Designing Elements for a Gaze Sensitive Object: Meet the CoffeePet

Siti Aisyah binti Anas, Matthias Rauterberg, Jun Hu
Eindhoven University of Technology,
Department of Industrial Design,
5600 MB, Eindhoven, The Netherlands
{s.a.b.anas, g.w.m.rauterberg, j.hu}@tue.nl

ABSTRACT
It is not difficult to design an object that is sensitive to our gaze. The challenging part is to make the user realize it. And since people are not used to interacting with an object by simply moving their eyes (people use eyes to see, not to trigger something), the interaction itself become unfamiliar to them. Based on the gaze behavior of people socializing with others, we believe that the feeling of being looked back when interacting with a gaze sensitive object is rather crucial in order to overcome the problem of unfamiliarity and to make people naturally realize that the object was sensitive to their gaze behavior. To achieve this feeling, we conclude that eyes need to be presence in the user's view. In this paper, we present an anthropomorphize coffee machine called the CoffeePet, attached with two, small OLED screen that displays animated eyes. These eyes are responsive towards the user's gaze behavior. Furthermore, the CoffeePet will automatically start to brew the drink if the user manages to maintain a prolonged eye contact with it. In three experiments, we investigated the impact of the animated eyes in aiding the participants to become familiar and to realize that their gaze behavior influences the CoffeePet to react. Without being told on how to interact with the CoffeePet, participants were randomly assigned to participate in one of three conditions. 1) CoffeePet with watching eyes (eyes with direct gaze), 2) CoffeePet with interactive gaze model, and 3) CoffeePet with interactive gaze following. The results showed that the interactive sharing gaze did, in fact, lead the participants to become familiar and to realize that they can interact with the object by simply moving their eyes.

Author Keywords
Eye tracking; gaze sensitive object; anthropomorphism; human-object interaction; emotional connection.

HAI '17, October 17–20, 2017, Bielefeld, Germany

ACM ISBN 978-1-4503-5113-3/17/10…$15.00
https://doi.org/10.1145/3125739.3125745

ACM Classification Keywords
H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous;

INTRODUCTION
Eye movements are an essential behavior that people depend on. We use our eyes to scan immediate surrounding, to regulate and exercise social control, to establish emotional connection and to indicate the target of our own visual interest [14]. Our eyes could also reflect our cognitive process; provides hints of what we are currently thinking and revealing our intentions [12]. For example, we often look at things before we decide to act on it [16]. Since 1970s, tremendous effort has been done in order to develop a system that uses gaze as input method [18]. The information gathers from eye movements provide greater opportunities to create more engaging, effortless user experiences. Furthermore, this nonverbal communication has been found to be the most straightforward interaction, the user needs to simply point with their eyes[9, 11, 24, 26]. However, every silver lining has a cloud. As eyes are used to obtain information, how can the system differentiate between gaze that is observing or watching and gaze that is intended to activate a command. Otherwise, every time we consciously or unconsciously move our eyes, a new function gets activated [10]. The most common solution to overcome this so-called Midas touch problem is to deal with the dwell time, where an intentionally prolonged gaze will activate the objects[9, 25]. Yet, if the user still does not realize that their gaze influences the object to react, the interaction will become useless to them. Based on our previous experiment [1, 2], when we design a gaze sensitive object, we did solve one problem (Midas touch problem) only to cause another. The user still did not realize the object reacted because of their prolonged gaze. These failures occurred due to the fact mention earlier, we use eyes to observe the environment, and it is difficult to let the eyes act as a double-role (to see and to trigger). We also found that since eyes are rarely used as an input method, we cannot simply let the user to explore and to understand the system by themselves without being told that the system is sensitive to their gaze.

To design a gaze sensitive object, we need to consider on how to bind two entities (the user and the object) into a singular connected system so both of them can understand each other. We believe that the important research issue is on how to make this *natural interaction* become reliable input method for the user to become familiar with the interaction itself. Hence, we need to understand in which situation people always realize that they need to depend on their eyes to interact. For example, during social interactions, people spend most of their time looking at each other's eyes [8]. It is easier and one the most efficient way to determine the emotional state of the others by simply looking at the eyes [21]. Eyes can also reveal the center of interest and also to build inferences about their mental states [3, 4]. Moreover, the ability to make joint or shared attention with their partner is crucial to make the interaction become more effective [6, 7]. Recent research has also revealed that without making eye contact, it is hard to establish shared attention [15]. Generally, people already know that their eye movements are important in order to socialize with other. They need to be seen in eye to eye. Based on these theories, we conclude that the feeling of being looked back followed by shared gazing is critical in designing for gaze sensitive object, to make people realize that they can socialize or interact with the object by depending on their eyes only. However, is it enough that the object give the feeling of being looked back or the user need to also build the relationship in order to discover the element (gaze as input method) behind the interaction?

MOTIVATION

In this paper, we propose a method to overcome the problem of eyes as double role and also the unfamiliarity face by the user when interacting with a gaze sensitive object. We decide to anthropomorphize a coffee machine by attaching a pair of animated eyes so that it can interact based on the user's eye gaze. As a result of implementing this method, both the user and the object can obtain knowledge about their current engagement. Such information not only aids the user on how to interact with the object, it also provides a significant source of information for determining when the object should execute a proper command based on the user's gaze behavior.

PROTOTYPE

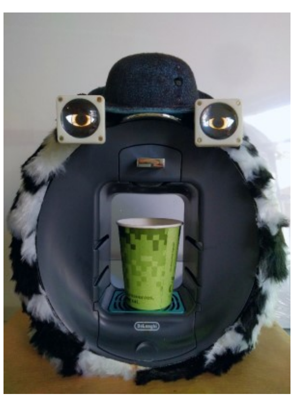

Figure 1. The CoffeePet, a gaze sensitive coffee machine that will start to brew the drink if the user is looking at it.

