a continuous interaction. In face-to-face communication, a person's pupillary area will enlarge about 1.5 times when he/she is in control of the communication (Sejima *et al.*, 2018). This finding showed that pupil response has relationships in eye-to-eye communication. Therefore, pupil dilation could be used as a signal for maintaining engagement.

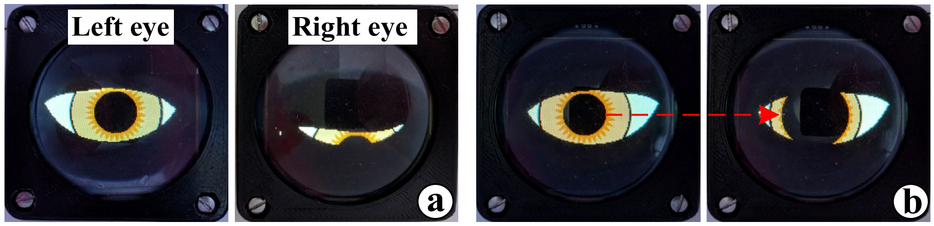


Figure 37 Eye expressions: (a) winking; (b) pupil dilation.

Apart from expressing intention with winking and pupil dilation, the object can also express its intention by paying attention to someone. This attention can be expressed in two visually observable ways: rolling the eyes to look at someone (Ruhland *et al.*, 2014) or turning to someone to look straight at him/her (Vázquez *et al.*, 2017). Gaze movement can be easily achieved using the direction of gaze of the eyes, as shown in Figure 22a, while body movement requires a mechanism of actuation. Figure 38a shows an example design of a mechanical rotating base consisting of a stepper motor mounted to a hub and gears and fixed to a round plate. The rotating base allows the body orientation of any object placed on it to be changed, as shown in Figure 38b.

Figure 39 shows the previously introduced coffee machine in Chapter 4 and is placed on the rotating base to control its body orientation. A compact Omron HVC-P2 (Mouser Electronics, 2016) camera module with a multi-function image sensor that natively supports face detection and gazes direction is used to search for users and identify their engagement.

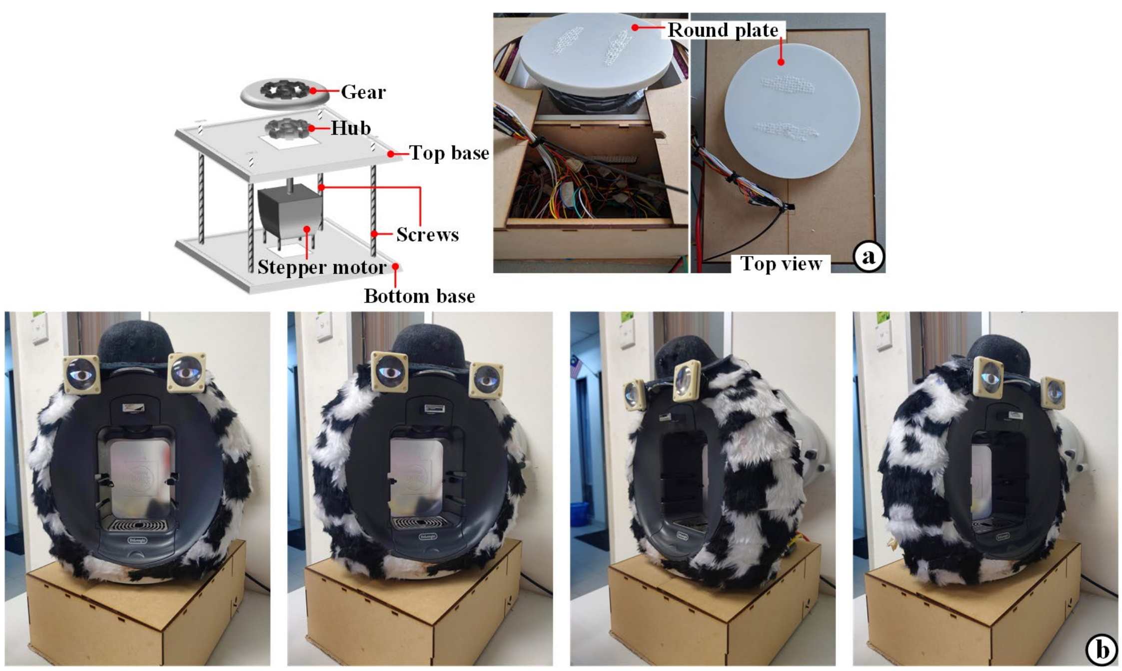


Figure 38 (a) Design of the mechanical rotating base. (b) The rotating base allows an object to orient its body towards the intended user.

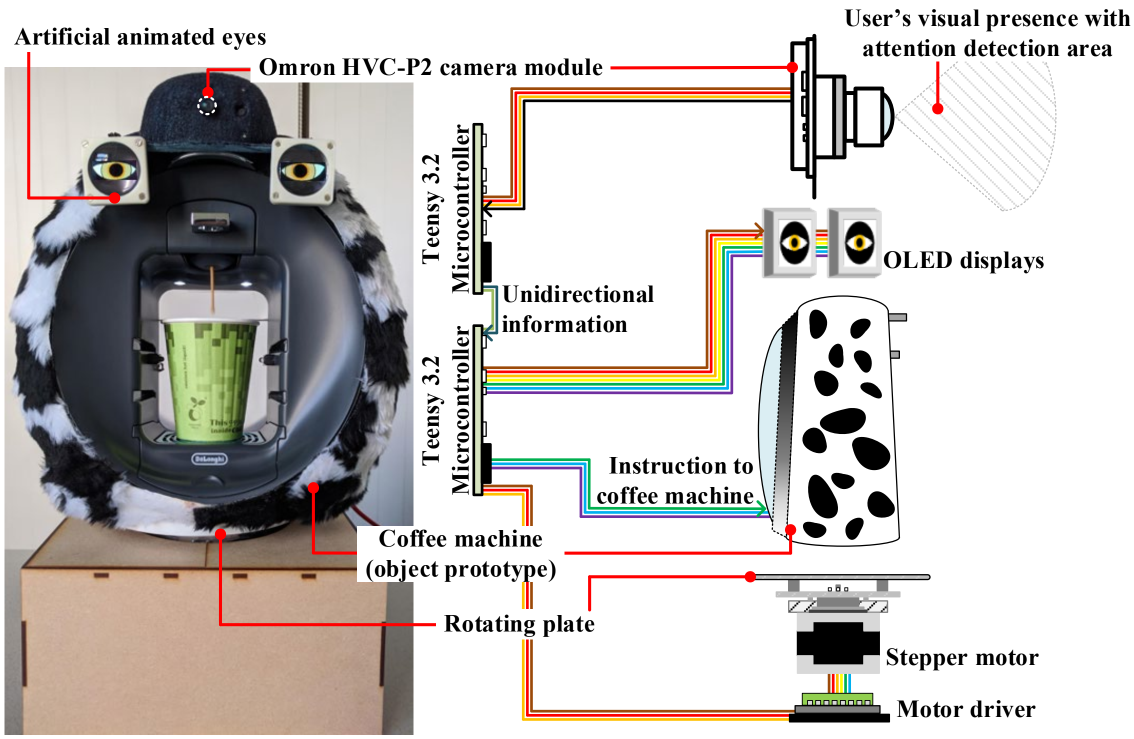


Figure 39 Prototype design of a coffee machine with proactive behavior.

5.4. Pilot Studies

Two pilot studies were conducted to investigate the viability of the prototype design. The coffee machine prototype was evaluated in both of these studies.

5.4.1. Pilot study 1: Visibility of the perceptual quality

The first pilot study investigated the perceptual quality's visibility for allowing the object to express the intention in a two-user scenario. Two situations were investigated:

1) *Eyes-Motion*: With the artificial eyes mounted, the coffee machine expressed its intention by reorienting its body and looking at the user, as shown in Figure 40a.
2) *Motion-Only*: Without the artificial eyes mounted, the coffee machine expressed its intention by reorienting its body toward the user, as shown in Figure 40b.

Fourteen participants (six males, eight females) aged from 24 to 35 years were recruited and separated into two equal-sized groups. One group encountered an *Eyes-Motion* object, and the other a *Motion-Only* object. Each participant was asked to describe their experiences after the situation.

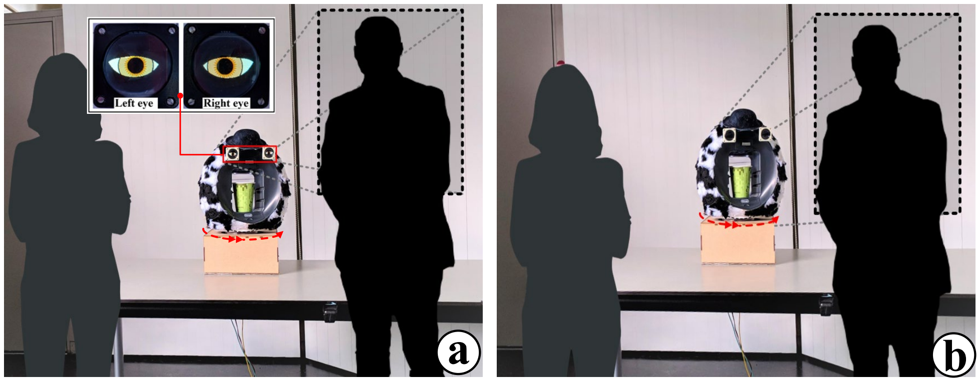


Figure 40 Pilot study 1: (a) Eyes-Motion coffee machine and (b) Motion-only coffee machine

Results

In the Eyes-motion study, participants mostly used the words "aware of me," "recognized me," "noticed me," "interacted with me," "observed me," "engaged with me" to describe the experiences. These statements show that they perceived the Eyes-Motion behavior from the object to interact with them. On the contrary, in the *Motion-Only* study, participants mostly used the words "followed me," "detected my movement," "reacted to my presence," and "responded to my movement" to describe their experiences, showing that they perceived the movement as a mechanism.

Several participants in the *Eyes-Motion* study investigated the possibility of another kind of interaction besides making eye contact, such as trying to blink at the same time as the coffee machine and speculating on how to get a cup of coffee through non-verbal input. In contrast, in the *Motion-Only* study, several participants attempted to understand the mechanism of the machine. They identified a sensor that could sense their presence and questioned the coffee machine's motivation reorienting its body toward them.

Discussion

Overall, the presented results in Pilot Study 1 suggest that the artificial eyes and the abstract motion are required to allow the coffee machine to express the intention to initiate interaction with the user.

5.4.2. Pilot study 2: Expression of intentions

The second pilot study investigated the interplay between eye gaze direction and the abstract motion to express the intention in two-user scenarios. Two situations were implemented:

1) *Motion-Gaze*: With the artificial eyes mounted, the coffee machine showed attention to a participant by reorienting its body and then gazing at him/her, as shown in Figure 41a.
2) *Gaze-Only*: With the artificial eyes mounted, the coffee machine showed attention to a participant by rolling its eyes to stare at him/her, as shown in Figure 41b.

Seven participants (four males, three females) aged from 23 to 33 years were recruited and participated in both situations in a counterbalanced order. In each situation, each participant was involved in a two-user scenario with another experimenter. The participants were led to believe that the experimenter was another user in the scenario. They were then asked to describe their experiences after each of the situations.

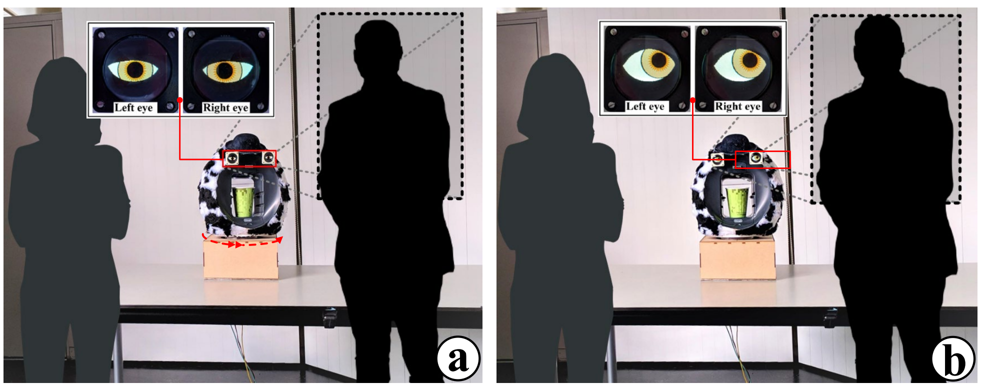


Figure 41 Pilot study 2: (a) Motion-Gaze coffee machine; (b) Gaze-Only coffee machine

Results

In the *Motion-Gaze* study, participants mostly used the words "instantly understood," "felt engaged and connected," "predictable behavior," and "capable of noticing people" to describe their experiences. These statements show that the attention in the *Motion-Gaze* study was perceived as an intention of the object. On the contrary, in the Gaze-Only study, participants mostly used the words "unpredictable," "misleading," "avoided engagement," "had a mind of its own," "unintelligent," and "suspicious" to describe their experiences. These statements show that the attention shown in the *Gaze-Only* study was not well perceived.

Participants in the *Motion-Gaze* experiment appeared to understand these behaviors immediately. They reacted positively to the coffee machine, mainly when it chose to engage with them rather than the other person in the scenario. Conversely, several *Gaze-Only* participants appeared confused since eye gaze direction did not indicate the coffee machine's intention to interact. Several of them suggested adding more output modalities to support the communication, such as using audio messages or an arrow in the graphical display indicating the machine's intention to interact.

Discussion

Overall, the presented results in Pilot Study 2 suggested that expressive artificial eyes are required to show the intention to initiate an interaction and engagement. The results also suggested that the coffee machine should appropriately deliver a more transparent intention through the combination of gaze and motion.

5.5. Proof-of-Concept Implementation and Example Scenario

This section explained the proof-of-concept implementation and example scenarios of a coffee machine initiating interacting with its potential user based on the developed SIPO conceptual model. The coffee machine's proactive behavior can be used or implemented to improve smart consumer equipment. As mentioned by Park *et al.* (2019), productivity studies suggest that periodically taking a proper break (such as 52 minutes of work followed by a break of 17 minutes) is suitable for a worker's health and productivity at work. Therefore, a proactive coffee machine is implemented that can initiate a coffee break to improve employee productivity.

Figure 42 shows the coffee machine initiates interaction with the users. The coffee machine starts to explore the environment looking for users using face tracking, as shown in Figure 42a. Referring to Figure 42b, after the coffee machine explores the environment, the coffee machine identified a potential user interested in interacting with the coffee machine. After successfully identifying the potential user, the coffee machine orientates its body towards the potential user, as shown in Figure 42c. Although the coffee machine could also avert its gaze at the potential user without orienting its body, the conducted pilot study 2 result shows that averted gaze did not indicate the coffee machine's intention to interact. Also, in recent work conducted by Kiilavuori *et al*. (2021), eye contact with head-turning elicits attentional reactions compared to averting gaze side-ways. Therefore, orienting and looking directly at the potential user is more suitable for capturing someone's attention and initiating an interaction. Once eye-to-eye contact is established, the coffee machine winks at the user as a friendly invitation of expressing an intention to initiate an interaction, as shown in Figure 42d. After winking and the user continuously engage with the coffee machine, the coffee machine dilates its pupils, as shown in Figure 42e. The pupils dilated as an acknowledgment of establishing perceptual crossing with the user. The coffee machine's pupil dilation also expresses the intention to dominate the interaction, attracting the user to maintain the engagement (Sejima *et al.*, 2018). The communication sessions between the coffee machine and the user can start afterward.

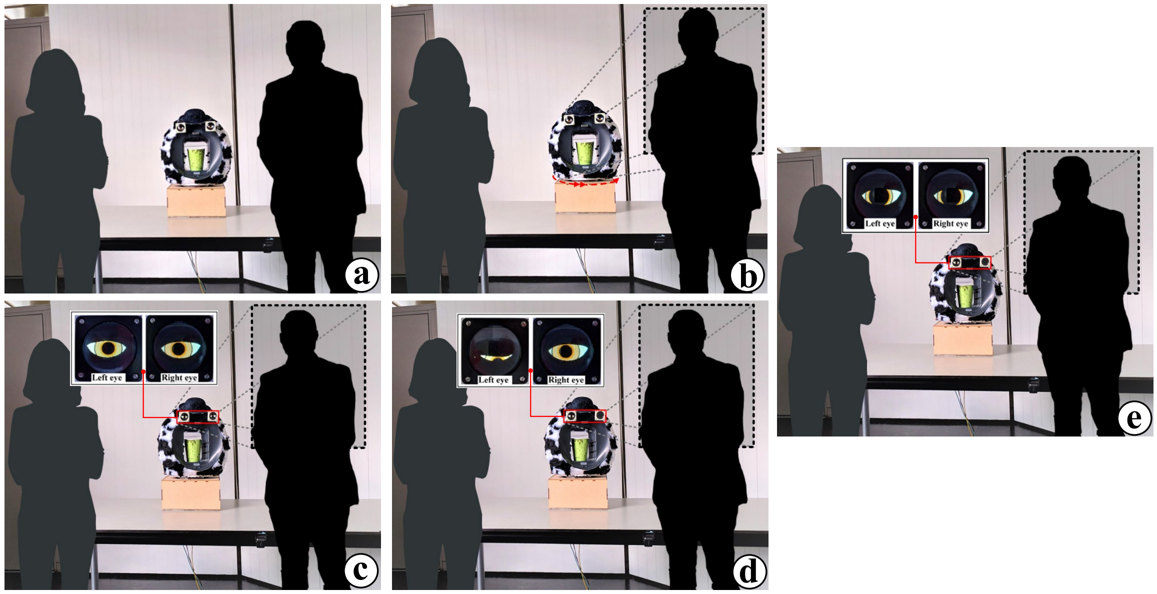


Figure 42 Initiating a session: (a) an idle coffee machine searches for potential users; (b) turns to a user of interest; (c) looks straight;(d) winks at him and (e) dilates its pupils and the communication session starts.

