A Place for Every Tool and Every Tool in Its Place: Performing Collaborative Tasks with Interactive Robotic Drawers

Brian Ka-Jun Mok*, Stephen Yang†, David Sirkin*, Wendy Ju*
*Department of Mechanical Engineering and †Department of Computer Science
Stanford University, Stanford, CA, USA
{brianmok, syang0, sirkin, wendyju} @stanford.edu

Abstract— In this study, we examined how participants (N=20) interacted and collaborated with a set of robotic drawers to accomplish an assembly task. The drawers’ behavior varied along two dimensions – proactivity and expressivity of motions. The results of our study indicate that participants consider an expressive robot to be more involved and interested in the interaction. We also found that while proactive or expressive robots could dominate the interaction, proactivity might negatively affect the participants’ perception of their social status relative to that of the robot’s, while expressiveness did not. This shows the importance of utilizing expressive movements when designing socially appropriate robots that collaborate with human users.

I. INTRODUCTION

Robotic systems, particularly domestic household robots, have become increasingly prevalent in our daily lives. As household robots continue to become more integrated into our homes and workspaces, we will need to interact and collaborate with them. However, Breazeal et al. [1] noted that both robots and human users have traditionally treated each other more like obstacles rather than social figures with whom to collaborate. Additionally, when people engage with robots to collaboratively accomplish a task, the means of interaction are often non-verbal and task-based. Thus, we believe that designing robots that can effectively work as a team with human users requires a deeper understanding of the social and non-verbal collaborative cues used during task-based activity.

Our current research focuses on improving the engagement with interactive furniture and appliances, particularly through gestures and other forms of non-verbal communication. In this study, we explore the interactions between human users and a set of robotic drawers that could collaborate together to accomplish an assembly task. We are interested in two variables that govern the robot’s behavior. The first variable is the robot’s expressiveness, with the robot being either non-expressive or expressive. Breazeal et al. [2] noted the importance of a robot’s expressivity in providing effective communication. In our study, we defined a non-expressive robot as one that performed only basic and fundamental movements to help complete a task. Conversely, an expressive robot carried out additional movements and animations for the purpose of communication. The second variable is the proactivity of the robot, with the robot being either reactive or proactive. Fink et al. [3] and Mubin et al. [4] indicated that proactivity could be an important factor, especially when a robot worked together with human users. We defined a reactive robot to be one that responded and executed actions according to the user’s gestures, while a proactive robot anticipated a user’s impending needs and initiated an action to complete a task. To explore the effects of these two dimensions, we performed a controlled study.

II. BACKGROUND

A. Related Work

There has been extensive research on different methods of improving the social interactions between robots and human users. One branch of research explores the use of anthropomorphism. Fink et al. [5] incorporated several anthropomorphic elements into a robot to create more positive interactions. Osawa et al. [6] examined the utilization of anthropomorphic features and frameworks for domestic robots. The study found that the implementation of anthropomorphism could help robots facilitate better communication with human users. However, one of the challenges with using anthropomorphism in task-based applications is that the form factor of the robot is often dictated by the task. When designers have limited ability to add any humanoid features and forms to the robot, the anthropomorphic approach becomes problematic.

Another branch of research focuses more on gestures and different forms of nonverbal communication. Hoffman et al.
indicated that most people are highly sensitive to physical movements. Hence, robots with well-designed motions might engage and communicate more effectively with human users. Sirkin et al. [8] examined the use of on screen and off screen motions for telepresence robotics. Their research found that the additional motions displayed by the telepresence robots improved the understanding of messages and the sense of collaboration. Sharma et al [9] and Szafir et al [10] examined the ability for flying robots to communicate intent through the use of expressive motion. Depending on the application, the ability to perform motions may already be present in the robot for pragmatic purposes. Therefore, it may be possible to augment these robots to have additional communicative and expressive purposes and abilities.