Figure 1 shows the prototype of a gaze sensitive coffee machine. The design of the CoffeePet's animated eyes is based on the design of the electronic animated eyes created by Burgees et al. [22]. We changed the original version of the eyes to make it more cartoonish and user-friendly. The animated eyes were controlled by a Teensy 3.2 microcontroller board and can be programmed to gaze in any directions. We also modified the mechanical push button of the coffee machine and rewired it to the same microcontroller board which gives the ability for our system to control the coffee machine to start brewing the drink automatically. Furthermore, we decided to cover the coffee machine with fur to avoid the user looking for buttons when interacting with the CoffeePet. To detect and track the user's eye movements, we use Omron HVC sensor module that can estimate the user's eye gaze in real time and we placed the sensor on top of the coffee machine. Since our objective is to let the user to figure out on how to interact with the CoffeePet, we hide the sensor underneath a small hat to avoid the user constructing internal interpretation regarding the function of the sensor. Moreover, to allow the sensor to communicate with our system, we connect it to the same microcontroller board. We also programmed the system to record the user's gaze fixation to get a better understanding of where the user is looking while interacting with the CoffeePet.

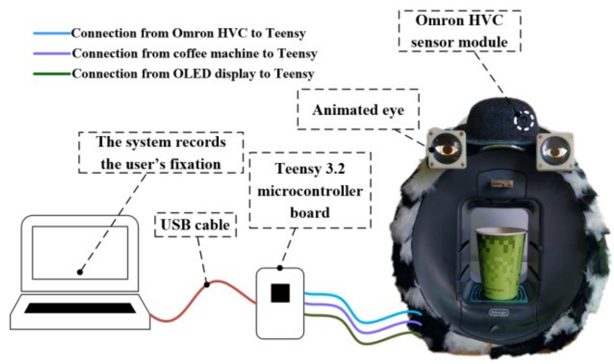

Figure 2. Diagram of the working system.

Figure 2 shows the overview of the systems**.** The sensor module will estimate the user's gaze behavior, and send the data to the microcontroller board. At the same time, the system will record and save the gaze data. The board will process the gaze data and animate the gazing directions according to the user's eye movements. If the user is fixating at the CoffeePet's eyes for a certain time, the microcontroller will activate the coffee machine to start brewing the drink automatically.

EVALUATION

We designed a study to test participant perceptions towards the CoffeePet and the effectiveness of the animated eyes gestures in guiding them to realize that their gaze behavior influenced the CoffeePet to start brewing the drink automatically.

Experimental Design

To evaluate our framework, we conducted a between-participants design involved a single independent variable with three conditions as stated below. Each participant is required to participate in one of these conditions.

Condition 1 (C1): CoffeePet with watching eyes, (static direct gaze).

Condition 2 (C2): CoffeePet with interactive gaze model. In this condition, CoffeePet established eye contact with the participant in between 1 to 2 seconds and averted its gaze randomly (by looking up, down, left or right, in any state) for about 2 to 6 seconds [5].

Condition 3 (C3): CoffeePet with interactive gaze following. CoffeePet looked at the participant when the participant is gazing at the CoffeePet's eyes to established eye contact. When the participant looked away and moved his/her attention to a new target, the CoffeePet will also averted its gaze and looked at the participant's current attention. In this condition, the CoffeePet and the participant needed to coordinate their gaze behavior in order to achieve stable interaction. We also delayed the animation of the eyes for 1 second to let the participants become aware of the direction the CoffeePet was gazing before both of them started to make eye contact again. For example, when both of them were making eye contact, and the participant started to avert his/her gaze to the left, the CoffeePet will also look to the left. When the participant looks at the CoffeePet again (to make eye contact), the CoffeePet's eyes were still gazing to the left and will look back at the participant after 1 second. Hence, the participant will know where did the CoffeePet was gazing before it looked back at him/her.

The fixation duration of 3.5 seconds (dwell time) was chosen based on the empirical data that we gathered from test runs in *C3*, as it gave participants enough time to realize that their gaze behavior influenced the system to react. For *C3*, after the participant and the CoffeePet managed to established eye contact for 3.5 seconds, CoffeePet will pour the hot drink into the paper cup. We implemented the same timing for *C1* and *C2*. For *C1*, the participant needed to fixate at the CoffeePet's static gaze for 3.5 seconds. For *C2*, the participant needed to look at the CoffeePet's eyes for 3.5 seconds even after the CoffeePet averted its gaze.

Hypotheses

We derived three hypotheses based on the existing theory on how gaze cue would affect people during face-to-face communication and how the perception of different gaze cue would guide the participants during interaction with the CoffeePet.

Hypothesis 1: Attaching animated eyes to an object will give participants a fundamental idea that their eye movements are crucial in order to interact with the CoffeePet.

Hypothesis 2: Mutual eye contact followed by shared gazing will make the participants to realize and to confirm that they can communicate using their eyes without being told on how to interact with the CoffeePet.

Hypothesis 3: The participants will manage to control their gaze behavior once they knew that the CoffeePet could also *see* them.

Experimental Setup

Figure 3 shows the experimental setup. We placed the CoffeePet on top of a brown box to make sure that the CoffeePet's eyes are almost on the same level with the participant's eyes. Moreover, we decided to ask the participants to assemble a set of Lego blocks while interacting with the CoffeePet. We placed the Lego blocks in a box in front of the CoffeePet and include an instructions booklet that they can refer to. We wanted the participants to be busy with their hands and to encourage them to interact with the CoffeePet by using their eyes. Compared to our previous setup [1], the benefit of using Omron HVC sensor module is that its ability to compensate head and body movements while estimating the participant's fixation point. Furthermore, the distance between the sensor and the participant can be up to 1.3 meters. With these features, the participants can be more relax during the experiment and we do not have to create any rules on how to interact with the CoffeePet. They can behave as they please. We also do not have to tell them where they should position themselves. We simply give them the flexibility to explore the environment on their own.