After the perceptual crossing is established, the user decides whether to further communicate with the coffee machine using the supported modality, such as voice commands or to terminate the communication session. Notably, when the coffee machine finds the user ignoring the invitation by not looking back, as shown in Figure 43a, it turns to another nearby user. It tries to initiate a new perceptual crossing session shown in Figure 43b. If the coffee machine cannot detect any potential users, the coffee machine goes into hibernating and shuts its eyes until the next coffee break, as shown in Figure 43c.

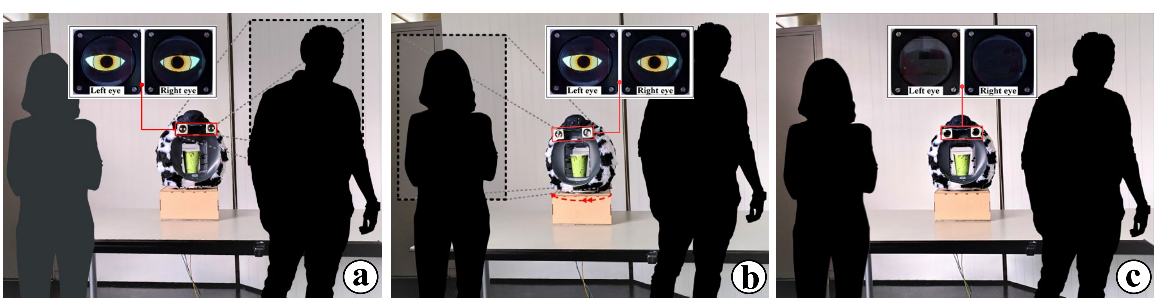


Figure 43 Terminating a session: (a) the user of interest looks away, so the coffee machine (b) turns to the next user of interest, and (c) hibernating by shutting its eyes when no potential users' present.

5.6. Evaluation: Real-Environment User Study

Instead of using the previously introduced coffee machine, a new minimalist prototype design is proposed. The minimal design is deliberately used to keep the user's focus on the expression rather than the object itself.

Figure 44a shows a physical Box mounted with a pair of artificial eyes and a compact Omron HVC-P2 camera module on top of it. The camera module that natively supports face detection and gaze direction allows the object to search for users and identify their engagement. The camera module is also mounted together with the artificial eyes, allowing the object to detect the user's direct visual attention when making eye contact. In addition to that, abstract motion is introduced. It consists of a rotating base controlled by a low-torque micro servo motor and is placed at the bottom of the physical Box (Figure 44b). The rotating base allows the physical Box to control its body orientation, as shown in Figure 44c. The servo motor is hidden inside the physical Box to reduce the servo motor noise and avoid creating an attention-drawing feature to the Box.

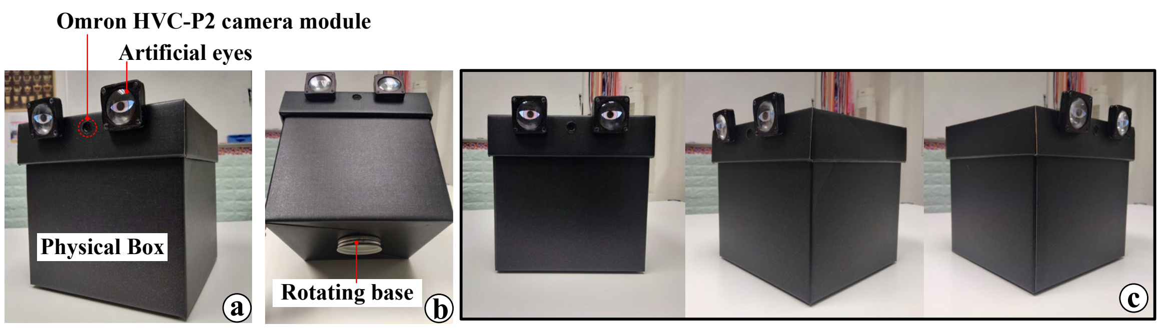


Figure 44 (a) Physical Box mounted with a pair of artificial eyes and Omron HVC-P2 camera module. (b) Rotating base attached at the bottom of the physical Box, (c) to allow the physical Box to orientate itself

5.6.1. Real-environment scenarios synthesis

Based on the pilot study results in Section 5.4, 10 scenarios were designed and evaluated. Each scenario consists of a proactive physical Box performing various *Motion-Gaze* expressions to show its intention in single- and two-user scenarios. In the two-user scenario, each participant will participate with another experimenter that acted as another user. Since the physical Box is simply a shape primitive with no resemblance to human features in its shape, this allows the *Motion-Gaze* expressions to be evaluated independently.

The *Motion-Gaze* expressions, *Turn (T)* and a *Wink (W)* were tested, in both *Single-User (SU)* and *Two-User (TU)* scenarios, where *TU* is considered as minimum engagement for multi-user scenario. Since the multi-user scenario involved three entities in a triadic interaction, the physical Box was acknowledged as the primary entity wanting to initiate an interaction with either the user (first-person) or the experimenter (third-person). In *SU* scenarios, the physical Box expresses its intention to engage with the user. Hence, two possible expressions were tested: 1) $SU(T_1,S_1)$ condition: turning toward the participant without winking, and 2) $SU(T_1,W_1,S_1)$ condition: turning towards the participant and winks. These expressions were followed with a stare (*S*) to maintain the interaction and engagement with the participant. In the *TU* scenario, the physical Box expressed its intention to interact with the participant or the experimenter. In addition to the cases in which physical Box turned to the intended participant directly with a wink condition: $TU_1(T_1,W_1,S_1)$ and $TU_3(T_3,W_3,S_3)$ or without wink conditions - $TU_1(T_1,S_1)$ and $TU_3(T_3,S_3)$, it was also interesting to see how people felt if the object also turned to another participant before turning to the intended participant (the other four cases). Hence, a total of eight conditions of expression were tested. Each expression was followed by a stare *(S)* to maintain the engagement between the physical Box and the participant/experimenter. Table 3 was referred to show the timing diagrams that illustrate the physical Box's behaviors for each condition in Figure 45 and Figure 46. For each condition, the physical Box remains idle for 3 *s*, whereas the turn, wink, and stare behaviors of the physical Box remain for 1 s before the physical Box stops reacting. Also, to justify the differences between eye blinking and eye winking, eye blinking is programmed to blink between 200 $ms \leq$ *eye blinking* < 1 *s* time is programmed for eye blinking and 1 *s* for eye winking.

Table 3. Timing diagram and sequence of behaviors in single- and two-user scenarios.

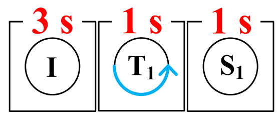
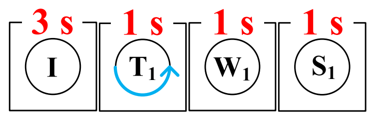
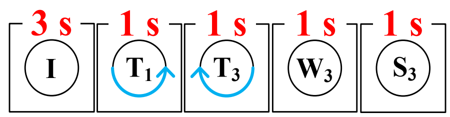
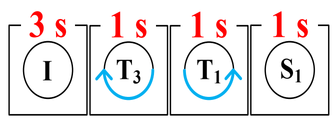
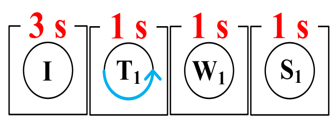
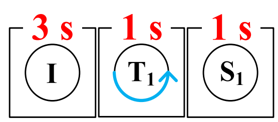
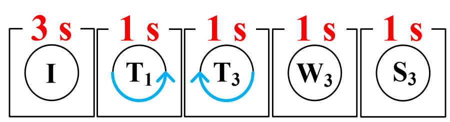
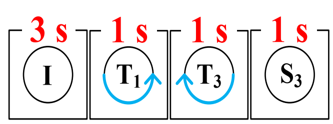
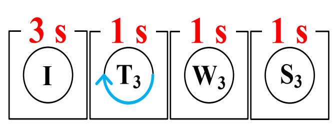
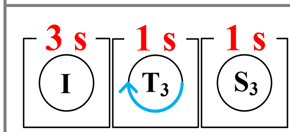
Scenarios	Sequence of behaviors	Conditions
Single-User		*SU(T₁,S₁)*
		*SU(T₁,W₁,S₁)*
Two-Users (Object initiating interaction with the first-person)		*TU₁(T₃,T₁,W₁,S₁)*
		*TU₁(T₃,T₁,S₁)*
		*TU₁(T₁,W₁,S₁)*
		*TU₁(T₁,S₁)*
Two-Users (Object initiating interaction with the third-person)		*TU₃(T₁,T₃,W₃,S₃)*
		*TU₃(T₁,T₃,S₃)*
		*TU₃(T₃,W₃,S₃)*
		*TU₃(T₃,S₃)*

Figure 45 Real-environment user study for Single-User (SU) scenarios

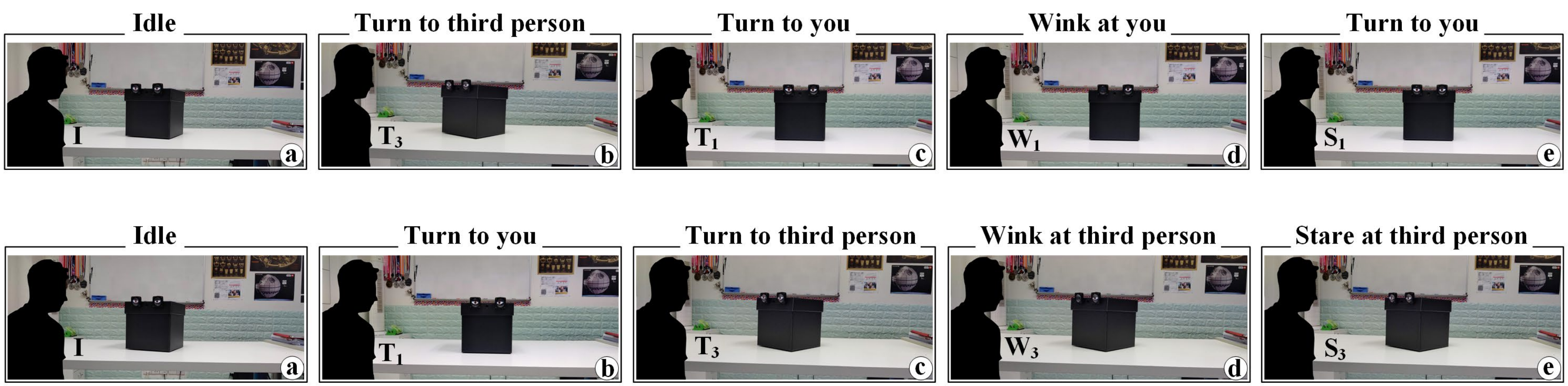


Figure 46 Real-environment user study for Two-User (TU) scenarios. The experimenter (on the left side) as the third-person

5.6.2. Participants, task, and stimuli

A total of 28 participants were recruited to participate in the real-environment user study. The participants were evenly separated into two groups of 14 participants: group A and group B. Initially, group A is involved with the physical Box, initiating interaction with the participants. In contrast, group B is involved with the physical Box initiating interaction with the third-person (experimenter), as presented in Table 4. Of the ten scenarios created according to Table 4, the eight scenarios in the *TU* were split into two sets. One set was related to initiating interaction with the participant *(TU_1),* while the other was related to initiating interaction with the third-person *(TU_3)*. TU_1 and TU_3 conditions were assigned to groups A and B, respectively. To avoid the between-group ordering effects in *SU* and *TU* conditions, seven participants in group A and group B experience the $SU(T_1,S_1)$ condition. In contrast, seven participants in group A and group B experience the $SU(T_1,W_1,S_1)$ as the first condition before proceeding with the other conditions presented in Table 4.

Table 4. Group A and group B conditions distribution.

Scenarios	Group	Conditions
First engagement	A (first half of the participants)	$SU(T_1,S_1)$
	B (first half of the participants)	
	B (second half of the participants)	$SU(T_1,W_1,S_1)$
	A (second half of the participants)	
Two-Users (Object initiating interaction with the first-person)	A	$TU_1(T_3,T_1,W_1,S_1)$
		$TU_1(T_3,T_1,S_1)$
		$TU_1(T_1,W_1,S_1)$
		$TU_1(T_1,S_1)$
Two-Users (Object initiating interaction with the third-person)	B	$TU_3(T_1,T_3,W_3,S_3)$
		$TU_3(T_1,T_3,S_3)$
		$TU_3(T_3,W_3,S_3)$
		$TU_3(T_3,S_3)$

5.6.3. Procedures

Each participant in group A and group B must interact and engage with six conditions. Two conditions are in the *SU* scenario, and four conditions are in the *TU* scenario. Before participants start to interact and engage with the physical Box, a set of instructions were given to the participants that explain the purpose of the study and the meaning of proactive and reactive behavior using examples of preference functions introduced by Lin and Carley (1993). The participants were also instructed to position themselves at a distance between 60 *cm* to 120 *cm* to have an eye-to-eye interaction and engagement with the physical Box. After interacting and engaging with the physical Box, the participants were asked to respond before moving to the next condition.

5.6.4. Measurement

Each participant was given a seven-point Likert scale of proactive-reactive measures, where a score of seven stands for 'very proactive' and one stands for 'very reactive' (see *Appendix C*). Each participant was also required to explain one sentence to each question about the reason for the rating given.

5.6.5. Results

A summary of the real-environment study results is given in Figure 47 and Table 5. In this section, the quantitative results obtained from the proactive-reactive measure are described, and the qualitative findings from the user's explanations are discussed.

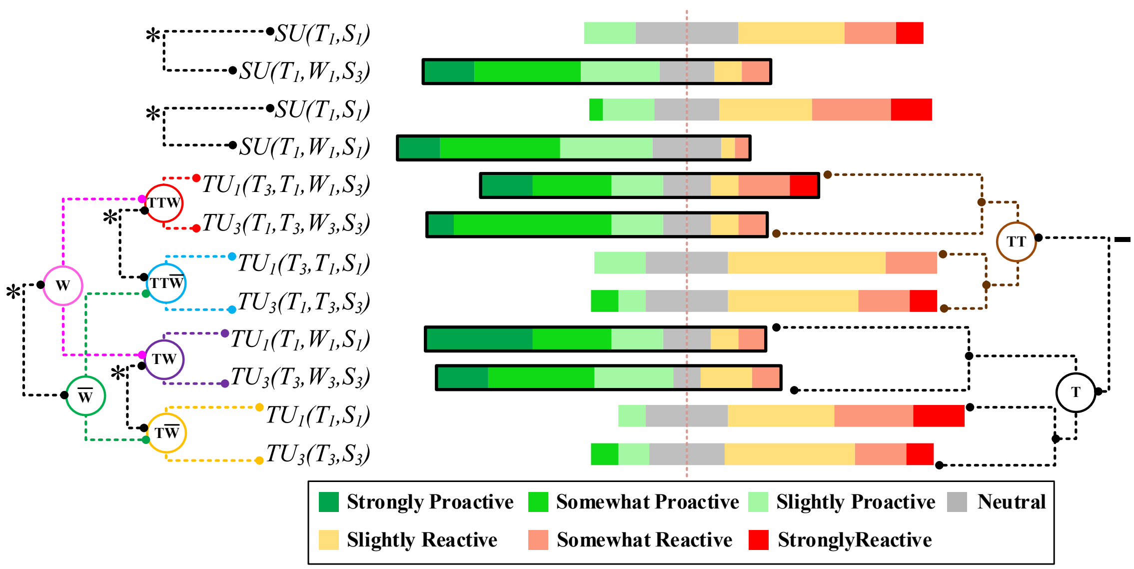


Figure 47 Diverging stacked bar chart of participants' responses to the seven-point Likert proactive-reactive measure. An asterisk highlights the significance of wink-related results (*), and non-significance results are highlighted by a dash (-): significant difference (p<.05), the insignificant difference (p>=.05).