B. Prior Work

In our previous study (Mok et al. 2014 [11]), we examined how to create socially appropriate interactions for non-anthropomorphic robots, particularly with interactive furniture and appliances. We are interested in how the robot should behave in response to a human partner’s actions and emotions. We used video prototypes that explored the interactions between a human user (actor) and a set of robotic drawers. Through gestures and nonverbal communication, we had the human actor and the robot each displayed one of five possible emotional states – Angry, Happy, Indifferent, Sad, and Timid. These video prototypes were then utilized in an online web-based survey with participants recruited from Amazon Mechanical Turk, a crowdsourcing service. After watching the video prototypes, the participants (N=40) then answered questions pertaining to how they perceived the interactions.

Participants significantly preferred empathetic drawers that displayed emotions matching those of the human actor to the drawers that displayed emotions orthogonal to the user’s emotions. Although neutral drawers were not preferred to empathetic drawers, they were significantly preferred over drawers displaying orthogonal emotions – for instance, drawers that acted happy when the human actor was sad. The results were reminiscent of those of Nass et al. [12] in which emotional pairing of a car navigational system’s voice to the human driver’s current emotional state significantly affected the performance of the driver. This reveals the importance of designing robots to display appropriate emotions.

C. Wizard of Oz Techniques

It is our goal to evaluate how people in a household environment will interact with our robot. To do so, we use Wizard of Oz techniques to conduct our research. By utilizing a trained human tele-operator, we can examine many of the questions we have about how people prefer a robot to behave without information that is critical to have prior to developing an autonomous system. Prior work in this field examining best practices and public acceptance in interaction, including Weiss et al. [13], has demonstrated the usefulness of Wizard of Oz in exploring user experience factors. Similarly, Fink et al. [3] used Wizard of Oz techniques to observe how a robot could interact with children, particularly exploring how to encourage children to clean up their rooms and put away their toys.

III. METHODOLOGY

A. Physical Prototype

The robotic drawers prototype was fabricated from a standard Ikea MICKE 4-drawer unit (Fig. 1). To allow the drawers to perform consistent and repeatable motions, the top 3 drawers were retrofitted with DC motors on a rack and pinion system. Additionally, spring-loaded rotational encoders, mounted against the drawers’ frame, allowed us to track each drawer’s position. The bottom drawer contained parts and hardware that drove the system, and was not used in this study. The actuation was controlled by an Arduino microcontroller communicating with a local client program over a USB cable.

The local client program provided a Wizard of Oz style remote control for the drawers. A human “Wizard” in an adjacent room could remotely operate the drawers prototype via keyboard hotkeys (Fig. 2). The UI presented the Wizard with 15 buttons and each button corresponded to a pre-programmed sequence of actions, or “animations” for the drawers to execute. The Wizard had a one-way audio/video feed of both the drawers and participants. This was needed for the operator to control the drawers appropriately.

B. Study Design

This research study explores how a robot’s expressivity and proactivity affect human-robot turn-taking. To examine this, we devised a Wizard of Oz study involving a set of robotic drawers. We conducted a between-subjects test with four scenarios. Each scenario followed the same script, but varied two factors yielding a 2x2 factorial design. The first variable is the expressivity of the drawers. The non-expressive set of drawers only opened and closed drawers whereas the expressive set of drawers expressed sentiment with additional animations. The second variable is the proactivity of the drawers. In the proactive case, the drawers led by initiating interactions. In the reactive case, the drawers waited for gestures by the participants before reacting. Below we describe the drawers and the study.

Figure 2: Diagram of study setup. The Wizard monitors and manipulates the Drawers from the observation room. The Participant collaborates with the Drawers in the study room.
C. Animations

We want to have robotic drawers that appear to be either non-expressive or expressive. Thus, animations were developed accordingly to achieve this effect. The animations were designed through an embodied design improvisation process with motion experts, in a process described in Hoffman et al. [7] and Sirkin et al. [14]. For a non-expressive drawer, we wanted to use the most basic and functional animations to interact with the participants. Therefore, there were only four animations for these conditions. Three of these animations were used to perform a Simple Open (Fig. 3a) for each of the three drawers. For this animation, each individual drawer opened at a constant speed. The last animation was used to perform a Simple Close (Fig. 3b), with all the drawers closing at the same constant speed and locking after closed.