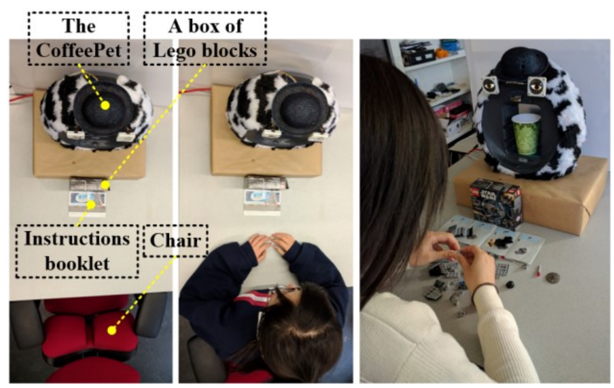

Figure 3. Overview of the experimental setup.

Experimental Procedures

Before starting the experiment, we handed out a written description of the objective and procedure of the experiment to the participants. They were also required to choose one of the ten flavors of the drinks that they would like to enjoy during the experiment. We purposely hide the design of the CoffeePet and participants were not given any information on how to interact with the CoffeePet. The only hint that we gave to the participants was to use any nonverbal communication skills that they found to be reliable during the experiment to attract the CoffeePet's attention to brewed them the chosen drink. The task that the participants needed to do during the experiment was to assembled a Lego set by following the given building instructions booklet. While completing the task, they also needed to interact with the CoffeePet. Participants were randomly assigned to one of three experimental conditions (*C1*, *C2* or *C3*). After reading

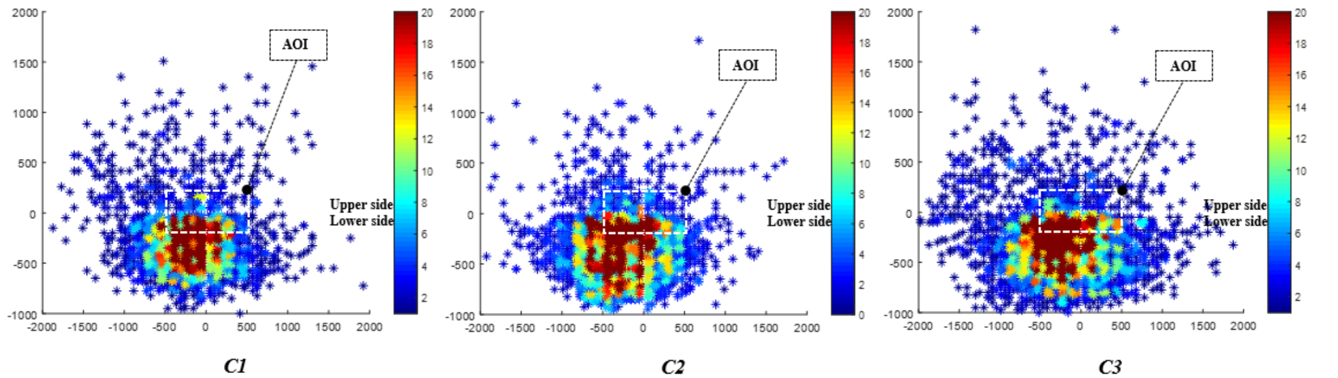

Figure 4: Cumulative heatmap for condition 1 (*C1*), condition 2 (*C2*), and condition 3 (*C3*).

the procedure, participants signed a consent form and the experimenter took them to the experiment room. The experimenter left the participants in the room alone with the CoffeePet and recorded the interaction using a webcam. The experiment lasted six minutes. At the end of the experiment, participants were asked to filled out a post-experiment questionnaire to measured their perception towards the CoffeePet. Finally, the experimenter stayed in the room to interview all participants regarding their experience interacting with the CoffeePet. In this three conditions design, our main interest was to test whether the CoffeePet with interactive gaze following (*C3*) will differ significantly from the CoffeePet with watching eyes (*C1*) and CoffeePet with interactive gaze model (*C2*).

Measurement

Our experimental design involved three manipulated independent variables, (1) watching eyes (static direct gaze), (2) interactive gaze model, and (3) interactive gaze following. The dependent variables involved objective and subjective measurements.

Objective –We recorded the participant's gaze data and the system calculated the fixation duration during the experiment. We defined the CoffeePet's eyes as the area of the interest (AOI). We measured the participants' interaction performance through observing the captured gaze data.

Subjective – We used a scale developed for evaluating life-like interface agents [20] to measure the participant's perception towards the CoffeePet. The questionnaire involved eight items and was made on 7-point Likert-type scales. Four items attributed to the CoffeePet's appearance: human-likeness, attractiveness, sociable, and intelligent. Two items attributed to the partnership with the CoffeePet: mutual liking and trustworthiness. Two items attributed to the level of interaction with the CoffeePet: difficultness and enjoyment.

After answering the questionnaire, we did the manipulation check using open-ended questions and asked the participants to list any nonverbal behavior that they depended on and whether they realized that their gazes influenced the CoffeePet to react. We also conducted a semi-structured interview to gain a richer understanding of participant's experience with the CoffeePet and to confirm the participant's answer and its relation to the recorded gaze data.