Table 5 The participants' responses to the seven-point scaled Likert proactive-reactive measure in the real-environment user study. Wink-related results are highlighted in bold font.

Impressions	Scenarios	*Mean (M)*	*Median (Mdn)*	*Standard Deviation (SD)*
First engagement	$SU(T_1,S_1)$	3.39	3	1.21
	$SU(T_1,W_1,S_1)$	**5.08**	**5**	**1.28**
Single-user scenarios	$SU(T_1,S_1)$	3.15	3	1.35
	$SU(T_1,W_1,S_1)$	**5.19**	**5**	**1.21**
Two-user scenarios	TTW $TU_1(T_3,T_1,W_1,S_1)$	**5.15**	**6**	**1.68**
	$TU_3(T_1,T_3,W_3,S_3)$	**5.08**	**6**	**1.38**
	TT$\overline{W}$ $TU_1(T_3,T_1,S_1)$	3.38	3	0.92
	$TU_3(T_1,T_3,S_3)$	3.31	3	1.26
	TW $TU_1(T_1,W_1,S_1)$	**5.31**	**6**	**1.65**
	$TU_3(T_3,W_3,S_3)$	**5.00**	**5**	**1.58**
	T$\overline{W}$ $TU_1(T_1,S_1)$	2.85	3	1.21
	$TU_3(T_3,S_3)$	3.31	3	1.32

TTW: Turn Twice and Wink
TT$\overline{W}$: Turn Twice and No Wink
TT: Turn Twice (TTW & TT$\overline{W}$)
T: Turn Once (TW & T$\overline{W}$)
TW: Turn Once and Wink
T$\overline{W}$: Turn Once and No Wink
W: Wink (TTW & TW)
$\overline{W}$: No wink (TT$\overline{W}$ & T$\overline{W}$)

Quantitative Results

Wink is validated as proactive behavior

In the *SU* scenario, the results of the Mann-Whitney U test show that the rating of $SU(T_1,W_1,S_1)$ condition *(Mdn=5, SD=1.21)* was significantly higher $(Z=-0.19, p<0.05)$ than $SU(T_1,S_1)$ condition *(Mdn=3, SD=1.35)*. Figure 47 shows that 74% of the participants in $SU(T_1,W_1,S_1)$ rated winking as proactive behavior, compared to only 19% in $SU(T_1,S_1)$ who rated the physical Box that stared at the participants as being proactive. The interacting and engaging participants rated the first impressions with the physical Box also concurred with the obtained result, which validates the winking expression as proactive behavior. To support the validated first impressions winking expression as proactive behavior, the results of the Mann-Whitney U test shows that the rating of $SU(T_1,W_1,S_1)_{first}$ condition *(Mdn=5, SD=1.28)* was significantly higher $(Z=-2.73, p<0.05)$ that $SU(T_1,S_1)_{first}$ condition *(Mdn=3, SD=1.21)*. Figure 47 shows that 69% of the $SU(T_1,W_1,S_1)_{first}$ condition found the physical Box as proactive when the physical Box winked at the participants. In contrast, in $SU(T_1,S_1)_{first}$ condition, 46% of participants were inclined to rate the physical Box as reactive, and 30% rated the physical Box as neither proactive nor reactive.

In the *TU* scenario, the results of the Mann-Whitney U test show that the rating of (W) conditions (physical Box winked at the participants) (*Mdn=6, SD=1.53*) was significantly higher $(Z=-0.63, p<0.05)$ than ($\overline{W}$) conditions (physical Box stared at the participants) ($\overline{W}$) (*Mdn=3, SD=1.18*). Figure 47 shows that 69% of the participants rated the physical Box with a wink, (W) as proactive behavior (rating scale between 5–7), compared to 19% who perceived winking as a reactive behavior (rating scale between 1–3). In contrast, 63% of the participants experienced a reactive behavior without a wink, while 13% consider the physical Box as proactive behavior. The conducted real-environment user study validates and confirms that the winking expression makes the participants described the physical Box as proactive instead of the physical Box that only stares at the participants.

Making more Turns, validated as neither proactive nor reactive

In the *TU* scenario, the results of a Mann-Whitney U test show no significant difference *(p=0.94)* between the rating of (TT) conditions (turning to each participant) (*Mdn=4, SD=1.60*) and (T) conditions (turning only to the user of interest) (*Mdn=4, SD=1.77*). Figure 47 shows that 42% of the participants (TT) perceived the physical Box as proactive, whereas 40% considered the physical

Box as reactive. While 40% of the participants who experienced the  conditions rated the physical Box as proactive, 42% of the participants rated the physical Box as reactive. The conducted real-environment user study validates and confirms that the number of turns does not influence the participants' score rating to infer the physical Box as proactive or reactive. Therefore, regardless of winking, the physical Box that turns to the engaging participants does not make the physical Box proactive.

First-person and third-person interaction, validated as perceiving similar proactiveness

In the *TU* scenario, the results of the Mann-Whitney U test show no significant difference *($p=0.39$)* between participants in group A (*$Mdn=4$, $SD=1.75$*) and participants in group B (*$Mdn=4$, $SD=1.63$*). Therefore, no effect was found on proactive measures when the participants engage with the Box as first-person or third-person. Figure 47 shows that for the  conditions, 53% of $TU_1(T_3,T_1,W_1,S_1)$ and 69% of the $TU_3(T_1,T_3,W_3,S_3)$ participants reported the physical Box as proactive. In  conditions, 69% of the $TU_1(T_1,W_1,S_1)$ condition and $TU_3(T_3,W_3,S_3)$ condition emphasized the physical Box as being proactive than reactive. With that, the conducted real-environmental user study validates and confirms that the participants perceived winking as proactive behavior regardless of whether the physical Box winked at the first-person or the third-person.

Also,  conditions state that 15% of the $TU(T_3,T_1,S_1)$ condition rated the physical Box as proactive, and 61% of the participants rated the physical Box as reactive. In contrast, in $TU_3(T_1,T_3,S_3)$ condition, 16% of the participants consider the physical Box as proactive, and 61% of the participants considered the physical Box as reactive. The results validate that without winking and making more turns, the participants tend to rate the physical Box as reactive. For  conditions, 69% of the $TU_1(T_1,S_1)$ condition and 61% in $TU_3(T_3,S_3)$ condition recognized staring and turning towards the person of interest as reactive behaviors. Therefore, the results validate that behavior without winking and fewer turns influence the participants' rating towards reactive behavior.

Qualitative Results

First Impressions, Physical Box Turning with a Wink

For participants interacting and engaging in $SU(T_1,W_1,S_1)$ conditions as their initial engagement, 7 out of 14 ranked the physical Box as a proactive object (rating scale between 5–7). Seven participants mentioned that the physical Box could initiate interaction when the physical Box wink at the participants. Therefore, seven reflections after the participants experience the interaction and engagement with the physical Box are presented in the following:

"I immediately felt that Box is trying to engage with me when it winked at me." (P2)
"Box winked at me to show that it wanted to engage with me." (P7)
"I felt confused when Box stared at me after turning towards me, but once it winked at me, I understand that it wanted to interact with me." (P5)
"When Box winked at me, I felt that it wanted to convey something to me." (P6)
"Box winked behavior showed that it wanted to connect with me." (P13)
"The box tried to befriend with me when it winked at me." (P8)
"Winking after turning and staring was an obvious signal of showing that Box tried to engage with me." (P10)

Besides that, two participants reported that the physical Box tried to influence them to collaborate. The following statements are the reflections from the participants:

"It was a signal of showing that the Box wanted me to do something. Maybe the Box needed me to transport it somewhere else." (P1)
"I winked back after the Box winked at me, and I felt that I am engaged with the Box" (P3)

However, two participants ranked the physical Box as a reactive object (rating scale between 1–3). Therefore, the following statements are the reflections from the participants:

"Box reacted by turned and winked at me to acknowledge my presence." (P9)
"Box turned and winked at me because it can detect and reacted to my presence." (P11)

First Impressions, Physical Box Turning without a Wink

For participants interacting and engaging in $SU(T_1,S_1)$ condition as their initial engagement, 2 out of 14 participants ranked the physical Box as proactive (rating scale 5). Therefore, the following statements are the reflections from the participants:

"Box made an effort to turn and looked at me without waiting for me to do anything else yet." (P7)
"Box turned towards me to engage with me, and we made eye contact." (P13)

Nonetheless, 7 out of 14 participants perceived the physical Box as a reactive object (rating scale 1–3). Four participants reported that the physical Box was aware of their presence and reacted by turning towards the participants. Therefore, the following statements are the reflections from the participants:

"Box suddenly reacted by turning and looked at me when I am sitting on the chair" (P2)
"Box engaged with me by turning itself towards me." (P3)
"Box reacted by turning towards me because it wanted to engage with me." (P9)
"Box sensed my presence, and that was why it turned and looked at me." (P6)

Three participants mentioned that the physical Box was acting in response to their presence. Therefore, the following statements are the reflections from the participants:

"Box was programmed to detect my presence by turning towards me." (P4)
"Box was looking somewhere else and reacted to my presence." (P11)
"Box turned and looked at me as a reaction to acknowledge my presence." (P12)

User Experiences Towards Physical Box Different Ways of Turning

Four participants in the TTW conditions and $TT\overline{W}$ conditions pointed out that the physical Box's turning behavior towards the participants was proactive. Therefore, the following statements are the reflections from the participants:

"I appreciated that Box realized the presence of the other person and me by turning towards both of us." (P3)
"Box proactively made an effort to turns toward the other person and then towards me to acknowledge our presence." (P4)
"I thought Box wanted to engage with me, but it turned towards the other person too. That was when I understand that Box wanted to recognize and acknowledged both of us." (P7)
"Box capable of acknowledged where we were by rotating itself towards us." (P8)

Nonetheless, five participants in (TTW) and (TT$\overline{W}$) mentioned that the physical Box was responsive and reactive to their presence. Therefore, the following statements are the reflections from the participants:

"Box turned toward me and to the other person. Seems like Box was responding to our presence." (P2)
"Box was reactive to the presence of both of us by turning its body and made eye contact." (P4)
"Box simply reacted by turning itself towards us to indicate that it knew there were two persons nearby." (P8)
"I felt that Box's behavior of turning was very reactive and not proactive. It reacted by orienting itself towards us." (P10)
"Box seems capable of detecting our presence and knew where to react to acknowledge our presence." (P12)

Whereas six participants in (TW) and (T$\overline{W}$) reported that the physical Box was proactive because it could make its own decisions. Therefore, the following statements are the reflections from the participants:

"Box already knew which participant it should interact with first by stayed engaged with the other person and not me." (P3)
"Box proactively choose to interact with me by turning towards me and not the other person" (P4)
"I was expecting that Box wanted to interact with the other person. But instead, it turned towards me and winked at me, and that was when I know that Box decided to engage with me" (P5)
"Box seems clever to turn itself and winked to whom it wanted to interact." (P9)
"The other person and I did not influence box decision to choose which one it wanted to interact. Box was able to decide on its own." (P10)
"I did not have to behave in certain ways to attract Box to engage with me, but Box picked me instead of the other person, which show that it was intelligent" (P12)

While two participants (T$\overline{W}$) mentioned that the physical Box reacted to them with no apparent action, therefore, the following statements are the reflections from the participants:

"I do not understand why Box turned and stared at me. I was expecting something was about to happen, but there was no proactive action from Box" (P7)

“Box’s behavior seems clueless after turning and staring at me. I felt clueless on how should I behave after it made eye contact with me” (P13)

5.7. Evaluation: The Crowd-Sourced Video-Based User Study

A crowd-sourcing video-based user study was conducted to verify the result gained in the real-environment user study presented in Section 5.6. The crowd-sourced user study is chosen for several reasons. First, the crowd-sourced presents a useful paradigm that enables the experimenters to gain results from a large group in such a way as to maximize cognitive diversity (Oluchukwu *et al.*, 2018)and enhance group performance (Wang, 2017). Second, representation of the future design in the form of video has been proven to be one of the best ways to gather the possible varieties of participants' responses when they have no direct experience about the design (Oogjes and Wakkary, 2017). Finally, the crowd-sourced study has comparable validity to a real-environment user study (Borgo *et al.*, 2018). The results obtained through the crowd-sourced user studies were comparable to laboratory studies (Borgo *et al.*, 2018). This approach's primary assumption is that participants' reactions to videos provide an efficient way to capture how they perceive an actual object. The participants' reaction to videos has also been assumed in other work (Lloyd, 2019; Sturdee *et al.*, 2019). Moreover, video-based user studies allow experimenters to exclude unwanted environmental factors and control the experimental parameters to precisely retain the design object's validity.

5.7.1. Video synthesis

The ten scenarios tested in the real-environment user study were also tested in the crowd-sourced video-based user study. Hence, ten pre-recorded videos were produced for evaluation. Each video consists of a proactive virtual Box mounted with a pair of artificial eyes on the top of it, performing various *Motion-Gaze* animations to show its intention, as shown in Figure 48. The design of the virtual Box is similar to the physical Box used in the real-environment user study. However, the virtual Box's expression is pre-recorded in a series of videos, and therefore, it cannot detect a potential user's presence in its surrounding. The video assumed that the user's eyes were focused on the center of the window; therefore, the virtual Box is located in a simulated three-dimensional space from a first-person perspective. The ground is provided as a reference for the one-dimensional rotation. An abstract figurine was used in the two-user scenarios to represent a third-person.

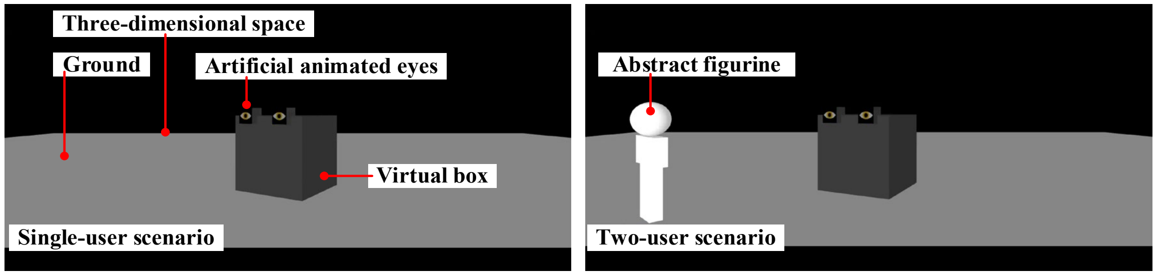


Figure 48 The design of virtual Box's video animation in single- and two-user scenario in three-dimensional virtual space.

5.7.2. Participants, task, stimuli, and measurement

A total of 240 participants were recruited from Amazon Mechanical Turk. The participants were evenly separated into two groups of 120 participants: group A and group B. Initially, group A is involved with the virtual Box initiating interaction with the participants. In contrast, group B is involved with the virtual Box initiating interaction with the third-person (abstract figurine), as presented in Table 4. Of the ten videos created according to Table 6, the eight videos in the *TU* scenario were split into two sets. One set was related to initiating interaction with the participant (TU_1), while the other was related to initiating interaction with the third-person (TU_3). TU_1 and TU_3 conditions were assigned to groups A and B, respectively. The 2 (*SU*) × 4 (*TU*) × two groups = 16 conditions were tested and counterbalanced using Latin Square to eliminate the between-group ordering effects in SU and TU scenarios. For instance, 120 (out of 240) participants watched $SU(T_1,S_1)$ condition as the first video and the other 120 watched $SU(T_1,W_1,S_1)$ condition as the first video. The crowd-sourced user study adapted a similar task, stimuli, and measurement in the real-environment user study. Figure 49 and Figure 50 illustrate a proactive virtual Box interacting in single-user (SU) and two-user (TU) scenarios.