For the expressive drawers, we added flair to the closing animation and more communicative gestures. The closing animation was now specific to the drawer that was currently opened. This would cause the other drawers to open and close slightly at the end of the animation, mimicking a ripple effect. This is called the Close Flair (Fig. 3c) animation. We also added animations that suggested the drawers wanted something put back in or taken out of an opened drawer. To indicate the former, the drawers had a Wiggle (Fig. 3d) animation in which an opened drawer wiggled in and out slightly two times. To indicate the latter, there was a Beckon (Fig. 3e) animation in which an opened drawer closed halfway and reopened at half speed. Lastly, we also added two animations to suggest sentiment, Chuckle and Happy. In the Chuckle (Fig. 3f) animation, all the drawers opened to random positions, wiggled twice, and closed. The Happy (Fig. 3g) animation involved all the drawers mimicking a ripple effect traveling down the drawers twice.

D. Participants

We recruited a total of 20 participants from the Stanford undergraduate and graduate student population and collected data over five days. Exactly 35% of participants were male and 65% were female. The ages of the participants ranged from 18 to 26 (M=20.8, SD=3.03). Each session required an average time of 45 minutes to complete and participants were compensated with a gift certificate.

E. Procedure

The participants were asked to take a seat at the construction station (Fig. 2). They were then given an introduction to the study. The task for the participants was to assemble a cube made from six acrylic sheets and an 80/20 aluminum frame (Fig. 4). Each face of the cube required the participants to use a different type of fastener (Flathead, Hex Bolt, Philips, Spanner, Socket Cap, and Security Torx) to attach it to the frame. The participants were informed that the tools needed to complete this task were stored inside the top three drawers. The location of each tool was marked with white outlines inside the drawers so that these tools could be returned to their proper place by the participants.

Depending on whether the drawers were in the proactive or the reactive condition, the participants were told that “the drawers will provide you with the tools that you need” or “you may gesture to the drawers to acquire the tools that you need,” respectively. Participants were also requested to take out a maximum of two tools at a time and to build the cube by following the instructions in order. (In earlier pilot testing, participants often took out all the tools at once, which greatly changed the participants’ interactions with the drawers. Instilling these constraints allowed the interactions to be more structured and it also ensured some degree of interaction occurred between the participants and the drawers). After all the instructions, the participants were advised to begin the task. When it was finished, participants were given a questionnaire and an interview.

Figure 3: Animations of Drawers: (a) Simple Open (b) Simple Close (c) Flair Close (d) Wiggle (e) Beckon (f) Chuckle (g) Happy

Figure 4: Acrylic and aluminum cube assembled using 6 different types of fasteners and tools
**F. Wizard’s Responsibilities**

Throughout the study, only one Wizard, who was trained to strictly follow the protocol, was used. The Wizard was tasked with actuating the set of drawers and performing the appropriate animations at the proper times. This was done by using the flow chart in Fig. 5. Given that the participants followed the order of tasks specified in the instructions and the “two tools” rule, the Wizard needed to perform at least eight total tool exchanges (take and return). How these interactions specifically occurred would depend on the expressivity and proactivity conditions of the drawers.

Along the first variable of expressivity, the Wizard had to use different animations to convey the condition. In the nonexpressive conditions, the Wizard used the Simple Open animation when a tool needed to be taken from or returned to the drawer by the participants. After the tool was taken or returned, the Wizard then used the Simple Close animation. However, when a tool needed to be taken from or returned to an expressive drawer, the Wizard opened a drawer using the Simple Open animation followed by the Wiggle animation. Similarly, the Wizard used the Close Flair animation instead of the Simple Close. In addition, the Wizard could make the drawers express certain sentiments. If the participants at any time made a mistake, such as dropping a fastener or picking up the wrong tool, the Wizard used the drawers’ Chuckle animation. Also, when the drawers congratulated the participants for completing milestones, the Wizard used the Happy animation. This was done when the frame was assembled and when the cube was completed.