Participants

33 volunteers (16 female, 17 male) between the ages of 20 to 35 years participated in the study. None of the volunteers were visually impaired. Of all the volunteers, 25 studied industrial design, 3 studied mechanical engineering, 3 studied electronic engineering and two of the volunteer studied software engineering. All volunteers never saw the design of the CoffeePet before and they had no clue that there was a tiny hidden camera to detect their eye movements.

RESULTS

Objective Measures

We used three main objective measures, (1) the participants' gaze data, (2) the time it took for them to trigger the CoffeePet to start brewing the drink, and (3) the number of participants who managed to get the drink during the experiment.

Heatmap [23] are one of many ways to represent the location of the participant's visual attention and are suitable for evaluating different user group [19]. A color spectrum was used to reveal intensity, make it easier to illustrate the region where the participants fixed their attention the most [17]. Figure 4 displays the cumulative heat map for all 11 participants of each condition. The color changes, from blue to dark red, meaning an increasing number of fixations from lowest to highest. The white, dash square are the AOI (the estimation position of CoffeePet's eye). It can be seen that *C3* presented the highest area of blue color (total fixation count: 4921), compared with *C1* (total fixation count: 3521) and *C2* (total fixation count: 4426) suggesting that the participants were attracted to make more eye fixations compared to other condition. They tried to explore, to interact and to experience more with the CoffeePet's eye in order to understand the underlying mechanisms of the system. These findings were in-line with our qualitative data; most participants reported during the interviews that at first, they did not know that the CoffeePet's eyes are following their gaze behavior. Once they realized it, they started to make more conscious eye movements to investigate and to confirm that the CoffeePet's eyes are influenced by their gaze behavior. After making eye contact for a certain time, the

CoffeePet's started to brew the drink and they knew instantly that eye contact and eye movements were necessary in order to interact with the CoffeePet.

In *C2*, the cumulative heat map shows that there was an increase of tracking activities on the lower side of the graph (fixation count: 2515) perhaps because the participants were paying more attention to doing the task and they were uninterested in looking at the CoffeePet. These findings were further supported by our semi-structured interview; participants reported that they tend to look away from the CoffeePet or looking at the same direction where the CoffeePet are gazing whenever the CoffeePet averted its gaze. Several participants felt that the CoffeePet was rude or look bored because it looks away several times when they attempted to make eye contact so they tried to wave their hands to attract the CoffeePet's attention but there was no obvious feedback and the gaze behavior showed was still random. Hence, they decided to focus more on completing the task rather than to further interacted with the CoffeePet.

As for *C1*, it can be seen from the cumulative heat map that the gaze recorded were more collected around the AOI compared to *C2*, suggesting that the participants were exploring the physical appearance of the CoffeePet. Based on the interview, several participants admitted that they were not willingly to complete the task but rather spent most of the time trying to figure out any hidden mechanism that they should know in order to interact with the CoffePet by looking around the design installation, by gently stroking the CoffeePet's fur, by waving or clapping their hand and even snapping their finger. Some of the participants were also looking behind them to find any components that they should consider since the CoffeePet's eyes was always watching straight ahead. However, the behavior showed by the CoffeePet was always the same which caused confusion and frustration among the participants.

Higher fixation outside the AOI on the lower side of the graph shows that the participants were focusing on the task (gazing downward). According to participants in *C3*, they felt encouraged to finish the task as a result of shared gazing displayed from the CoffeePet (see Figure 5).

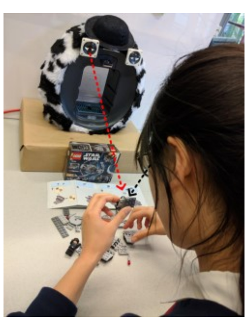

Figure 5. CoffeePet's eyes was looking down whenever the participants were also gazing down to complete the task.

In *C1*, participants were more inclined to ignore the task and spent most of the time trying to figure out how to interact with the CoffeePet. As for *C2*, participants simply gave up and focus more on completing the task rather than to pay attention to the CoffeePet. However, if we look at the cumulative heatmap dispersion in *C2* (on the lower side), the gaze data were not much collected compared to *C3*. The sensor lost its ability to track the participant's fixation point suggesting that participants in *C2* were completing the task on the area where the sensor could not detect their gazing point (see Figure 6). Based on our qualitative data, participants in *C3* reported that they insisted on doing the task on one small area because they wanted the CoffeePet to also *see* what they were doing during the experiment with the intention to develop joint attention between them. Since the CoffeePet does not have any features of body movement, they tried to make sure that they assembled the Lego blocks in front of the CoffeePet's eyes.

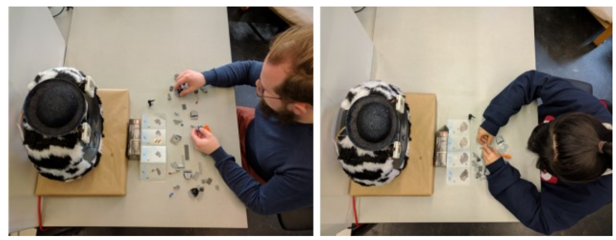

Figure 6. Participant in *C2* (left) and participant in *C3* (right). In *C2*, since participant was looking outside of the tracking area, the sensor lost its ability to track the gaze data.

Figure 7 shows the average of time it took to get the system to react based on the participant's gaze behavior. For example, in *C1*, participants managed to trigger the CoffeePet to start brewing the drink at the average of 2 min. and 46 sec. after the experiment started. *C3* has the fastest average time compared to *C1* and *C2* suggesting that the CoffeePet's behavior showed in *C3* influenced the participant to realize within the average of 2 min. and 12 sec., which nonverbal behavior that they needed to use in order to interact with the CoffeePet. However, we expected to get the result of no significant differences for the three conditions [$F(2,19) = 0.646$, $p = 0.535$] since not all participants in *C1* (rate of success: 5 participants) and *C2* (rate of success: 7 participants) managed to figure out how to interact with the CoffeePet within the duration of the experiment (6 minutes).