Table 6 Group A and group B conditions distribution

Scenarios	**Group**	**Conditions**	**Video links**
First engagement	A (first half of the participants)	$SU(T_1,S_1)$	https://youtu.be/95MjWrO-Hho
	B (first half of the participants)		
	B (second half of the participants)	$SU(T_1,W_1,S_1)$	https://youtu.be/J25vEjk7m5c
	A (second half of the participants)		
Two-Users (Object initiating interaction with the first-person)	A	$TU_1(T_3,T_1,W_1,S_1)$	https://youtu.be/EPbP2QP9aIc
		$TU_1(T_3,T_1,S_1)$	https://youtu.be/nqsQOouDOZ0
		$TU_1(T_1,W_1,S_1)$	https://youtu.be/J25vEjk7m5c
		$TU_1(T_1,S_1)$	https://youtu.be/dFoCnPkzSZ8
Two-Users (Object initiating interaction with the third-person)	B	$TU_3(T_1,T_3,W_3,S_3)$	https://youtu.be/ev3SlSHeNz4
		$TU_3(T_1,T_3,S_3)$	https://youtu.be/JoLXOwcFsdI
		$TU_3(T_3,W_3,S_3)$	https://youtu.be/s6OYFw3yS0o
		$TU_3(T_3,S_3)$	https://youtu.be/bmhnW_OxmDc

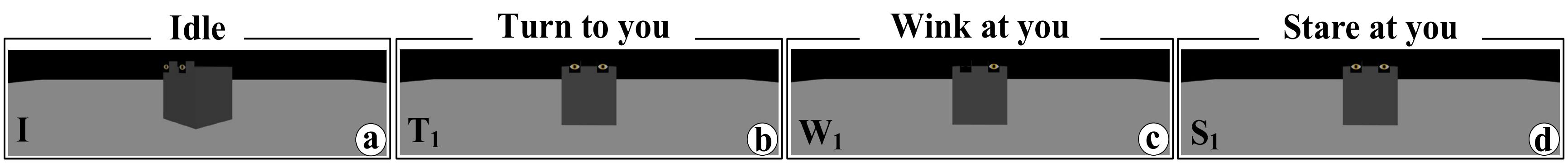


Figure 49 Video-based user study for Single-User (SU) scenarios.

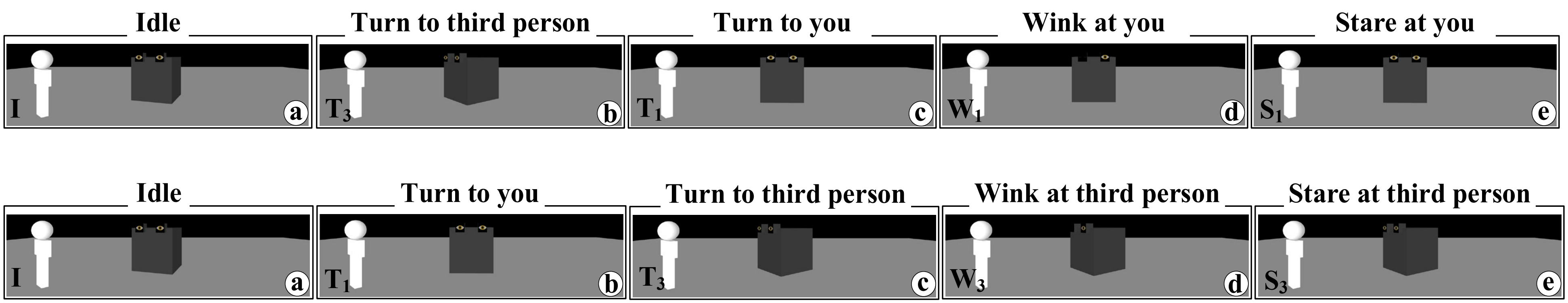


Figure 50 Video-based user study for Two-User (TU) scenarios. Abstract figurine (on the left side) acts as the third-person

5.7.3. Procedures

Each participant in group A and group B, must-watch six videos. Two videos in the *SU* scenario and four videos in the *TU* scenario. Each participant first watched the two videos in the *SU* scenario and then watched a set of four videos in the *TU* scenario. Before participants started to observe the videos, they read a set of instructions that introduced the study's purpose and explained the meaning of proactive and reactive behavior using examples of preference functions (Lin and Carley, 1993). The instructions also mentioned where the user is the first-person, the abstract white figurine as the third-person. The virtual Box is introduced as the interactor who wants to interact with the first- or the third-person. All videos have a video cover image of the virtual Box staring at the participant (Figure 51). Before pressing the play button, participants were instructed to position themselves at a distance to engage in eye-to-eye interaction with the virtual Box. After watching each video condition, the participants were asked to respond before moving to the next video condition. After approval, compensation of USD 2.50 was given to each participant who submitted a completed survey.

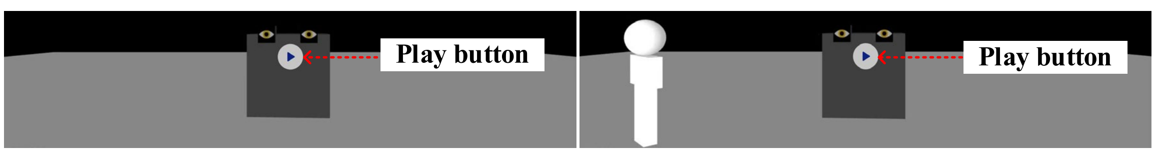


Figure 51 Video cover image in single-user (left) and two-user (right) scenarios.

5.7.4. Results

A summary of the crowd-sourced study results is given in Table 7 and Figure 52. In this section, the quantitative results obtained from the proactive-reactive measure are described, and the qualitative findings from the user's explanations are discussed.

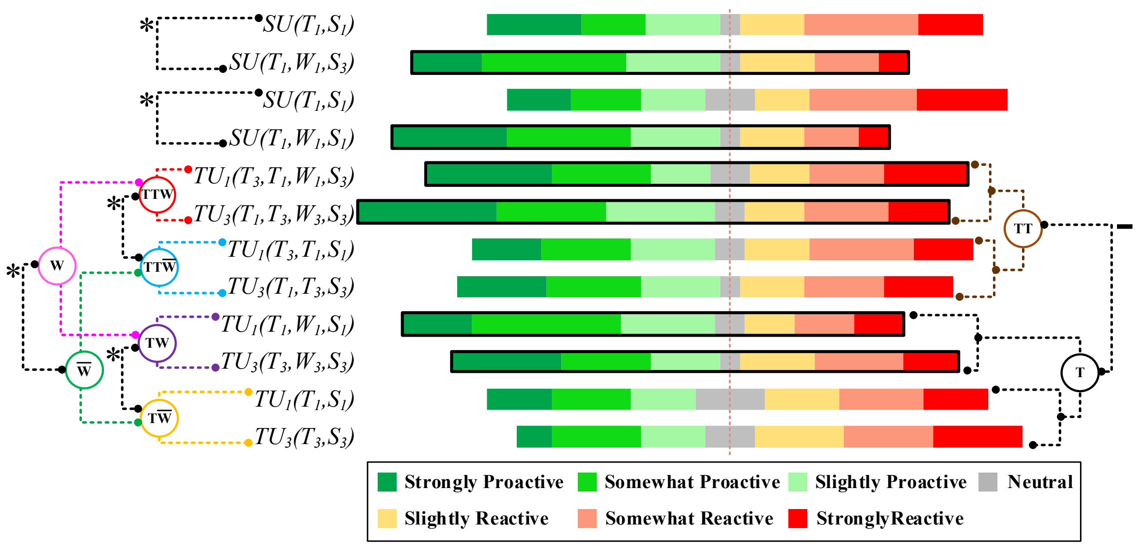


Figure 52 Diverging stacked bar chart of participants' responses to the seven-point Likert proactive-reactive measure. An asterisk highlights the significance of wink-related results (*), and non-significance results are highlighted by a dash (-): significant difference ($p<.05$), the insignificant difference ($p>=.05$).

Table 7 The participants' responses to the seven-point scaled Likert proactive-reactive measure in the crowd-sourced study. Wink-related results are highlighted in bold font.

Impressions	Scenarios	*Mean (M)*	*Median (Mdn)*	*Standard Deviation (SD)*
First engagement	$SU(T_1,S_1)$	4.03	4	2.15
	$SU(T_1,W_1,S_1)$	**4.63**	**5**	**1.86**
Single-user scenarios	$SU(T_1,S_1)$	3.90	3.5	2.08
	$SU(T_1,W_1,S_1)$	**4.71**	**5**	**1.91**
Two-user scenarios (TTW)	$TU_1(T_3,T_1,W_1,S_1)$	**4.57**	**5**	**2.07**
	$TU_3(T_1,T_3,W_3,S_3)$	**4.49**	**5**	**2.20**
($TT\overline{W}$)	$TU_1(T_3,T_1,S_1)$	4.19	5	2.13
	$TU_3(T_1,T_3,S_3)$	4.04	4	2.05
(TW)	$TU_1(T_1,W_1,S_1)$	**4.32**	**5**	**2.13**
	$TU_3(T_3,W_3,S_3)$	**4.58**	**5**	**1.96**
($T\overline{W}$)	$TU_1(T_1,S_1)$	3.62	3	1.95
	$TU_3(T_3,S_3)$	3.98	4	1.98

TTW: Turn Twice and Wink
$TT\overline{W}$: Turn Twice and No Wink
TT: Turn Twice (TTW & $TT\overline{W}$)
T: Turn Once (TW & $T\overline{W}$)
TW: Turn Once and Wink
$T\overline{W}$: Turn Once and No Wink
W: Wink (TTW & TW)
$\overline{W}$: No wink ($TT\overline{W}$ & $T\overline{W}$)

Quantitative results

Wink is generally considered proactive

In the *SU* scenario, the Shapiro-Wilk test results indicated that the ratings' distribution was not statistically normal in both videos (p>0.05). The results of the Mann-Whitney U test showed that the rating of $SU(T_1,W_1,S_1)$ condition *(Mdn=5, SD=1.91)* was significantly higher *(Z=–4.13, p<0.05)* than $SU(T_1,S_1)$ condition *(Mdn=3.5, SD=2.15)*. Figure 52 shows that 64% of the participants in the $SU(T_1,W_1,S_1)$ condition rated winking as a proactive behavior, compared to only 39% in $SU(T_1,S_1)$ condition who rated the virtual Box that stared at them as being proactive. The answers for the first video regarding the first impressions of the participants also concurred with this result. The results of the Mann-Whitney U test showed that the participants who watched $SU(T_1,W_1,S_3)$ condition *(Mdn=5, SD=1.86)* first also rated it significantly higher *(Z=–2.07, p<0.05)* than those who started from $SU(T_1,S_1)$ condition *(Mdn=4, SD=2.15)*. Figure 52 shows that 63% of the participants in the SU(T1,W1,S3) condition found the virtual Box to be proactive when it winked at them for the first impression. In contrast, in $SU(T_1,S_1)$ condition, the participants were equally inclined to rate the virtual Box as proactive or reactive (48% for each).

In the *TU* scenario, the Shapiro-Wilk test results indicated that the ratings' distribution was not statistically normal in all eight videos (p<0.05). The results of the Mann-Whitney U test showed that the rating of all videos condition with a wink (W) (*Mdn=5, SD=2.08*) was significantly higher *(Z=4.02, p<0.05)* than all videos condition without a wink ($\overline{W}$) (*Mdn=4, SD=2.03*). Based on the Likert scale shown in Figure 52, 60% of the participants rated the winking behavior (W) as slightly too strongly proactive than 35% who perceived winking as a reactive behavior (rating scale between 1–3). Without a wink ($\overline{W}$), 48% of the participants experienced staring as reactive behavior, while 44% considered the virtual box proactive. The results suggest that a proactive virtual Box that winks makes the user see it as more proactive than the virtual Box that does not wink.

Making more turns does not affect the perception

In the *TU* scenario, the results of the Mann-Whitney U test show no significant difference *(p=0.12)* between turning to each participant (TT) (*Mdn=5, SD=2.12*) and turning only to the user of interest (T) (*Mdn=4, SD=2.03*). Figure 52 reveals that 54% of the participants perceived the virtual Box that acknowledged both participants as proactive, whereas 41% of the participants considered making more turns to be reactive behaviors. 49% of the participants who experienced fewer turns

(T) reported the virtual Box as proactive, while 43% reported the virtual Box as reactive. In summary, the results suggest that making fewer or more turns does not influence the participants' rating score to infer the virtual Box as more proactive or reactive. Therefore, regardless of winking, making one more turn does not make the virtual Box appear more proactive. Considering making more turns in multi-user scenarios could be time-consuming, directly turn to the user of interest is more time-saving.

First-person and third-person interaction perceived similar proactiveness

In the *TU* scenario, the results of the Mann-Whitney U test show no significant difference *($p=0.49$)* between participants in group A (*$Mdn=5$, $SD=2.06$*) and participants in group B (*$Mdn=5$, $SD=2.10$*). Therefore, in the *TU* scenario, regardless of who is the user of interest, there is no effect on the virtual Box's perceived proactiveness.

For the (TTW) conditions, the diverging stacked bar in Figure 52 shows that 57% of the $TU_1(T_3,T_1,W_1,S_1)$ condition reported virtual Box as proactive. The same goes for the participants in the $TU_3(T_1,T_3,W_3,S_3)$ condition; 60% of them rated the virtual Box as proactive. In (TW) scenarios, 63% of the $TU_1(T_1,W_1,S_1)$ condition and 53% of the $TU_3(T_3,W_3,S_3)$ condition emphasized the virtual Box as more proactive than reactive. Therefore, these results suggested that regardless of virtual Box winked at the participants or the third-person, the participants perceived winking as proactive behavior.

For ($TT\overline{W}$) conditions, 48% of the $TU_1(T_3,T_1,S_1)$ condition rated the virtual Box as proactive, and 46% of the participants rated the virtual Box as reactive. In contrast, in $TU_3(T_1,T_3,S_3)$, 53% of the participants considered the virtual Box to be proactive, and 43% of the participants considered the virtual Box as reactive. The results show that participants categorize the virtual Box as either proactive or reactive without winking and making more turns. Therefore, no significant differences were found between participants in group A and group B. For ($T\overline{W}$) scenarios, 64% of the $TU_1(T_1,S_1)$ condition and 53% of the $TU_3(T_3,S_3)$ condition recognized staring and turning as reactive behaviors. The results show that behavior without winking and less turning influenced the participants to rate the virtual Box as more reactive.

Qualitative results

First impressions, Box turning with a wink

For participants experiencing $SU(T_1,W_1,S_1)$ condition as their starting point, 75 out of 120 ranked the virtual Box as a proactive object (rating scale between 5–7). Thirty-three participants mentioned that the virtual Box was capable of initiating interaction when it winked at them. Examples are illustrated below.

"Box turned in my direction and tried to communicate by winking at me." (P16)
"Box turned and winked at me. I felt like Box initiated a contact." (P99)

Seven participants reported that the virtual Box tried to influence them to collaborate. The following statements show the participants' reflection.

"I was surprised Box looked at me and then winked, so I winked back." (P35)
"When Box winked, I believe it was trying to invite me to do something." (P79)
"Box noticed me, and we interacted by winking. It tried to get me involved in something interesting." (P109)

Five participants realized that they needed to cooperate with the virtual Box. Example statements are presented below.

"Box winked, trying to get my attention and ready for a command from me." (P3)
"When Box winked, it affected me, and I felt that I needed to do something to respond." (P141)

However, 40 participants perceived the virtual Box as a reactive object (rating scale between 1–3). Twenty-seven participants mentioned winking as reactive behavior. Example statements are given below.

"Box's movements felt orchestrated, and I thought that it only winked at me because it was told to." (P6)
"Box seems only to react when I am making eye contact with him." (P81)

First impressions, Box turning without a wink

Of the participants experiencing $SU(T_1,S_1)$ condition as their starting point, 57 (out of 120) participants ranked the virtual Box as being proactive (rating scale between 5–7). Thirty-four participants mentioned that the virtual Box took the initiative and made eye contact, meaning that it was proactive. The following statements show the participants' reflection.

"Box turned toward me, and we made eye contact." (P51)
"Box engaged with me first and looked at me." (P125)

Nonetheless, 58 (out of 120) participants perceived the virtual Box as a reactive object (rating scale 1–3). 33 (out of 120) participants reported that the virtual Box was aware of their presence and reacted by turning towards them. Examples are illustrated below.

"Box seemed more reactive as though it became aware of my presence and then looked directly at me." (P18)
"Box felt very reactive. It sensed my presence and turned to look at me." (P48)

From the quotations below, nine (out of 120) participants stated that the virtual Box was acting according to their presence. Example statements are presented below.

"Box was being reactive as if it was waiting for me to act first before doing its next action." (P42)
"Box did not do much, but it did seem at least slightly interested in what I needed to do first." (P50)

User experiences towards Box different ways of turning

Thirty-seven (out of 60) participants in the TTW and TT$\overline{W}$ conditions pointed out that acknowledgment was proactive. Examples are illustrated below.

"Box displayed proactive action because it acknowledged both our presences." (P23)
"Box considered both of us by acknowledging us." (P7)

Also, 21 (out of 60) participants in TTW and TT$\overline{W}$ conditions reported that the virtual Box showed proactive behavior because it gave equal attention to both of them. The following statements show the participants' reflection.

"I appreciate that Box gave its attention to both of us." (P9)
"Box seemed to be balancing out engaging both of us." (P165)

Nonetheless, 82 participants in TTW and TT$\overline{W}$ conditions mentioned that since the virtual Box was responsive to their presence and incapable of making its own decisions, it was reactive. Examples are reflected in the following statements.

"Box responded to our presence but looked confused as to which one it should interact with." (P11)

"Box made a clear effort to observe both of us, but I think it was unable to decide." (P159)

11 (out of 30) participants in $T\overline{W}$ conditions speculated that the Box was proactive because it showed the intention to receive a reply from them. Examples are illustrated below.