On the second variable of proactivity, the Wizard changed the timing of the drawers’ behaviors. For the proactive, the Wizard anticipated when the participants required a tool (based on where they were in the instructions) and opened that drawer (with the tool needed) in advance. The Wizard then closed the drawers when the participants took the tool. If the participants had two tools already and required a third one, the Wizard opened one of the two former drawers to ask for a tool back. In the reactive condition, the drawers only opened and closed based on the participant’s gestures. So, the participants needed to indicate to the drawers which one they wanted to open. The Wizard then opened the drawer and kept it open until the participants gestured for it to be closed or after 10 seconds had passed. Hence, the participants in the reactive scenario had more freedom as they were allowed to both request for a wrong tool and to violate the “two tools” rule. The participants in the proactive scenario were coerced to follow the drawers’ lead, thus, they could not violate the above rule or request an incorrect tool.

**G. Questionnaire**

To help evaluate how the participants perceived their interactions with a set of drawers, we adapted and utilized the Dillard’s Relational Message Scale from Dillard et al. [15]. The relational factors that we incorporated were immediacy, affect, similarity, receptivity, composure, dominance, and equality. We adapted 26 statements regarding the participants’ interactions with the robot from the above relational factors. Participants were asked how much they agreed or disagreed with each of these statements based on a Seven-Point Likert Scale (1=Strongly Disagree, 7=Strongly Agree). Additionally, all of the participants’ open-ended comments were also collected.

**IV. ANALYSIS AND RESULTS**

First, the sets of survey responses were imported into R for statistical analysis. We used a two way ANOVA Test to determine if either the proactivity variable or the expressiveness variable had a significant effect on how the robot interactions were perceived by the participants. Some questions that appear to show a strong or significant difference in response with respect to either proactivity or expressiveness are summarized in the following tables. Statistically significant results are highlighted in bold. Several other suggestive, though not statistically significant (to p<0.05), results are also included. As many questions did not show a strong difference, analysis on the relational factors as an indices was not performed.

**A. Immediacy and Affect**

<table>
<thead>
<tr>
<th>Question</th>
<th>Proactivity</th>
<th>Expressiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td>The drawers found the interaction stimulating.</td>
<td>$F(1,16) = 1.39$ $p = 0.256$</td>
<td>$F(1,16) = 20.1$ $p &lt; 0.01$</td>
</tr>
<tr>
<td>The drawers were interested in interacting with me</td>
<td>$F(1,16) = 0.138$ $p = 0.715$</td>
<td>$F(1,16) = 3.45$ $p = 0.082$</td>
</tr>
</tbody>
</table>

*Table 1: Immediacy and Affect Relational Factors*

Questions from the Immediacy relational factor are related to the engagement and involvement of the interaction. Similarly, those of the Affect relational factor are related to the interest of the interaction. We can see that drawers exhibiting expressiveness are perceived by the participants to be more engaged and interested in the interaction compared to non-expressive drawers. In response to the statement, “The drawers found the interaction stimulating,” expressive drawers have greater Immediacy than non-expressive drawers (Fig. 6).
Questions from the Similarity relational factor are related to the familiarity and friendliness of the interaction. Again it appears that the expressive drawers are perceived to be more familiar and friendlier than non-expressive drawers. When reviewing the question “The drawers seemed to desire further interaction,” drawers that displayed expressive motions appeared to participants as having greater similarity (Fig. 8). As with Immediacy and Affect, the proactivity again does not appear to produce a significant difference.

Figure 6: Responses to “The drawers found the interaction stimulating” with respect to condition.

Additionally, regarding whether “The drawers were interested in interacting with me,” the expressive drawers similarly have greater Affect than non-expressive drawers (Fig. 7). Overall, proactivity does not cause a significant difference in the perception of Immediacy and Affect.

Figure 7: Responses to “The drawers were interested in interacting with me” with respect to condition.

B. Similarity

<table>
<thead>
<tr>
<th>Question</th>
<th>Proactivity</th>
<th>Expressiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td>The drawers seemed to desire further interaction.</td>
<td>F(1,16) = 0.023 p = 0.882</td>
<td>F(1,16) = 14.2 p = 0.017</td>
</tr>
</tbody>
</table>

Table 2: Similarity Relational Factor

C. Domination

The Domination relational factor and its questions help determine the leader or influential figure in the interaction. We can see that either exhibiting expressive or proactive behaviors can lead to the drawers appearing more dominant and influential. This is expected for proactive drawers since they are always in a leading role and do not yield in response to the participants’ gestures. For expressive drawers, participants may feel that the animations are trying to get their attention, thus influencing them. Expressivity shows a strong difference in regards to the questions of “The drawers tried to dominate me” (Fig. 9) and “The drawers did not attempt the influence me” (Fig. 10). Note that the question “The drawers did not attempt to influence me” is reverse coded, with larger value indicating more dominance.