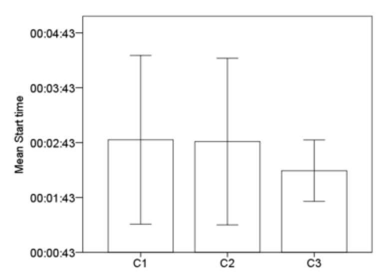

Figure 7. The time it took for participants to trigger the CoffeePet to start brewing the drink.

We continued the analysis by performing a Chi Square test to examine the relation between *C1*, *C2*, and *C3* with the number of drink served during the experiment. The relation between these variables was significant, $CHI^2(1,33) = 9.25$, $p = 0.01$. As can be seen in Figure 8(a), all participants in *C3* were successfully figure out how to make the CoffeePet to start brewing the drink compare to *C1* with 5 participants and *C2* with 6 participants. However, our contingency analysis for the manipulation check showed that significantly fewer participants reported identified gaze as input to interact with the CoffeePet in *C1* and *C2* than in *C3* [$CHI^2(1,33) = 21.928$, $p < 0.01$], as shown in Figure 8(b). Even though 5 participants in *C1* did managed to get the drink but only one participant realize that the CoffeePet was influenced by his/her gaze behavior. According to the participants who managed to get the drink but did not managed to realize gaze as input, they were very confused on why suddenly the CoffeePet started to brew the drink.

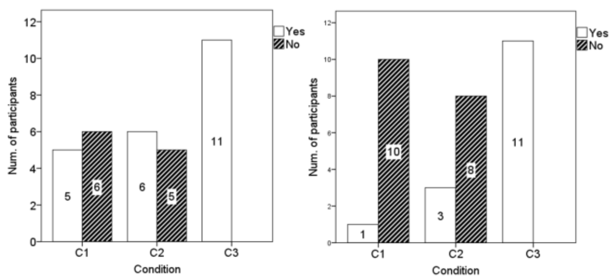

Figure 8. (a) Number of participants who managed to get the drink during the experiment. (b) Number of participants who reported identifying gaze as input to interact with the CoffeePet.

Subjective Measures

A One-way between subjects analysis of variance (ANOVA) was conducted to compare the effect of gaze displayed against the (1) CofeePet's appearance, (2) the sense of partnership and (3) the level of interaction between the participants and the CoffeePet in *C1*, *C2,* and *C3*.

The analysis showed that no differences in participants' rating in human-likeness [$F(2,30) = 1.162$, $p = ns$] for all conditions as can be seen in Figure 9(a). This result is consistent with our qualitative data, participants mainly associated any object that had human characteristics as humanlike. Even though the CoffeePet's eyes were not fully responding towards them in *C1* and *C2*, they somehow felt that the CoffeePet's eyes create a sense of being watched.

On the other hand, the analysis produced a significant effect on the judgments of attractiveness (Figure 9(b)) for the three conditions [$F(2,30) = 3.889$, $p = 0.03$]. Post hoc comparisons using the Tukey HSD test indicated that the mean score for *C3* ($M = 5.64$, $SD = 1.12$) was significantly different than *C1* ($M = 4.27$, $SD = 1.10$). However, *C2* ($M = 4.73$, $SD = 1.27$) did not significantly differ from *C1* and *C3*. These results suggested that the interactive shared gazing displayed by the CoffeePet's eyes really did have an effect on the judgment of attractiveness. Specifically, our results suggest that participants significantly judge the CoffeePet as more appealing when the CoffeePet's eyes were interacting with them. However, it should be noted that the participants needed to realize that the animation of the CoffeePet's eyes was really influenced by their gaze behavior to see an effect. Medium level of interaction (interactive gaze model in *C2*) did not appear to significantly increased the level of attractiveness.

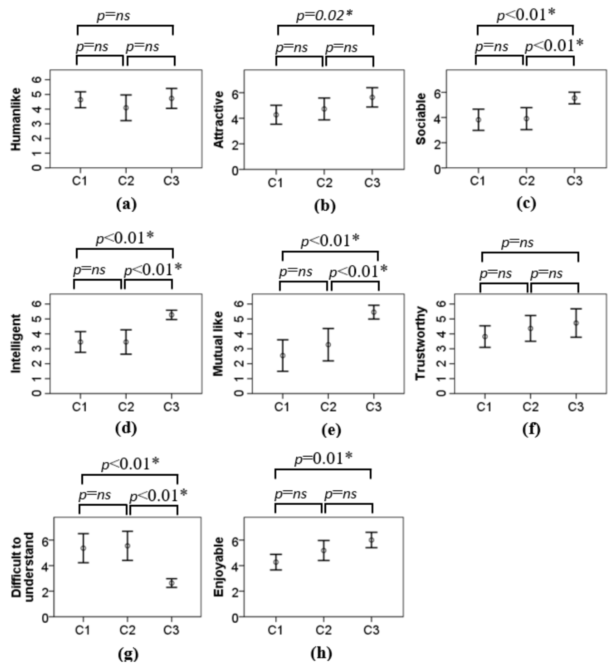

Figure 9: Results on subjective measures: (a) Participants' rating in human-likeness, (b) participants' judgements on the attractiveness of the CoffeePet, (c) participants' judgements on how sociable they feel while interacting with the CoffeePet, (d) participants' rating on the level of the CoffePet's intelligent, (e) participants' judgements of mutual liking, (f) participants' judgement of trustworthiness, (g) participants' rating on the level of difficulty to interact with the CoffeePet, and (h) participants' rating on the level of enjoyment while socializing with the CoffeePet. (*) denotes statistically significant probabilities