"Box acknowledged me; I assume it was waiting for me to do something." (P17)
"Box turned its body to square up with my gaze. I suppose Box was waiting for me to make a move." (P73).

Whereas 44 (out of 60) participants in TW and $T\overline{W}$ conditions reported that the virtual Box was proactive because it could make its own decisions independently. The following statements show the participants' reflection.

"Box seemed more active and interested in interacting with me compared to the other person." (P50)
"Box was proactive by aiming its attention at the other person and ignored me." (P198)

Nonetheless, 24 (out of 60) participants in TW and $T\overline{W}$ conditions stated that since the virtual Box was unable to notice all the observers' presence, it was reactive. Example statements are given below.

"Box only ever turned towards the other person and never myself." (P144)
"Box was reactive because it did not acknowledge the other person and was staring blankly at myself." (P81)

28 (out of 60) participants in $T\overline{W}$ conditions mentioned that the virtual Box reacted to them with no explicit action. The following statements show the participants' reflection.

"Box turned to me without giving any clue as to why it was looking at me." (P19)
"Box only looked at the other person but did not act towards him or engage." (P168)

5.8. Discussion

The Real-Environment User Study versus The Crowd-sourced Video-Based User Study

This section discusses the overall findings of the proactive-reactive behavioral measures of an object mounted with artificial eyes and abstract motion based on the real-environment user study and the crowd-sourced video-based user study.

In the real-environment user study, a physical Box mounted with a pair of artificial eyes and a camera module on top of it was developed. The physical Box is also attached with a micro servo motor at the bottom of it, giving the Box the capability to turn towards the user of interest by detecting the user's presence using the mounted camera module. The physical performed various *Motion-Gaze* expression. Each participant was required to rate the physical Box using a seven-point Likert scale of proactive-reactive measures.

The crowd-sourced video-based user study was conducted to verify the result gained in the real-environment user study. In the crowd-sourced video-based user study, participants were required to watch and evaluate a series of pre-recorded videos which consists of a virtual Box mounted with a pair of artificial eyes on top of it. The virtual Box representing the physical Box used in the real-environment user study is also designed to turn towards the user of interest. The pre-recorded videos contain a virtual Box performing various *Motion-Gaze* expressions. Each participant was required to rate the virtual Box after watching each pre-recorded video using a seven-point Likert scale of proactive-reactive measures.

The overall results of the real-environment user study and the crowd-sourced video-based user study are presented in Table 8. Table 8 shows the mean and median finding for the first impression, single-user scenarios, and two-user scenarios, which shows no significant difference ($p_M=0.82$, $p_{Mdn}=0.64$) when the scenarios are compared in both user studies. Therefore, the correlation coefficient, *r*, shows a positive relationship ($0.75 \leq r \leq 0.92$) between the real-environment user study and the crowd-sourced video-based user study.

Although more comprehensive variability sources were found on the standard deviation ($p_{SD}<0.05$, $r=0.091$) between the real-environment user study and the crowd-sourced video-based user study, the variability of the standard deviation did not impact the nature of the object's proactive-reactive behaviors. The result can be verified by observing the mean and median findings in Table 8, reflecting the object's winking expression as proactive initiative-taking towards the interacting and engaging user. To further justify the 240-participant rated the object's winking expression as proactive during the crowd-sourced video-based user study, a random data of 28 participants that rated the object's winking expression as proactive has been selected. The random data of 28-participant from the crowd-sourced video-based user study is selected to compare and further validate the correlation coefficient and significant findings with the 28-participant in the real-environment user study. The results of the 28-participant from the crowd-sourced video-based user study show no significant difference ($p_M=0.84$, $p_{MDN}=0.85$, $p_{SD}=0.83$) with the 28-participant in the real-environment user study. The correlation coefficient also shows a positive relationship *($0.77 \leq r \leq 0.90$)* between the 28 randomly selected participants from the crowd-sourced video-based user study and the 28 participants in the real-environment user study. Therefore, the presented results show that winking is a useful expression that makes the user feel that the object is proactive and can encourage them to take reciprocal action to engage in the interaction.

Table 8 Comparison between the crowd-sourced video-based user study and the real-environment user study. Wink-related results are highlighted in bold font.

Impressions	**Scenarios**	***Mean (M)***			***Median (Mdn)***			***Standard Deviation (SD)***		
		User study			**User study**			**User study**		
		Crowd-sourced *(N=240)*	**Real-environment *(N=28)***	***Random* Crowd-sourced *(N=28)***	**Crowd-sourced *(N=240)***	**Real-environment *(N=28)***	***Random* Crowd-sourced *(N=28)***	**Crowd-sourced *(N=240)***	**Real-environment *(N=28)***	***Random* Crowd-sourced *(N=28)***
First engagement	$SU(T_1,S_1)$	4.03	3.39	2.92	4	3	3	2.15	1.21	1.33
	$SU(T_1,W_1,S_1)$	**4.63**	**5.08**	**4.85**	**5**	**5**	**5**	**1.86**	**1.28**	**1.35**
Single-user scenarios	$SU(T_1,S_1)$	3.90	3.15	3.73	3.5	3	4	2.08	1.35	1.40
	$SU(T_1,W_1,S_1)$	**4.71**	**5.19**	**4.92**	**5**	**5**	**5**	**1.91**	**1.21**	**1.30**
Two-user scenarios	$TU_1(T_3,T_1,W_1,S_1)$	**4.57**	**5.15**	**4.85**	**5**	**6**	**5**	**2.07**	**1.68**	**1.75**
	$TU_3(T_1,T_3,W_3,S_3)$	**4.49**	**5.08**	**4.62**	**5**	**6**	**5**	**2.20**	**1.38**	**1.44**
	$TU_1(T_3,T_1,S_1)$	4.19	3.38	3.92	5	3	4	2.13	0.92	1.07
	$TU_3(T_1,T_3,S_3)$	4.04	3.31	3.38	4	3	3	2.05	1.26	1.27
	$TU_1(T_1,W_1,S_1)$	**4.32**	**5.31**	**4.62**	**5**	**6**	**5**	**2.13**	**1.65**	**1.55**
	$TU_3(T_3,W_3,S_3)$	**4.58**	**5.00**	**4.69**	**5**	**5**	**5**	**1.96**	**1.58**	**1.49**
	$TU_1(T_1,S_1)$	3.62	2.85	4.23	3	3	4	1.95	1.21	1.12
	$TU_3(T_3,S_3)$	3.98	3.31	4.31	4	3	4	1.98	1.32	1.20

Mean: $r=0.92$, $p=0.82$ (Crowd-sourced vs. Real-environment); $r=0.77$, $p=0.84$ (Real-environment vs. Random Crowd-sourced)

Median: $r=0.75$, $p=0.64$ (Crowd-sourced vs. Real-environment); $r=0.86$, $p=0.85$ (Real-environment vs. Random Crowd-sourced)

Standard Deviation: $r=0.091$, $p<0.05$ (Crowd-sourced vs. Real-environment); $r=0.90$, $p=0.83$ (Real-environment vs. Random Crowd-sourced)

5.9. Conclusion

A proactive object mounted with artificial eyes on top and abstract motion has been tested and validated through the real-environment and the crowd-sourced video-based user studies. Both user studies show that the artificial eyes winking expression has successfully convinced the participants to rate the object as proactive. Furthermore, winking expression influenced the participant to take reciprocal action, enabling them to experience back-and-forth interaction (i.e., bidirectional interaction). However, without winking expression, the crowd-sourced user study participants were inclined to rate the object as neither proactive nor reactive. In contrast, the participants in the real-environment user study tend to rate the object as reactive. Nevertheless, the object that only stares at the user was not considered as proactive.

Turning to acknowledge the users in the environment should make the object proactive. However, the participants in the real-environment user study found the object's turning behavior as being reactive. In contrast, the crowd-sourced user study participants were inclined to rate the object as neither proactive nor reactive. The results suggest that turning behavior does not influence the participants to rate the object as more proactive. Therefore, regardless of winking, making more turns does not make the object appear more proactive. Making more turns in multi-user scenarios also prolonged the interaction; directly turn to the user of interest is more effective and time-saving. For instance, in a multi-user scenario, the object can straight away turn to the intended user and winked to initiate the interaction, and this minimizes the interaction design. Hence, making more turns as a proactive gesture to acknowledge the users in multi-user scenarios can be excluded from the interaction design.

Chapter 6.

Discussions and Conclusions

This chapter provides an overall discussion regarding the bidirectional and proactive human-object interaction design based on the perceptual crossing paradigm. This chapter also briefly presents the limitations that can be considered to enhance this research work, explain future work possibilities, and summarise the thesis's contributions.

6.1. Answer to the Research Question 1

"How can an object proactively initiate an interaction and maintain a stable and continuous bidirectional engagement with an intentional person?"

Based on the perceptual crossing paradigm, bidirectional and proactive communication is essential for distinguishing whether the interaction is with an intentional entity or a reactive entity. Therefore, to exploit the perceptual crossing paradigm into the human-object interaction design practice, detecting the person's visual presence with attention is an effective medium for an object to:

- differentiate an intentional person with an unintentional entity (e.g., passerby, other objects), and
- establish the bidirectional interaction with the intentional person.

In Study 1 (Chapter 3), a prototype integrated with 1) an eye tracker to detect the participants' visual presence with attention and 2) abstract motion to allow the object to proactively show its intention to interact is developed, tested, and validated. The results show that the object's abstract motion successfully attracted the participants to interact and engage. However, the medium of visual attention as the interaction method shows the participants did not achieve bidirectional interaction with the object. Based on the perceptual crossing paradigm, even if the communication channel is reduced to a bare minimum, two intentional entities can still recognize each other as long as they experienced the same perceptual environment. Therefore, with visual attention as the only perceptual quality used to interact with the object (i.e., eyes), the object should also show the same perceptual quality to enable the interaction to occur in the same perceptual environment. Therefore, to let the participants realize the bidirectional interaction, expressive and recognizable visual feedback similar to that of human eye contact (i.e., a pair of artificial eyes) is introduced.

In Study 2 (Chapter 4), an object prototype mounted with artificial eyes and a camera module for detecting the user's visual attention is developed, tested, and validated. The results show that the object's artificial eyes that stare and blink have successfully attracted the participants to achieve bidirectional interaction. However, to maintain a continuous bidirectional interaction, the object's artificial eyes need to convey its proactive behavior besides staring and blinking. Therefore, the artificial eyes are proposed with proactive expressions such as winking and pupil dilation to allow the participant to maintain continuous bidirectional interaction with the object.

In Study 3 (Chapter 5), an object prototype mounted with artificial eyes that display proactive expression (i.e., winking) is developed, tested, and validated. The results show that the object's artificial eyes winking expression has successfully made the object proactively show its intention to interact with the participant. The object's artificial eyes winking expression also influenced the participants to take reciprocal action (e.g., winked back), enabling them to maintain a stable and continuous bidirectional interaction and engagement.

With the interaction method reduced to a bare minimum, based on the perceptual crossing paradigm and the presented results in Study 1, Study 2, and Study 3, for an object to proactively initiate interaction and maintain a stable and continuous bidirectional engagement with a person, the object needs to it

With unimodal communication cue, for an object to show its intention to interact with a user based on the perceptual crossing paradigm, the object needs to:

1) differentiate an intentional user from a potential user (e.g., nearby user) or an unintentional entity (e.g., passerby, other objects).

2) visibly augment its perceptual quality that reflects the chosen communication cue to enable the user to experience the interaction in the same perceptual environment.

3) express proactive intention through the augmented perceptual quality to invite the user to reciprocate the object's intention using the same perceptual quality.

6.2. Answer to the Research Question 2

"What is the signaling protocol that can be implemented for an object to proactively initiate an interaction and maintain a continuous bidirectional engagement with an intentional person?"

In Study 3 (Chapter 5), a conceptual model called Session Initiation for Proactive Object (SIPO) is introduced (Figure 34a and Figure 36), implemented, tested, and validated to improve the object's proactive behavior towards people in its surrounding. The results show that the implemented SIPO conceptual model successfully achieved a proactive and continuous bidirectional interaction and engagement. Based on the perceptual crossing paradigm, the SIPO conceptual model outline a protocol of interaction that allows an object to proactively initiate interaction and maintain a continuous bidirectional engagement with a person. The SIPO conceptual model protocol consists of four state conditions, which are 1) exploring, 2) expressing and 3) engaging, and 4) terminating state condition. In the 1) exploring state condition, a proactive object explores the environment to search for a user of interest. This state condition is required to allow the proactive object to differentiate an intentional entity (user of interest) from a reactive entity (e.g., passerby, other objects) and to identify a potential user of interest within its environment. After the proactive object explores and identifies the user of interest, the state condition changed to the 2) expressing state condition. This state condition is required to allow the proactive object to express its intention to initiate an interaction and engage with the user of interest. Once the intention is initiated towards the person, the state condition changed to the 3) engaging state condition. This state condition is required to allow the user to reciprocate by responding to the proactive object's action with another equivalent action within a time threshold. The perceptual crossing is established once the proactive object receives a reciprocating action from the intentional person. A continuous bidirectional engagement and communication session between the proactive object and the user is maintained. Based on the specified time threshold, the proactive object state condition changed to the 4) terminating state condition when the user of interest ignores the proactive object's intention to interact during the 2) expressing state condition and exceeded the time threshold. The terminating state condition is also executed when the engaging user decided to terminate the ongoing continuous bidirectional engagement with the proactive object during the 3) engaging state condition. After a session is terminated, the proactive object explores the environment to search for another potential user.

6.3. Discussion

6.3.1. Visibility of the perceptual quality

A pair of artificial eyes that allow the user to interpret the intention to interact has been investigated. The results identified three expressions, blinking, winking, and staring, as the key elements to enable the design of a proactive object. With the motion-enabled platform, gaze direction is no longer necessary. Two OLED screens can be replaced by a pair of sparse yet expressive (e.g., 5×5) LED matrices that show a symbolic representation of the eyes. Ultimately, two LED bulbs with a careful design of the lighting behavior (Song and Yamada, 2018) could work too. Still, users may not treat such an oversimplification as expressive perceptual quality, thus disabling the perceptual crossing. In all, the design space between expressiveness and simplicity should be further investigated.

A proactive object can also be augmented by other expressive perceptual quality modalities, such as a pair of ears (Nanayakkara *et al.*, 2018). For example, an object with ears can turn its body orientation to show attention to the user and wiggle its ear to invite the users to provide reciprocal input, such as talking to it. Such a design could make voice user interfaces such as Amazon Echo both more proactive and discreet (Torta *et al.*, 2014), as it would not need to make sounds to attract users.

6.3.2. Expressing intention for initiating interaction

Winking is a subtle form of visible communication used to make a silent agreement between two subjects (Goldstein, 2019; Kowalczyk and Sawicki, 2019). More importantly, winking is the most straightforward action that can be realized using a simple infrastructure, i.e., an add-on LED matrix. This work confirms the effectiveness of applying winking in proactive object design. Other plausible gestures for expressing confirmation or agreement, such as head nodding or hand gestures, could also be considered, but these designs require more mechanical joints or an extra display. Different forms of confirmation, for example, providing GUIs such as icons or progress bars on the visual display, can also express confirmation, but this introduces another layer of information that may break natural eye-to-eye contact and gaze engagement. Hence, this thesis stays in the context of eye expressions without adding further complexity.

This thesis design exploration uses one degree-of-freedom abstract motion to express intention due to its practicality and simplicity. It can be implemented by simple mechanisms driven by a single motor. It is a simple yet effective way to express intention based on the coordination of two persons' movement patterns (Sun *et al.*, 2019). Although an object's expressiveness can be enriched by increasing the number of degrees of freedom, such as adding more joints, the object might be perceived more like a robot. This research attempts to present design guidelines that are generalizable to object design. Even an abstract primitive (such as Box) may be designed to take the initiative in an engagement.

It should be noted that the abstract motion platform was based on a servo motor and a noisy stepper motor that could be unpleasant for people working in a quiet environment. Silent actuation methods such as joule-heating actuators (Wang *et al.*, 2018) and shape memory alloy (Oh *et al.*, 2017) are possible solutions. Still, the low torque could mean that the applications are limited to lightweight objects.

6.3.3. The appropriateness of winking

Winking may be perceived as a controversial behavior and can convey various messages that might be either harmless or offensive (Ren and Zhi-peng, 2014). For example, embedding winking eyes into a gender-specific object might create sexual implications (e.g., flirting). Winking eyes on a data-sensitive object might create an inappropriate expression of secret-sharing or agreement-making, as winking is often interpreted as meaningful behavior by a receiver (Clodic *et al.,* 2017). The meaning behind winking can also vary depending on the situation, culture, and gender differences. For example, winking, especially coming from the opposite sex, is perceived as an impolite gesture in many Asian countries (Goldstein, 2019). Still, the results may vary if a non-gendered object winks as a way of establishing engagement. Therefore, this work deliberately designed artificial eyes with a gender-neutral appearance. In this way, the eyes can be easily mounted on any object that can be considered gender-neutral in most cases. The crowd-sourced user study and the real-environment user study evaluated the perceived proactiveness in a gender-neutral setting (i.e., physical/virtual Box and an abstract figurine). The setting is created to prevent affecting the participants' emotions, which may influence their ratings of the perceived scenarios.