As can be seen in Figure 9(c) and Figure 9(d), we also found a significant effect on the judgement of sociable [$F(2,30) = 8.37$, $p < 0.01$] and intelligent [$F(2,30) = 13.16$, $p < 0.01$]. Our post hoc comparisons test for the judgement of sociable indicated that the mean score for *C3* ($M = 5.55$, $SD = 0.69$) was significantly different than *C1* ($M = 3.82$, $SD = 1.25$) and *C2* ($M = 3.91$, $SD = 1.30$). Same goes for the judgement of intelligent, the mean score for *C3* ($M = 5.27$, $SD = 0.47$) was significantly different that *C1* ($M = 3.45$, $SD = 1.04$) and *C2* ($M = 3.45$, $SD = 1.21$). No differences were observed for both judgments of sociable and intelligent across *C1* and *C2*. Taken together, these results suggest that the synchronization behavior showed by the CoffeePet did have an impact on the level of sociable and intelligent. We believe that participants rated the CoffeePet as more sociable because of the CoffeePet's ability to perform shared gazing by looking at the same direction as they were and more intelligent due to the fact that the CoffeePet able to start brewing the drink automatically when the participant made a steady eye contact with the CoffeePet's eyes.

We found a significant effect on the judgement of mutual liking (Figure 9(e)) for the three conditions [$F(2,30) = 13.595$, $p < 0.01$]. The post hoc comparisons test revealed that the mean score for *C3* ($M = 5.45$, $SD = 0.69$) was significantly different than *C1* ($M = 2.55$, $SD = 1.57$) and *C2* ($M = 3.27$, $SD = 1.62$). As expected, the result showed no significant differences across *C1* and *C2*. This test revealed that participants might found it easier to interact with the CoffeePet in *C3* who they believed to give them positive feelings and encouragement to complete the task.

Further analysis showed no significant differences detected on the judgement of trustworthiness [$F(2,30) = 1.429$, $p = ns$] (Figure 9(f)). We believe that since most of the participants in *C1* and *C2* managed to get the drink, we suspected that the participants found the CoffeePet fairly reliable enough to be trusted even though majority of them did not realize on how they managed to trigger the CoffeePet to start brewing the drink.

Figure 9(g) shows a significant effect on the level of difficulty to interact with the CoffeePet for the three conditions [$F(2,30) = 14.650$, $p < 0.01$]. The mean score for *C3* ($M = 2.64$, $SD = 0.50$) was significantly different than *C1* ($M = 5.36$, $SD = 1.69$) and *C2* ($M = 5.55$, $SD = 1.70$). No significant differences were found across *C1* and *C2*. These findings confirm the usefulness of reciprocal gaze behavior displayed between the CoffeePet and the participant in guiding the participants to became aware and making them understand that their eye movements were being used as input to the system.

Moreover, the analysis also produced a significant effect on the level of enjoyment (Figure 9(h)) [$F(2,30) = 8.262$, $p = 0.01$]. According to the post hoc comparisons test, the mean score for *C3* ($M = 6.00$, $SD = 0.89$) was significantly different than *C1* ($M = 4.27$, $SD = 0.91$). However, *C2* ($M = 5.18$, $SD = 1.17$) did not significantly differ from *C1* and *C3*. It is interesting to note that the CoffeePet's eyes following the participant's gazes do increase the level of enjoyment. The results suggested that participants in *C3* were delighted to complete the task while interacting with the CoffeePet since the CoffeePet was showing interest by gazing on the same area the participants were looking during assembling the Lego blocks. Gazed displayed by the CoffeePet in *C2* that are partially interacted with the participants did not appear to significantly increase the level of enjoyment.

DISCUSSION

The present results aimed to discover, under which animated eyes can aid participants to realize that they can interact with the system simply by moving their eyes. In particular, we wanted to explore whether putting eyes could overcome the problem of eyes as double role and the unfamiliarity face by the user when interacting with a gaze sensitive object. We evaluated our hypotheses in three experimental studies with a total combined sample of 33 participants.

Consistent with Hypothesis 1 about the outcome of attaching animated eyes to an object, the CoffeePet's eyes indeed give participants a clue on the possibility of using their eyes to interact with the CoffeePet by making eye contact or eye movements. It was based on their past experience while socializing with others. Even though the animated eyes displayed in the three conditions were different, it demonstrates that the projected eyes undoubtedly spark a sense of being watched among the participants. This suggests that the feeling of being watched or looked back should matter in designing for gaze sensitive object, to give the user a fundamental idea that they can interact with the object nonverbally by using their eyes.

However, attaching watching eyes (*C1*) and partially interacted eyes (*C2*) on an object become useless to the participants in aiding them in discovering what trigger the system to react. As compared to *C1* and *C2*, interactive following gaze (*C3*) led to significantly greater participants reported identified their gaze as input to interact with the CoffeePet. In fact, all participants in *C3* (100%) do realize that they should use their eyes in order to communicate with the CoffeePet as compared to *C1* (9%) and *C2* (27%). Furthermore, participants in *C3* started to grasp the idea of eye movements as input to the system when the CoffeePet's eyes were looking at the same Lego block or at the same area they were fixating. We also found significant differences on the judgment of mutual liking and the level of intelligent in *C3* compared to *C1* and *C2*. These results supported our second hypothesis. Participants performed better in *C3* when the CoffeePet was starting to make eye contact with them followed by sharing a focus on the same object or area. According to our qualitative data, since the CoffeePet was paying much attention to what they were doing at that moment, they felt that the CoffeePet was showing curiosity in joining them to assemble the Lego blocks by means of eye-gazing. And when the participants were showing interest with the CoffeePet by making prolonged eye contact (3.5 sec.), it started to brew the drink, which they consider the CoffeePet to be intelligent. Participants also reported that they were thinking about getting the drink from the CoffeePet, perhaps by making eye contact, the CoffeePet will understand their intention. Here, the participants realized they needed to look at the CoffeePet's eyes and by doing that will trigger the CoffeePet to start brewing the drink. They managed to discover that they can use their eyes not only to observed the environment but also to activate a command (eyes as double role). Furthermore, we found a significant effect on the level of difficulty where participants in *C3* rated interaction with the CoffeePet were easier to understand compared to *C1* and *C2*. Even though the CoffeePet in *C2* did manage to make eye contact with the participants, but once the CoffeePet averted its gaze while they were still looking at the CoffeePet's eyes, participants began to feel that their gaze behavior did not influence the CoffeePet to react. Participants in *C1* and *C2* found that the behavior showed by the CoffeePet was very difficult to understand which led to confusion and frustration.

We also found strong effects on the judgment of sociable and the level of enjoyment in *C3* which supported our third hypothesis. Once they knew that the CoffeePet's eyes were interacting with their gaze behavior, they started to make conscious eye movements to socialize more with the CoffeePet. They thought that the CoffeePet was friendly. Even though the CoffeePet could not speak, they felt that the movements of the CoffeePet's eyes gave a strong indication that it was willing to engage and to interact with them. It is interesting to note that the interactive gaze displayed in *C3* also affected the participants to regulate their behavior by making sure that they completed the task in front of the area where the CoffeePet could also *see*. These result has further strengthened our hypothesis that not even the participants managed to control their gaze behavior, they also managed to act accordingly in order to make sure that the CoffeePet can also get involved in completing the task (sharing gaze), thus increased the level of enjoyment. The ability of the CoffeePet to show interest encouraged the participants to focus more on doing the task and to interact with the CoffeePet at the same time.

Design and Research Implications

It is worth noting that the feeling of being watched or looked back follow by the establishment of shared gazing between the object and the user can be use to overcome the problem of unfamiliarity and eyes as double role in designing for gaze sensitive object. Our study shows that through active exploration during the interaction process, participants manage to build the relationship and to realize that they can interact with the CoffeePet using their eyes which lead to a better understanding and decrease the level of confusion and frustration among the participants. This work also informs that emotional connection between the user and the object could be crucial if we want to use eye movements as input to a system because interaction with eyes mostly leads to confusion and frustration (based on our previous work [2] and also other works [13, 27], etc.). For example, to design an object that is sensitive to our gaze is rather straightforward, but the challenging part is to make people realize the interaction itself (eyes as input). People are more easily relate to a product when they are able to create a bond with the object at a personal level. Hence, it is important to consider how people can build a social relationship with the object within the context of eye gazing. Furthermore, this work extends our understanding of how the participants interpret and react towards the animated eyes in aiding them to realize that the object is responding based on their gaze behavior.

Limitations

The system that we develop does not detect and track the user's face which limits the CoffeePet's eyes to notice if there is someone near it. For example, if the user is somewhere in front of the detection ranges, but not in face to face with the CoffeePet, the CoffeePet does not have the ability to gaze at the person's face to attract that person to position him/herself in front of it. In order to interact with the CoffeePet, the user needs to discover it on their own that they need to be in front of the CoffeePet's eyes which they manage to do so (after some time) during the experiment. Therefore, we plan to implement face detection combine with gaze tracking to make the system more intelligent and to create a more natural interaction between the user and the object.

CONCLUSION

During social interactions, people always looking at their partner's eyes. They already knew that their eye movements play an important role to make the communication become more effective. Based on this scenario, we proposed a method (the feeling of being looked back) to make people discover on their own that they can also depend on their eyes to interact with a gaze sensitive object. To verify our method, we attached a pair of animated eyes on a coffee machine that will brew the drink automatically when the participants were fixating at the animated eyes (making prolonged eye contact). Based on the results, we found that it is not enough for the participants to get the feeling of being looked back, they also need to build the relationship by making shared gazing in order to make them realize that they can use their eyes to interact with the object (like in a normal face to face conversation). While the system does not have the capability to understand if the participants were just observing or to activate a command (eyes as double role), the animated eyes did guide the participants to grasp the idea that in order to make the CoffeePet to start brewing the drink, prolonged eye contact is needed. It was based on their previous experience, making eye contact to show interest (to connect with the system), look away to break the communication.

ACKNOWLEDGMENTS

We thank our colleagues at the Design Intelligence research group who give suggestions for this project. This research received funding from the Ministry of Higher Education, Malaysia and Universiti Teknikal Malaysia Melaka (UTeM).

REFERENCES

1. Anas, S.A. binti, Qiu, S., Rauterberg, M. and Hu, J. 2016. Exploring gaze in interacting with everyday objects with an interactive cup. *In Proceedings of the Fourth International Conference on Human Agent Interaction (HAI '16)* (2016).
2. Anas, S.A. binti, Qiu, S., Rauterberg, M. and Hu, J. 2016. Exploring social interaction with everyday object based on perceptual crossing. *In Proceedings of the Fourth International Conference on Human Agent Interaction (HAI '16)* (2016).

3. Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y. and Plumb, I. 2001. The "Reading the Mind in the Eyes" Test Revised Version: A Study with Normal Adults, and Adults with Asperger Syndrome or High-functioning Autism. *Journal of Child Psychology and Psychiatry*. 42, 2 (2001), 241–251.
4. Bayliss, A.P. and Tipper, S.P. 2005. Gaze and arrow cueing of attention reveals individual differences along the autism spectrum as a function of target context. *British Journal of Psychology*. 96, 1 (2005), 95–114.
5. Bee, N., Wagner, J., André, E., Vogt, T., Charles, F., Pizzi, D. and Cavazza, M. 2010. Discovering eye gaze behavior during human-agent conversation in an interactive storytelling application. *International Conference on Multimodal Interfaces and the Workshop on Machine Learning for Multimodal Interaction (ICMI-MLMI '10)* (2010).
6. Brennan, S.E., Chen, X., Dickinson, C.A., Neider, M.B. and Zelinsky, G.J. 2008. Coordinating cognition: The costs and benefits of shared gaze during collaborative search. *Cognition*. 106, 3 (2008), 1465–1477.
7. Carletta, J., Hill, R.L., Nicol, C., Taylor, T., Ruiter, J.P. de and Bard, E.G. 2010. Eyetracking for two-person tasks with manipulation of a virtual world. *Behavior Research Methods*. 42, 1 (2010), 254–265.
8. Hsiao, J.H. and Cottrell, G. 2008. Two Fixations Suffice in Face Recognition. *Psychological Science*. 19, 10 (2008), 998–1006.
9. Jacob, R.J.K. 1995. Eye Tracking in Advanced Interface Design. W. Barfield and T.A. Furness, eds. Oxford University Press. 258–288.
10. Jacob, R.J.K. 1991. The Use of Eye Movements in Human-computer Interaction Techniques: What You Look at is What You Get. *ACM Trans. Inf. Syst.* 9, 2 (Apr. 1991), 152–169.
11. Jacob, R.J.K. and Karn, K.S. 2003. Eye Tracking in Human–Computer Interaction and Usability Research: Ready to Deliver the Promises. J. Hyona, R. Radach, and H. Deubel, eds. Elsevier B.V. 573–603.
12. Just, M.A. and Carpenter, P.A. 1976. Eye fixations and cognitive processes. *Cognitive Psychology*. 8, 4 (1976), 441–480.
13. Khamis, M., Bulling, A. and Alt, F. 2015. Tackling Challenges of Interactive Public Displays Using Gaze. *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers* (Osaka, Japan, 2015), 763–766.
14. Kleinke, C.L. 1986. Gaze and Eye Contact: A Research Review. *Psychological Bulletin*. 100, 1 (1986), 78–100.
15. Koike, T., Tanabe, H.C., Okazaki, S., Nakagawa, E., Sasaki, A.T., Shimada, K., Sugawara, S.K., Takahashi, H.K., Yoshihara, K., Bosch-Bayard, J. and Sadato, N. 2016. Neural substrates of shared attention as social memory: A hyperscanning functional magnetic resonance imaging study. *NeuroImage*. 125, (2016), 401–412.
16. Land, M.F. and Furneaux, S. 1997. The knowledge base of the oculomotor system. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 352, 1358 (1997), 1231–1239.
17. Lorigo, L., Haridasan, M., Brynjarsdóttir, H., Xia, L., Joachims, T., Gay, G., Granka, L., Pellacini, F. and Pan, B. 2008. Eye tracking and online search: Lessons learned and challenges ahead. *Journal of the American Society for Information Science and Technology*. 59, 7 (2008), 1041–1052.
18. Majaranta, P. and Räihä, K.-J. 2002. Twenty Years of Eye Typing: Systems and Design Issues. *Proceedings of the 2002 Symposium on Eye Tracking Research & Applications* (New Orleans, Louisiana, 2002), 15–22.
19. Mitterer-Daltoé, M.L., Queiroz, M.I., Fiszman, S. and Varela, P. 2014. Are fish products healthy? Eye tracking as a new food technology tool for a better understanding of consumer perception. *{LWT} - Food Science and Technology*. 55, 2 (2014), 459–465.
20. Parise, S., Kiesler, S., Sproull, L. and Waters, K. 1999. Cooperating with life-like interface agents. *Computers in Human Behavior*. 15, 2 (1999), 123–142.
21. Peterson, M.F. and Eckstein, M.P. 2012. Looking just below the eyes is optimal across face recognition tasks. *Proceedings of the National Academy of Sciences*. 109, 48 (2012), E3314–E3323.
22. Philip Burgess. Electronic Animated Eyes using Teensy 3.1/3.2. 2015. Retrieved September 17, 2016 from https://learn.adafruit.com/animated-electronic-eyes-using-teensy-3-1/overview.
23. Reeder, R.W., Pirolli, P. and Card, S.K. 2001. WebEyeMapper and WebLogger: Tools for Analyzing Eye Tracking Data Collected in Web-use Studies. *CHI '01 Extended Abstracts on Human Factors in Computing Systems* (Seattle, Washington, 2001), 19–20.
24. Stampe, D.M. and Reingold, E.M. 1995. Selection By Looking: A Novel Computer Interface And Its Application To Psychological Research. *Eye Movement ResearchMechanisms, Processes, and Applications*. R.W. John M. Findlay and R.W. Kentridge, eds. North-Holland. 467–478.
25. tephanidis, C. 2009. *The Universal Access Handbook*. CRC Press, Inc.
26. Sundstedt, V. 2010. Gazing at games: using eye tracking to control virtual characters. *ACM SIGGRAPH 2010 Courses*. 5, (2010), 1–160.
27. Zhang, Y., Chong, M.K., Müller, J., Bulling, A. and Gellersen, H. 2015. Eye Tracking for Public Displays in the Wild. *Personal and Ubiquitous Computing*. 19, 5 (2015), 967–981.