6.3.4. Communication session during eye engagement

This study's main focus was to enable an object to initiate eye-contact engagement in perceptual crossing to experience bidirectional and proactive interaction. By referring to the proposed SIPO model (Figure 34a), once the eye-contact engagement is detected, feedback (e.g., pupil dilation) should be provided to confirm the engagement. After that, during the human-object communication session, natural interactions such as bodily gestures, facial expressions, or voice responses could enhance the embodiment while maintaining the eye-to-eye contact interaction. For example, a simple nod from the user to show an agreement while conversing with each other. The SIPO session initiation technique provides a solid foundation for these interaction schemes.

6.3.5. Scalability of the interaction model

The viability of the SIPO interaction model was confirmed using two-user scenarios. Based on the model, proactive objects were designed for initiative-taking, and both users in the environment could recognize the intention and look back to establish the perceptual crossing. This model extends the traditional perceptual crossing paradigm from a dyadic relationship to a triadic one and can be extended to a multi-triadic scenario (Figure 53a) where a single object interacts with many users. However, it is insufficient to further extend the perceptual crossing models for fully connected many objects, many user scenarios (Figure 53b) due to the one-to-one perceptual crossing conflicts among these objects. Inter-object connectivity and more accurate sensory and object-level negotiation are requirements for such solutions (Vertegaal, 2003).

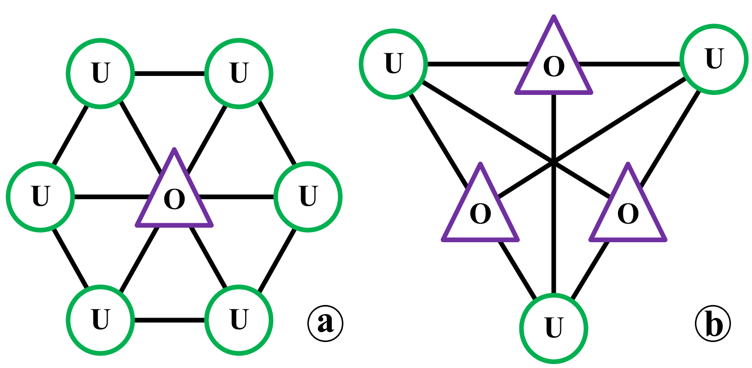


Figure 53 (a) Multi-triadic scenario (1-object, many users). (b) many objects, many users scenario.

6.3.6. Privacy

Camera modules used for gaze and face track generally lead to privacy issues (Oliveira *et al.*, 2018), especially when a device is internet-connected. Data protection and obfuscation services are required, and the system can also drain the data without making it for other uses. However, these solutions need to be made explicit to the users to be aware of them.

6.3.7. Limitation and future work

The object's limited hardware capability also limits the exploration in the real world since the accuracy of the sensor module's face, and gaze recognition depends on the distance. The signal processing has a 0.5–1*s* latency that could reduce the real-time object-user interaction's responsiveness. The latency problem could be improved by integrating a better body and gaze tracker with a more powerful computer running a more reliable body (Cao *et al.*, 2017) and a gaze detection algorithm (Kellnhofer *et al.*, 2019). However, factors affecting the camera and the computer need to be carefully considered to avoid the object being perceived as a computer. A more recent machine vision, the OpenMV Cam H7 (Abdelkader, 2018), could be used to overcome the limitations of the real-time face and eye-tracking latency at 60 frames per second.

Future research should explore other perceptual-to-perceptual interactions, as shown in Figure 54. The proactive expression for each perceptual quality should be investigated to allow the object to express its intention to interact. For instance, an expanded and dilated nostril of a nose could be used as a proactive expression of an object to show the intention to invite the user to smell something.

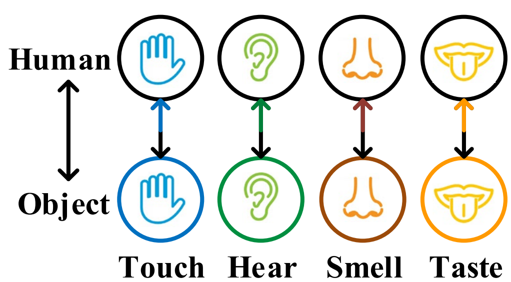


Figure 54 Perceptual quality

Future research should also consider enabling the proactive object to be aware of the context in which it is situated. Even though the initiation occurs silently in the background, it might disturb an ongoing discussion or introduce an unnecessary distraction. Identifying and adapting to the social norms of the context could further improve session initiation's social appropriateness. For instance, a proactive object aware of its potential users who are currently having a serious or a leisure discussion can interpret the situation and decide how to invite them for interaction without creating uneasiness.

6.4. Summary of Contributions

The significant contributions of this thesis are as described below:

Firstly, continuous object prototype improvement and enhancement are necessary to understand the bidirectional and proactive human-object interaction based on the perceptual crossing paradigm. Before the improvement and enhancement can be initiated, a preliminary object prototype is developed to conduct a user-experience experiment. The preliminary prototype and the UEQ questionnaires are used to study the participants' experience interacting with the prototype. The conducted study focuses on users experiencing bidirectional and proactive interaction to improve and enhance the prototype**.** However, the participants were unable to realize the bidirectional and proactive interaction expressed by the prototype. When the interaction method is limited to only one perceptual quality (i.e., eyes), the object should also have the same perceptual quality to enable the participant to involve in the same perceptual environment.

Next, a new prototype mounted with recognizable and expressive perceptual quality similar to that of human eye contact (i.e., a pair of artificial eyes) is developed. The artificial eyes are introduced to allow the prototype to create an engagement, intention, and bidirectional interaction with the user. A user-experience experiment is then conducted to evaluate the users' experience interacting with the prototype. The life-like interface agent questionnaires are used to study the participants' experience interacting with the prototype. The results show that the prototype mounted with a pair of artificial eyes can create an engagement, intention, and bidirectional interaction with the user. However, due to the reactive expression from the prototype, the perceptual crossing is not achieved.

Following the success of the participants experiencing bidirectional interaction with the prototype mounted with a pair of artificial eyes, a Session Initiation for Proactive Object (SIPO) conceptual model is introduced. The SIPO conceptual model is developed to reform the prototype's reactive behavior into proactive, which permits the prototype to have the intention to explore, express, and engage with the user proactively. The SIPO conceptual model emphasizes the bidirectional and proactive interaction between object and human and vice versa. The SIPO conceptual model operation stage is validated for single- and two-user interaction scenarios via the real-environment and the crowd-sourced video-based user studies. These studies evaluate the participants' perception of proactive and reactive behavior of an object mounted with a pair of artificial eyes.

6.5. Conclusion

Based on the conducted literature review on the perceptual crossing paradigms, proactive behavior is essential to maintain the bidirectional interaction and engagement between two entities. However, in previously conducted research, the element of bidirectional and proactive interaction was not focused. Therefore, this thesis presents the design, development, investigation, and user-experience experiments to achieve the bidirectional and proactive human-object interaction based on the perceptual crossing paradigm. In the following, an overall summary is briefly explained to provide achievement for each study.

Firstly, in Study 1 (Chapter 3), a prototype is developed with bidirectional and proactive behavior. The results in section 3.4 show the participants unable to realize the bidirectional and proactive interaction with the prototype. Therefore, in Study 2 (Chapter 4), the prototype was further improved with artificial eyes mounted on top. The results show the artificial eyes that animate staring eyes with blinking expression help the participants experience bidirectional interaction and engagement. Even though the eyes staring animation can create bidirectional interaction, the participants responded that the object did not act proactively to initiate the interaction and engagement. Thus, to integrate proactive behavior, a Session Initiation for Proactive Object (SIPO) conceptual model based on the perceptual crossing paradigm was introduced to conceptualize the bidirectional and proactive human-object interaction and engagement. The SIPO conceptual model was also introduced with an additional winking expression to improvise the prototype proactive behavior.

Study 3 (Chapter 5) presents a primitive shape prototype mounted with artificial eyes on top and can express proactive gestures (i.e., winking) to the user of interest. The prototype is used to validate the SIPO conceptual model. Based on the developed SIPO conceptual model, the results show that the prototype successfully engages in bidirectional eye-to-eye interaction and can express proactive gestures by winking to initiate interaction. The SIPO conceptual model is also tested in the crowd-sourced video-based user study for further validation. The results confirm that winking can be used as a proactive gesture that allows a proactive object to initiate interaction. The winking expression also influenced the users to take reciprocal action (e.g., winked back), enabling them to maintain a stable and continuous bidirectional interaction. Therefore, the perceptual crossing between the object and the user is achieved.

Auvray *et al.*'s perceptual crossing paradigm and the variation of works derived from this paradigm provide useful insights into understanding social interaction dynamics. The paradigm proved that people in one-dimensional space and without any other visual representation can still interact with each other in real-time by only depending on one perceptual input. As long as both are involved in the same perceptual environment and experience bidirectional and proactive interaction, they can effectively communicate and understand each other. This thesis results from the useful insight of the perceptual crossing paradigm in which the paradigm's outcome is implemented into improving the human-object interaction. Depending on one perceptual quality, an object can show its intention to interact bidirectionally with a user when both are in the same perceptual environment. In this way, the design space for improving human-object interaction can be minimized and simplified, which would benefit the designer into using the only required perceptual quality for an object to interact with the user proactively. Overall, by referring to the developed Session Initiation for Proactive Object (SIPO) conceptual model, this thesis contributes a new standpoint for future researchers to explore how to create smartness in proactive objects that interact with people.

References

Abramova, E., Slors, M.. (2019). Mechanistic Explanation for Enactive Sociality. Phenomenology and the Cognitive Sciences 18, pp. 401–424. https://doi.org/10.1007/s11097-018-9577-8

Abdelkader, I. “OpenMV Cam H7.” OpenMV, Aug. 2018, https://openmv.io/products/openmv-cam-h7.

Admoni, H., Scassellati, B. (2017). Social Eye Gaze in Human-Robot Interaction: A Review. Journal of Human-Robot Interaction 6, pp. 25–63. https://doi.org/10.5898/JHRI.6.1.Admoni

Ahson, S.A., Ilyas, M. (2018). SIP Handbook: Services, Technologies, and Security of Session Initiation Protocol. CRC Press.

Alexander, J., Roudaut, A., Steimle, J., Hornbæk, K., Bruns Alonso, M., Follmer, S., Merritt, T. (2018). Grand Challenges in Shape-Changing Interface Research. In Proceedings of the SGCHI conference Human Factors Computing Systems - CHI’18. Association for Computing Machinery, Montreal QC, Canada, pp. 1–14. https://doi.org/10.1145/3173574.3173873

Aljaroodi, H.M., Adam, M.T.P., Chiong, R., Teubner, T. (2019). Avatars and Embodied Agents in Experimental Information Systems Research: A Systematic Review and Conceptual Framework. Australasian Journal of Information Systems, vol. 23, pp. 1–37. https://doi.org/10.3127/ajis.v23i0.1841

Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, P., Suh, J., Iqbal, S., Bennett, P.N., Inkpen, K., Teevan, J., Kikin-Gil, R., Horvitz, E. (2019). Guidelines for Human-AI Interaction. In Proceedings of the SGCHI conference Human Factors Computing Systems - CHI’19. Association for Computing Machinery, Glasgow, Scotland UK, pp. 1–13. https://doi.org/10.1145/3290605.3300233

Anas, S.A.B., Liang, R.H., Hu, J., Rauterberg, M. (2020). Designing Proactive Objects with Artificial Eyes Based on Perceptual Crossing Paradigm. In 29th IEEE International Conference on Robot and Human Interactive Communication, RO-MAN’20, Naples, Italy, pp. 237–244. https://doi.org/10.1109/RO-MAN47096.2020.9223528

Anas, S.A.B., Qiu, S., Rauterberg, M., Hu, J. (2016). Exploring Social Interaction with Everyday Object Based on Perceptual Crossing. In Proceedings of the 4th International Conference Human-Agent Interaction, HAI’16. Association for Computing Machinery, Biopolis, Singapore, pp. 11–18. https://doi.org/10.1145/2974804.2974810

Anas, S.A.B., Rauterberg, M., Hu, J. (2017). Designing Elements for a Gaze Sensitive Object: Meet the CoffeePet. In Proceedings of the 5th International Conference Human-Agent Interaction, HAI’17. Association for Computing Machinery, Bielefeld, Germany, pp. 223–231. https://doi.org/10.1145/3125739.3125745

Andrist, S., Gleicher, M., Mutlu, B. (2017). Looking Coordinated: Bidirectional Gaze Mechanisms for Collaborative Interaction with Virtual Characters. In Proceedings of the SGCHI conference Human Factors Computing Systems - CHI’17. Association for Computing Machinery, Denver, Colorado, USA, pp. 2571–2582. https://doi.org/10.1145/3025453.3026033

Angelini, L., Caon, M., Lalanne D., Abou khaled, O., Mugellini, E. (2015). Towards an Anthropomorphic Lamp for Affective Interaction. In Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction, TEI'15. Association for Computing Machinery, New York, USA, pp. 661–666. **https://doi.org/10.1145/2677199.2687914**

Atzori, L., Iera, A., Morabito, G. (2017). Understanding the Internet of Things: Definition, Potentials, and Societal Role of a Fast Evolving Paradigm. Ad Hoc Networks, vol. 56, pp.122–140. **https://doi.org/10.1016/j.adhoc.2016.12.004**

Auvray, M. (2019). Multisensory and Spatial Processes in Sensory Substitution. Restorative Neurology and Neuroscience, vol. 37, pp. 609–619. **https://doi.org/10.3233/RNN-190950**

Auvray, M., Lenay, C., Stewart, J. (2009). Perceptual Interactions in a Minimalist Virtual Environment. New Ideas in Psychology, vol. 27, pp. 32–47. **https://doi.org/10.1016/j.newideapsych.2007.12.002**

Barone, P., Bedia, M.G., Gomila, A. (2020). A Minimal Turing Test: Reciprocal Sensorimotor Contingencies for Interaction Detection. Frontiers in Human Neuroscience, vol. 14(102), pp. 1–19. **https://doi.org/10.3389/fnhum.2020.00102**

Barreiros, C.A.S, Veas, E.E, Pammer, V. (2017). BioIoT: Communicating Sensory Information of a Coffee Machine Using a Nature Metaphor. Proceedings of the CHI Conference Extended Abstracts on Human Factors in Computing Systems - CHI'17. Association for Computing Machinery, New York, USA, pp. 2388–2394. **https://doi.org/10.1145/3027063.3053193**

Belk, R., Kniazeva, M. (2018). Morphing Anthropomorphism: An update. Journal of Global Scholars of Marketing Science, vol. 28(3), pp. 239–247. **https://doi.org/10.1080/21639159.2018.1466659**

Borgo, R., Micallef, L., Bach, B., McGee, F., Lee, B. (2018). Information Visualization Evaluation Using Crowdsourcing. Computer Graphics Forum, vol. 37(3), pp. 573–595. **https://doi.org/10.1111/cgf.13444**

Bryant, D., Borenstein, J., Howard, A. (2020). Why Should We Gender? The Effect of Robot Gendering and Occupational Stereotypes on Human Trust and Perceived Competency. In Proceedings of the ACM/IEEE International Conference Human-Robot Interaction, HRI'20. Association for Computing Machinery, Cambridge, United Kingdom, pp. 13–21. **https://doi.org/10.1145/3319502.3374778**

Buck, J.W., Perugini, S., Nguyen, T.V. (2018). Natural Language, Mixed-Initiative Personal Assistant Agents. In Proceedings of the 12th International Conference Ubiquitous Information Management Communication, IMCOM'18. Association for Computing Machinery, Langkawi, Malaysia, pp. 1–8. **https://doi.org/10.1145/3164541.3164609**

Cao, Z., Simon, T., Wei, S.-E., Sheikh, Y. (2017). Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields. In IEEE Conference on Computer Vision and Pattern Recognition, CVPR'17, Honolulu, Hawaii, pp. 1302-1310. **https://doi.org/10.1109/CVPR.2017.143.**

Cena, F., Console, L., Matassa, A., Torre, I. (2019). Multi-dimensional Intelligence in Smart Physical Objects. Information Systems Frontiers, vol. 21, pp. 383–404. **https://doi.org/10.1007/s10796-017-9758-y**

Cha, E., Kim, Y., Fong, T., Mataric, M.J. (2018). A Survey of Nonverbal Signaling Methods for Non-Humanoid Robots. Foundations and Trends in Robotics, vol. 6(4), pp. 211–323. **http://dx.doi.org/10.1561/2300000057**

Chuang, Y., Chen, L.-L., Liu, Y. (2018). Design Vocabulary for Human–IoT Systems Communication. In Proceedings of the SGCHI conference Human Factors Computing Systems - CHI'18. Association for Computing Machinery, Montreal QC, Canada, pp. 1–11. **https://doi.org/10.1145/3173574.3173848**

Chung, D., Funk, M., Liang, R.-H., Chen, L.-L. (2018). Explorations on Reciprocal Interplay in Things Ecology. In Proceedings of the ACM Conference Companion Publication Designing Interactive Systems, DIS'18 Companion. Association for Computing Machinery, Hong Kong, China, pp. 51–56. **https://doi.org/10.1145/3197391.3205411**

Ciechanowski, L., Przegalinska, A., Magnuski, M., Gloor, P. (2019). In the Shades of the Uncanny Valley: An Experimental Study of Human–Chatbot Interaction. Future Generation Computer Systems, vol. 92, pp. 539–548. **https://doi.org/10.1016/j.future.2018.01.055**

Clodic, A., Pacherie, E., Alami, R., Chatila, R. (2017). Key Elements for Human-Robot Joint Action. In Hakli, R., Seibt, J. (Eds.), Sociality Normativity Robots Philosophical Inquiries Human-Robot Interactions. Springer International Publishing, Cham, pp. 159–177. **https://doi.org/10.1007/978-3-319-53133-5_8**

Deckers, E., Lévy, P., Wensveen, S., Ahn, R., Overbeeke, K. (2013). Designing for Perceptual Crossing: Applying and Evaluating Design Notions. International Journal of Design, vol. 6, pp. 41–55.

Deckers, E., Wensveen, S., Ahn, R., Overbeeke, K. (2011). Designing for Perceptual Crossing to Improve User Involvement. In Proceedings of the SGCHI conferenceHumanFactors Computing Systems - CHI'11. Association for Computing Machinery, Vancouver, BC, Canada, pp. 1929–1938. **https://doi.org/10.1145/1978942.1979222**

Deschamps, L., Lenay, C., Rovira, K., Le Bihan, G., Aubert, D. (2016). Joint Perception of a Shared Object: A Minimalist Perceptual Crossing Experiment. Frontiers in Psychology, vol. 7(1059), pp. 1–15. **https://doi.org/10.3389/fpsyg.2016.01059**

Dove, G., Halskov, K., Forlizzi, J., and Zimmerman, J. (2017). UX Design Innovation: Challenges for Working with Machine Learning as a Design Material. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, CHI'17. Association for Computing Machinery, New York, USA, pp. 278–288. **https://doi.org/10.1145/3025453.3025739**

Eclipse Foundation, Inc. (2020). The Community for Open Innovation and Collaboration. The Eclipse Foundation. Eclipse. Retrieved from **http://www.eclipse.org**

Fortino, G., Russo, W., Savaglio, C., Shen, W., Zhou, M. (2018). Agent-Oriented Cooperative Smart Objects: From IoT System Design to Implementation. IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 48, pp. 1939–1956. **https://doi.org/10.1109/TSMC.2017.2780618**

Frich, J., Mose Biskjaer, M., Dalsgaard, P. (2018). Twenty Years of Creativity Research in Human-Computer Interaction: Current State and Future Directions. In Proceedings of the 2018 Designing Interactive Systems Conference, DIS'18. Association for Computing Machinery, Hong Kong, China, pp. 1235–1257. **https://doi.org/10.1145/3196709.3196732**

Froese, T., Di Paolo, E.A. (2009). Modeling Social Interactions as Perceptual Crossing: An Investigation into Dynamics of the Interaction Process. Connection Science, vol. 22(1), pp. 43–68. **https://doi.org/10.1080/09540090903197928**

Froese, T., Zapata-Fonseca, L., Leenen, I., Fossion, R. (2020). The Feeling Is Mutual: Clarity of Haptics-Mediated Social Perception Is Not Associated With the Recognition of the Other, Only

With Recognition of Each Other. Frontiers in Human Neuroscience, vol. 14(560567), pp. 1–8. https://doi.org/10.3389/fnhum.2020.560567

Galitsky, B., Ilvovsky, D. (2020). Interrupt Me Politely: Recommending Products and Services by Joining Human Conversation. In Proceedings of the Workshop Natural Language Processing E-Commerce. Association for Computational Linguistics, Barcelona, Spain, pp. 32–42.

Ghazali, A.S., Ham, J., Barakova, E.I., Markopoulos, P. (2017). Pardon the Rude Robot: Social Cues Diminish Reactance to High Controlling Language. In 26th IEEE International Conference on Robot and Human Interactive Communication, RO-MAN'17, Lisbon, Portugal, pp. 411–417. https://doi.org/10.1109/ROMAN.2017.8172335

Goldstein, S. (2019). Cross-Cultural Explorations: Activities in Culture and Psychology, 3rd ed. Routledge. https://doi.org/10.4324/9780429197086

Hadley, L.V., Brimijoin, W.O., Whitmer, W.M. (2019). Speech, Movement, and Gaze Behaviours During Dyadic Conversation in Noise. Scientific Reports, vol. 9(10451), pp. 1–8. https://doi.org/10.1038/s41598-019-46416-0

Hasbro, Inc. (2020). Furby Official Website, Furby Boom, Hasbro. Retrieved from https://furby.hasbro.com/en-us

Hayashi, K., Mizuuchi, I. (2017). Investigation Of Joint Action: Eye Blinking Behavior Improving Human-Robot Collaboration. In 26th IEEE International Conference on Robot and Human Interactive Communication, RO-MAN'17, Lisbon, Portugal, pp. 1133–1139. https://doi.org/10.1109/ROMAN.2017.8172446

Hermans, K.S.F.M., Kasanova, Z., Zapata-Fonseca, L., Lafit, G., Fossion, R., Froese, T., Myin-Germeys, I. (2020). Investigating Real-Time Social Interaction in Pairs of Adolescents with the Perceptual Crossing Experiment. Behavior Research Methods, vol. 52, pp.1929–1938. https://doi.org/10.3758/s13428-020-01378-4

Hetherington, N.J. (2020). Design and Evaluation of Nonverbal Motion Cues for Human-Robot Spatial Interaction. Electronic Theses and Dissertations (ETDs). University of British Columbia. https://doi.org/http://dx.doi.org/10.14288/1.0394282

Hietanen, J.K. (2018). Affective Eye Contact: An Integrative Review. Frontiers in Psychology, vol. 9(1587), pp. 1–15. https://doi.org/10.3389/fpsyg.2018.01587

Hinderks, A., Schrepp, M., Thomaschewski, J. (2018). User Experience Questionnaire (UEQ). UEQ User Experience Questionnaire. Retrieved from https://www.ueq-online.org/

Jabarin, B., Wu, J., Vertegaal, R., Grigorov, L. (2003). Establishing Remote Conversations Through Eye Contact with Physical Awareness Proxies. In Extended Abstracts on Human Factors in Computing Systems - CHI'03. Association for Computing Machinery, New York, USA, pp. 948–949. https://doi.org/10.1145/765891.766087

Jarick, M., Bencic, R. (2019). Eye Contact Is a Two-Way Street: Arousal Is Elicited by the Sending and Receiving of Eye Gaze Information. Frontiers in Psychology, vol. 10(1262), pp. pp. 1–14. https://doi.org/10.3389/fpsyg.2019.01262

Johansen, S. A., Agustin, J. S., Tall, M., Skovsgaard, H. (2013, October 12). The Eye Tribe. Retrieved from https://theeyetribe.com/theeyetribe.com/about/index.html

Kao, H. C. and Schmandt, C. (2014). MugShots: Everyday Objects as Social Catalysts. In Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication. New York, USA, pp 75–78. https://doi.org/10.1145/2638728.2638739

Kellnhofer, P., Recasens, A., Stent, S., Matusik, W., Torralba, A. (2019). Gaze360: Physically Unconstrained Gaze Estimation in the Wild. In Proceedings of the IEEE/CVF International Conference on Computer Vision, ICCV'19, Seoul, Korea (South), pp. 6911–6920. **https://doi.org/10.1109/ICCV.2019.00701**

Kiilavuori, H., Sariola, V., Peltola, M.J., Hietanen, J.K., 2021. Making Eye Contact With a Robot: Psychophysiological Responses to Eye Contact With a Human and With a Humanoid Robot. Biological Psychology, vol. 158(107989), pp. 1–13. **https://doi.org/10.1016/j.biopsycho.2020.107989**

Knote, R., Janson, A., Söllner, M., Leimeister, J.M. (2019). Classifying Smart Personal Assistants: An Empirical Cluster Analysis. In Proceedings of the 52nd Hawaii International Conference on System Sciences, HICSS'19, pp. 2024–2033.

Kowalczyk, P., Sawicki, D. (2019). Blink And Wink Detection As a Control Tool in Multimodal Interaction. Multimedia Tools and Applications, vol. 78, pp. 13749–13765. **https://doi.org/10.1007/s11042-018-6554-8**

Laugwitz B., Held T., Schrepp M. (2008) Construction and Evaluation of a User Experience Questionnaire. In Holzinger A. (Eds.) HCI and Usability for Education and Work, USAB'08. Lecture Notes in Computer Science, vol. 5298, pp. 63–76. Springer, Berlin, Heidelberg. **https://doi.org/10.1007/978-3-540-89350-9_6**

Lego System. (2020). Lego. Rebuild the World. Retrieved from **https://www.lego.com**

Lenay, C. (2017). Explanatory Schemes for Social Cognition - A Minimalist Interaction-Based Approach. Pragmatism Today, vol. 8(1), pp. 63–86.

Lenay, C., Stewart, J. (2012). Minimalist Approach to Perceptual Interactions. Frontiers in Human Neuroscience, vol. 6, pp. 98–115. **https://doi.org/10.3389/fnhum.2012.00098**

Levillain, F., Zibetti, E. (2017). Behavioral Objects: The Rise of The Evocative Machines. J. Hum.-Robot Interact, vol. 6(1), pp. 4–24. **https://doi.org/10.5898/JHRI.6.1.Levillain**

Lin, Z., Carley, K. (1993). Proactive or Reactive: An Analysis of the Effect of Agent Style on Organizational Decision-Making Performance. Intelligent Systems in Accounting, Finance and Management, vol. 2(4), pp. 271–287. **https://doi.org/10.1002/j.1099-1174.1993.tb00047.x**

Liu, Y., Chuang, Y., Lee, Y., Liang, R., Chen, L. (2017). Designing the Expressiveness of Point Lights for Bridging Human-IoT System Communications. In Proceedings of the Conference on Design and Semantics of Form and Movement: Sense and Sensitivity, DeSForM'17, pp. 619–706. **https://doi.org/10.5772/intechopen.71131**

Lizuka, H., Marocco, D., Ando, H., Maeda, T. (2012). Turn-Taking Supports Humanlikeness And Communication In Perceptual Crossing Experiments - Toward Developing Human-Like Communicable Interface Devices. IEEE Virtual Reality Workshops, VRW'12, Costa Mesa, California, pp. 1–4. **https://doi.org/10.1109/VR.2012.6180953**

Lloyd, M. (2019). The Non-Looks of the Mobile World: A Video-Based Study of Interactional Adaptation in Cycle-Lanes. Mobilities, vol. 14(4), pp. 500–523. **https://doi.org/10.1080/17450101.2019.1571721**

Lopez, A., Paredes, R., Quiroz, D., Trovato, G., Cuellar, F. (2017). Robotman: A Security Robot for Human-Robot Interaction. In 18th International Conference Advanced Robotics, ICAR'17. pp. 7–12. **https://doi.org/10.1109/ICAR.2017.8023489**

Michelle M.E. Van Pinxteren, Mark Pluymaekers, Jos G.A.M. Lemmink. (2020). Human-Like Communication in Conversational Agents: A Literature Review and Research Agenda. Journal of Service Management, vol. 31(2), pp. 203–225. https://doi.org/10.1108/JOSM-06-2019-0175

Majaranta, P., Räihä, K.-J., Hyrskykari, A., Špakov, O. (2019). Eye Movements and Human-Computer Interaction. In Klein C., Ettinger U. (Eds.) Eye Movement Research. Studies in Neuroscience, Psychology and Behavioral Economics. Springer, pp. 971–1015. https://doi.org/10.1007/978-3-030-20085-5_23

Marti, P. (2012). Materials of Embodied Interaction. In Proceedings of the 1st workshop on Smart Material Interfaces: A Material Step to the Future, SMI'12. Association for Computing Machinery, Santa Monica, California, pp. 1–6. https://doi.org/10.1145/2459056.2459058

Marti, P., Iacono, I., Parlangeli, O., Stienstra, J. (2020). Social Coordination in Human-Robot Interaction Through Reciprocal Engagement. In Rehm M., Saldien J., Manca S. (Eds.) Project and Design Literacy as Cornerstones of Smart Education. Smart Innovation, Systems and Technologies, vol. 158. Springer, Singapore, pp. 273–285. https://doi.org/10.1007/978-981-13-9652-6_25

Marti, P., Iacono, I., Stienstra, J., Tittarelli, M. (2014). Exploring Movement Qualities in Reciprocal Engagement. In 4th International Conference on Development and Learning and on Epigenetic Robotics, Genoa, pp. 124–129. https://doi.org/10.1109/DEVLRN.2014.6982966.

Marti, P., Stienstra, J.T. (2013). Exploring Empathy in Interaction : Scenarios of Respectful Robotics. GeroPsych: The Journal of Gerontopsychology and Geriatric Psychiatry, vol. 26(2), pp. 101–112. https://doi.org/10.1024/1662-9647/a000086

Mouser Electronics, Inc. (2016, August 31). Omron Electronics B5T HVC-P2 Image Sensor Modules. Omron Electronic Components. Retrieved from https://eu.mouser.com/new/omronelectronics/omron-b5t-hvc-p2-sensors/

Muller, V. (2019, May 27). Ulo. Mu Design. Retrieved from https://mu-design.lu/ulo

Nanayakkara, S., Huber, J., Sridhar, P. (2018). Augmented Sensors. In Huber, J., Shilkrot, R., Maes, P., Nanayakkara, S. (Eds.), Assistive Augmentation. Cognitive Science and Technology. Springer, Singapore, pp. 7–21. https://doi.org/10.1007/978-981-10-6404-3_2

Newn, J., Singh, R., Allison, F., Madumal, P., Velloso, E., Vetere, F. (2019). Designing Interactions with Intention-Aware Gaze-Enabled Artificial Agents. In Lamas D., Loizides F., Nacke L., Petrie H., Winckler M., Zaphiris P. (Eds.) Human-Computer Interaction – INTERACT'19. Lecture Notes in Computer Science, vol. 11747. Springer, Cham, pp. 255–281. https://doi.org/10.1007/978-3-030-29384-0_17

Oh, H., Kim, J., Morales, C., Gross, M., Eisenberg, M., Hsi, S. (2017). FoldMecha: Exploratory Design and Engineering of Mechanical Papercraft. In Proceedings of the 11th International Conference Tangible, Embedded, Embodied Interaction, TEI'17. Association for Computing Machinery, Yokohama, Japan, pp. 131–139. https://doi.org/10.1145/3024969.3024991

Oliveira, L.B., Pereira, F.M.Q., Misoczki, R., Aranha, D.F., Borges, F., Nogueira, M., Wangham, M., Wu, M., Liu, J. (2018). The Computer for the 21st Century: Present Security and Privacy Challenges. Journal of Internet Services and Applications, vol. 9(24), pp. 1–25. https://doi.org/10.1186/s13174-018-0095-2

Oluchukwu, I.H., Mustafa, I., Mohammed, A.A., Festus, V.B. (2018). Crowd-Sourcing (Who, Why and What). International Journal of Crowd Science vol. 2(1), pp. 27–41. https://doi.org/10.1108/IJCS-07-2017-0005

Oogjes, D., Wakkary, R. (2017). Videos of Things: Speculating on, Anticipating and Synthesizing Technological Mediations. In Proceedings of the SGCHI conference Human Factors Computing Systems - CHI'17. Association for Computing Machinery, Denver, Colorado, USA, pp. 4489–4500. **https://doi.org/10.1145/3025453.3025748**

Oracle Corporation. (2020). Oracle Integrated Cloud Applications and Platform Services. Oracle. Retrieved from **https://www.oracle.com**

Oertel, C., Castellano, G., Chetouani, M., Nasir, J., Obaid, M., Pelachaud, C., Peters, C. (2020). Engagement in Human-Agent Interaction: An Overview. Frontiers in Robotics and AI vol. 7(92), pp. 1–21. **https://doi.org/10.3389/frobt.2020.00092**

Parise, S., Kiesler, S., Sproull, L., Waters, K. (1999). Cooperating with Life-Like Interface Agents. Computers in Human Behavior, vol. 15, pp. 123–142. **http://dx.doi.org/10.1016/S0747-5632(98)00035-1**

Park, S., Lee, J.-H., Lee, W. (2019). The Effects of Workplace Rest Breaks on Health Problems Related to Long Working Hours and Shift Work among Male Apartment Janitors in Korea. Safety and Health at Work, vol. 10, pp. 512–517. **https://doi.org/10.1016/j.shaw.2019.10.003**

Petrov, V., Mikhaylov, K., Moltchanov, D., Andreev, S., Fodor, G., Torsner, J., Yanikomeroglu, H., Juntti, M.J., Koucheryavy, Y. (2017). When IoT Keeps People in the Loop: A Path Towards a New Global Utility. In IEEE Communications Magazine, vol. 57(1), pp. 114-121. **https://doi.org/10.1109/MCOM.2018.1700018**

Ren, Zhi-peng. (2014). Body Language in Different Cultures. US-China Foreign Language 12(12), pp. 1029–1033.**https://doi.org/10.17265/1539-8080/2014.12.008**

Rini, J.F., Ochoa, J. (2020). Behavioral Implications of Temporal Lobe Epilepsy on Social Contingency. Epilepsy an Behavior, vol. 110(107101), pp. 1–8. **https://doi.org/10.1016/j.yebeh.2020.107101**

Rosen, E., Whitney, D, Fishman, M., Ullman, D., Tellex, S. (2020). Mixed Reality as a Bidirectional Communication Interface for Human-Robot Interaction. IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS'20. Las Vegas, NV, USA, pp. 11431-11438. **https://doi.org/10.1109/IROS45743.2020.9340822**

Ruhland, K., Andrist, S., Badler, J., Peters, C., Badler, N., Gleicher, M., Mutlu, B., Mcdonnell, R. (2014). Look Me in The Eyes: A survey of Eye and Gaze Animation for Virtual Agents and Artificial Systems. In Eurographics 2014 State Art Reports. Strasbourg, France, pp. 69–91. **https://doi.org/10.2312/egst.20141036**

Saitoh S., Iizuka H., Yamamoto M. (2017) Increasing Stability of Human Interaction Against Time Delay on Perceptual Crossing Experiment. In Leu G., Singh H., Elsayed S. (Eds.) Intelligent and Evolutionary Systems. Proceedings in Adaptation, Learning and Optimization, vol. 8. Springer, Cham, pp. 361–369. **https://doi.org/10.1007/978-3-319-49049-6_26**

Schönherr, J., Westra, E. (2017). Beyond 'Interaction': How to Understand Social Effects on Social Cognition. The British Journal for the Philosophy of Science, vol. 70(1), pp. 27–52. **https://doi.org/10.1093/bjps/axx041**

Sejima, Y., Egawa, S., Maeda, R., Sato, Y., Watanabe, T. (2018). A Speech-Driven Pupil Response System with Affective Expression Using Hemispherical Displays. In 27th IEEE International Conference on Robot and Human Interactive Communication, RO-MAN'18, Nanjing, China, pp. 228–233. **https://doi.org/10.1109/ROMAN.2018.8525764**

Sharmin, S., Hoque, M.M., (2020). Developing an Empirical Robotic Framework to Establish Bidirectional Eye Contact. In Vasant, P., Zelinka, I., Weber, G.-W. (Eds.), IntelligentComputing Optimization. Springer International Publishing, Cham, pp. 389–398. **https://doi.org/10.1007/978-3-030-33585-4_39**

Siposova, B., Carpenter, M. (2019). A New Look at Joint Attention and Common Knowledge. Cognition, vol. 189, pp. 260–274. **https://doi.org/10.1016/j.cognition.2019.03.019**

Song, S., Yamada, S. (2018). Designing Expressive Lights and In-Situ Motions for Robots to Express Emotions. In Proceedings of the 6th International Conference Human-Agent Interaction, HAI'18. Association for Computing Machinery, Southampton, United Kingdom, pp. 222–228. **https://doi.org/10.1145/3284432.3284458**

Sturdee, M., Everitt, A., Lindley, J., Coulton, P., Alexander, J. (2019). Visual Methods for the Design of Shape-Changing Interfaces. In Lamas, D., Loizides, F., Nacke, L., Petrie, H., Winckler, M., Zaphiris, P. (Eds.), Human-Computer Interaction, INTERACT'19. Lecture Notes in Computer Science, vol. 11748. Springer, Cham, pp. 337–358. **https://doi.org/10.1007/978-3-030-29387-1_19**

Sun, Y., Shaikh, O., Won, A.S. (2019). Nonverbal Synchrony in Virtual Reality. PLOS ONE, vol. 14(9), pp. 1–28. **https://doi.org/10.1371/journal.pone.0221803**

Todorovic, D. (2019). Effects of Changes of Observer Vantage Points on the Perception of Spatial Structure in Perspective Images: Basic Geometric Analysis. Axiomathes, pp. 1–27. **https://doi.org/10.1007/s10516-019-09421-6**

Torta, E., van Heumen, J., Piunti, F., Romeo, L., & Cuijpers, R. (2014). Evaluation of Unimodal and Multimodal Communication Cues for Attracting Attention in Human-Robot Interaction. International Journal of Social Robotics, vol. 7(1), pp. 89–96. **ttps://doi.org/10.1007/s12369-014-0271-x**

Vázquez, M., Carter, E.J., McDorman, B., Forlizzi, J., Steinfeld, A., Hudson, S.E. (2017). Towards Robot Autonomy in Group Conversations: Understanding the Effects of Body Orientation and Gaze. In 12th ACM/IEEE International Conference Human-Robot Interaction, HRI'07. pp. 42–52.

Vertegaal, R. Attentive User Interfaces. (2003) Editorial, In Special Issue on Attentive User Interfaces, Communications of ACM, vol. 46(3). ACM Press, pp. 30–33

Vertegaal, R., Slagter, R., van der Veer, G., and Nijholt, A. (2001) Eye Gaze Patterns in Conversations: There is More to Conversational Agents than Meets the Eyes. In Proceedings of Conference on Human Factors in Computing Systems - CHI'01. ACM Press, pp. 301–308. **https://doi.org/10.1145/365024.365119**

Wang, G., Cheng, T., Do, Y., Yang, H., Tao, Y., Gu, J., An, B., Yao, L. (2018). Printed Paper Actuator: A Low-cost Reversible Actuation and Sensing Method for Shape Changing Interfaces. In Proceedings of the SGCHI conference Human Factors Computing Systems - CHI'18. Association for Computing Machinery, Montreal QC, Canada, pp. 1–12. **https://doi.org/10.1145/3173574.3174143**

Wang, R. (2017). Leveraging Team Diversity toward Collaborative Crowdsourcing Success. Academy of Management Journal, vol. 2017(1), pp. 225–244 **https://doi.org/10.5465/AMBPP.2017.11225abstract**

Wicked Cool Toys. (2020). Cabbage Patch Kids. Toy Cabbage Patch Kids. Retrieved from **https://cabbagepatchkids.com/**

Zapata-Fonseca, L., Dotov, D., Fossion, R., Froese, T., Schilbach, L., Vogeley, K., Timmermans, B. (2019). Multi-Scale Coordination of Distinctive Movement Patterns During Embodied Interaction Between Adults with High-Functioning Autism and Neurotypicals. Frontiers in Psychology, vol. 9(2760), pp. 1–7. **https://doi.org/10.3389/fpsyg.2018.02760**

Zapata-Fonseca, L., Froese, T., Schilbach, L., Vogeley, K., Timmermans, B. (2018). Sensitivity to Social Contingency in Adults with High-Functioning Autism during Computer-Mediated Embodied Interaction. Behavioral Sciences, vol. 8(22), pp. 1–14. **https://doi.org/10.3390/bs8020022**

Zhang, X., Sugano, Y., Bulling, A. (2017). Everyday Eye Contact Detection Using Unsupervised Gaze Target Discovery. In Proceedings of the 30th Annual ACM Symposium User Interface Software Technology, UIST'17. Association for Computing Machinery, Québec City, QC, Canada, pp. 193–203. **https://doi.org/10.1145/3126594.3126614**

Zhang, Y., Beskow, J., Kjellström, H. (2017). Look but Don't Stare : Mutual Gaze Interaction in Social Robots. In 9th International Conference on Social Robotics, ICSR'17, vol. 10652, pp. 556–566. **https://doi.org/10.1007/978-3-319-70022-9_55**

Zualkernan, I.A., Aloul, F., Algebail, E., Refaay, M.E., Ali, A., Sabaa, O.E. (2020). Little Genius: An Experiment in Internet of Tangible Learning Things. In IEEE 20th International Conference Advanced Learning Technologies, ICALT'20, pp. 69–71. **https://doi.org/10.1109/ICALT49669.2020.00028**

Appendix A.

User Experience Questionnaire

	−3	−2	−1	0	+1	+2	+3		
annoying	O	O	O	O	O	O	O	enjoyable	1
not understandable	O	O	O	O	O	O	O	understandable	2
dull	O	O	O	O	O	O	O	creative	3
difficult to learn	O	O	O	O	O	O	O	easy to learn	4
inferior	O	O	O	O	O	O	O	valuable	5
boring	O	O	O	O	O	O	O	exciting	6
not interesting	O	O	O	O	O	O	O	interesting	7
unpredictable	O	O	O	O	O	O	O	predictable	8
slow	O	O	O	O	O	O	O	fast	9
conventional	O	O	O	O	O	O	O	inventive	10
obstructive	O	O	O	O	O	O	O	supportive	11
bad	O	O	O	O	O	O	O	good	12
complicated	O	O	O	O	O	O	O	easy	13
unlikeable	O	O	O	O	O	O	O	pleasing	14
usual	O	O	O	O	O	O	O	leading edge	15
unpleasant	O	O	O	O	O	O	O	pleasant	16
not secure	O	O	O	O	O	O	O	secure	17
demotivating	O	O	O	O	O	O	O	motivating	18
does not meet expectations	O	O	O	O	O	O	O	meet expectations	19
inefficient	O	O	O	O	O	O	O	efficient	20
confusing	O	O	O	O	O	O	O	clear	21
impractical	O	O	O	O	O	O	O	practical	22
cluttered	O	O	O	O	O	O	O	organize	23
unattractive	O	O	O	O	O	O	O	attractive	24
unfriendly	O	O	O	O	O	O	O	friendly	25
conservative	O	O	O	O	O	O	O	innovative	26

Appendix B.

Life-Like Interface Agent Questionnaire

		1	2	3	4	5	6	7
Appearance	humanlike	O	O	O	O	O	O	O
	attractive	O	O	O	O	O	O	O
	sociable	O	O	O	O	O	O	O
	intelligent	O	O	O	O	O	O	O
Partnership	mutual like	O	O	O	O	O	O	O
	trustworthy	O	O	O	O	O	O	O
Interaction	difficult to understand	O	O	O	O	O	O	O
	enjoyable	O	O	O	O	O	O	O

Appendix C.

Proactive-Reactive Measures Questionnaire

You want to interact with Box. Box turn towards you. Do you perceive Box as a proactive object or a reactive object?

O	7 – Proactive
O	6 – Somewhat Proactive
O	5 – Slightly Proactive
O	4 – Neither Proactive nor Reactive
O	3 – Slightly Reactive
O	2 – Somewhat Reactive
O	1 – Reactive

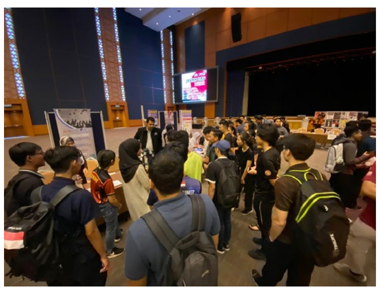
Freely describe the experiences with Box (one sentence minimal):

Appendix D.

Proactive Object In-Crowd







Curriculum Vitae

Siti Aisyah binti Anas was born on the 17th of May 1985 in Selangor, Malaysia. She received her Bachelor's Degree in Electronic and Computer Engineering from Universiti Teknikal Malaysia Melaka (UTeM) in 2008. In the same year, she was soon offered to work as a tutor in UTeM in the Department of Computer Engineering, Faculty of Electronic Engineering and Computer Engineering (FKEKK) professionally. In 2010, she completed her Master of Engineering program in Communication and Computer Engineering from Universiti Kebangsaan Malaysia (UKM). Upon completing her study, she was appointed as a lecturer and became an active member at the Advance Sensors & Embedded Controls System (ASECS), UTeM. Her active participation was reflected when she was awarded 11 research grants and managed to secure one utility pattern and four copyright works, contributing two gold, three silver, and 13 bronze awards from various invention competitions.

In January 2015, she started a Ph.D. project in the Department of Industrial Design at the Eindhoven University of Technology (TU/e), Netherlands, fully funded by the Malaysia Ministry of Education (MoE) and UTeM. Her Ph.D. project was carried out under the supervision of prof. Dr. Matthias Rauterberg, Assoc. Prof. Dr. Jun Hu and Asst. Prof. Dr Rong-Hao Liang. This thesis results from her Ph.D. research on the topic of "Perceptual Crossing with Artificial Eyes".

Since 2019, she resumed her position as a lecturer at UTeM and continuing her research, focusing on designing for human-object interaction, and managed to secure two research grants. In the same year, she completed her Malaysian Skills Diploma in Electronic Product Development from the Department of Skills Development, Ministry of Human Resources, Malaysia. In 2020, she became a member of The Institution of Engineers, Malaysia (IEM). She was certified and registered as a Professional Engineer in the branch of Electronic Engineering from the Board of Engineers Malaysia (BEM).

Acknowledgment

As my PhD journey ends, I would like to acknowledge the support and encouragement I received during my doctoral research.

Matthias, you are the one who introduced me to the perceptual crossing paradigm and challenged me to dig deeper and understand the unique outcome of the paradigm and make it the main topic of my doctoral research. I was struggling. With no design background, this was the most challenging research hurdle that I have ever faced. But you managed to make sure that I am on the right path. Your constructive criticism pushed me to become a better designer, and this means a lot to me.

Jun Hu, during my first year of PhD, our weekly meeting has exposed me to much new industrial design information that I never heard before. You were always giving me something to read to improve my design knowledge. At first, I found it very difficult for me to process the way designers creatively think to solve a problem. Still, I managed to use the knowledge that I gained and implemented it into my doctoral research. Thank you very much for introducing me to the design world.

Rong-Hao, we met in the middle of my third year PhD. Whenever you saw me, you always asked me, "do you want to talk today?". This brings so much joy to me—your willingness to discuss and brainstorm my doctoral research has made my day brighter. You have invested a lot of effort in making sure that I could move forward, and I very much appreciated the times you have supported me throughout the last years of my PhD.

To Matthias, Jun Hu, and Rong-Hao, I am sorry that I did not manage to finish my PhD as we expected. It seems like everything around me required my full attention, and to some extend, I unable to prioritize what is important and what can be avoided. Thank you very much for not giving up on me, and I am genuinely grateful to have all of you as my supervisors. I could not have done it without all of you. Thanks again for making sure that I finished my PhD.

I would also like to express my sincere appreciation to the reading committee members: prof.dr.ir. L.M.G. Feijs, prof.dr. P. Marti, dr. M. Li, and dr. R. Vertegaal. Their insightful comments and suggestions have signicantly improved the quality of my thesis.

My thanks also go out to Sabine Smits, Marly Sluismans, Anne Jueken, Maartje Mulder, and Hanneke Driessen. They were always so helpful and provided me with their assistance throughout my PhD journey.

To my mom, I am incredibly grateful for your constant love, support, and prayers. I am truly sorry that I left you and went abroad for four years to further my study. I knew it was hard for you to adapt to the situation, but we managed to pull it out together. And I promise I will make it up to you. To my best friend, Ain, thank you for taking care of my mom during my absence. I will always treasure our friendship and will never forget your kindness to my mom. Thank you so much, Ain.

I am also grateful to Universiti Teknikal Malaysia Melaka (UTeM) and the Ministry of Higher Education, Malaysia, for giving me the opportunity to further my studies and sponsor my PhD. To my best buddy, Ranjit, thank you for your kind words and encouragement, always backing me up, and always there for me when my spirits need a little lift; they mean a lot to me, and I truly appreciate it. Last but not least, my sincerest thanks to Banan, Banza, Maziah, Kak Do, Pidah, Ah, Ammar, Abang Aathif and Amna, Ekin, Abam Zairi, Aakif and Amna, Kak Yati, Amir, Hasif and Nurin, Bang Juff, Mak Raf, Adham and Sara, Bang Nidzam, Wani, Rifqi and Naurah, Yura, Kadian, Evans, Eunice, Fikile, Monroe, Ning, Jingya, Yuan, Patrick, Bin Yu, Nan, Mark and to all the people in my life that fill my days with warmth, smiles, and friendship. Thank you for being there for me.