Figure 8: Responses to “The drawers seemed to desire further interaction” with respect to condition

<table>
<thead>
<tr>
<th>Question</th>
<th>Proactivity</th>
<th>Expressiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td>The drawers tried to dominate me.</td>
<td>F(1,16) = 3.35 p = 0.086</td>
<td>F(1,16) = 6.84 p = 0.019</td>
</tr>
<tr>
<td>The drawers did not attempt the influence me.</td>
<td>F(1,16) = 18.8 p &lt; 0.01</td>
<td>F(1,16) = 26.6 p &lt; 0.01</td>
</tr>
</tbody>
</table>

Table 3: Domination Relational Factor
The Equality relational factor describes whether the participants felt that they were being treated on an equivalent social standing/status or not. Despite the results from the Dominance relation factor, we can see that expressiveness does not have an effect, with no strong difference in equality. Conversely, proactivity has a very notable difference. Regarding “The drawers considered us equal,” proactive drawers created a greater feeling of inequality to the participants compared to reactive drawers. Participants felt that the proactive drawers did not treat them as equals in social standing, while reactive drawers did (Fig. 11).

D. Equality

<table>
<thead>
<tr>
<th>Question</th>
<th>Proactivity</th>
<th>Expressiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td>The drawers consider us equals</td>
<td>F(1,16) = 7.94</td>
<td>F(1,16) = 0.025</td>
</tr>
<tr>
<td></td>
<td>p = 0.013</td>
<td>p = 0.878</td>
</tr>
</tbody>
</table>

Table 4: Equality Relational Factor

V. DISCUSSION

One of the findings of this research is that expressivity and proactivity greatly affect the robot’s persuasiveness and influence on participants. In addition, we see a strong interaction effect between proactivity and expressiveness in the relational factor of Dominance. Proactive drawers cause participants to feel that the robot does not treat them as equal in social status (as shown by the strong difference in the Equality relation factor). From post study interviews, participants report that the robot appears to be “like a boss” figure, relegating participants to a lower social standing. In one case, a participant indicated that “I was the builder, the drawers should not command me to do things. I will do it when I am ready.” Another participant felt that “It was distracting when I was trying to understand what’s going on, I knew it was trying to get me to get the tool for the next step, but I didn’t know what to do yet. So I just ignored it until I was ready.” Conversely, when the robot exhibited expressiveness in addition to proactivity, it did not create this feeling of inequality in participants interacting with proactive drawers. One participant noted that “It was like a fiery little Scottish Terrier trying to pull me its way.” Several other participants also noted that the robot was “just like a pet.” So, by incorporating an expressive nature into its actions, the robot can still help lead interactions while making the participants feel like an equal. Our study suggests that this is important as perceptions of the robot’s inappropriate performance regarding status can cause frustration and discomfort for the human user.

Through post study interviews, we also found that the participants are able to readily determine the intents and messages of the robot’s actions. As noted in some of the quotes above, participants know that the robot is trying to
get them to move onto the next step by either asking them to take or return a tool. However, as also noted, the participants wanted to work on the task at their own pace. So, they often ignored the robot's suggestions until they were ready to move onto the next step. Another important insight that we observed is the effectiveness of the expressive movement displays, despite what the participants were focused on. For example, from reviewing the video data, participants were often looking at the cube when a Happy animation was occurring. However, in the post study interview, participants consistently recalled seeing the animation, with a large percentage realizing that “it was trying to congratulate me.” Similar observations are seen with the Chuckle animation. One participant noticed that “it was warning me of an error.” Even without directly looking at the expressive drawers, participants still experienced the drawers’ movements. This finding reaffirms Hoffman’s [7] argument that humans are sensitive to movement and that small well designed motions can be used for effective communication.

In addition, we found another insight from monitoring the participants’ interactions with the reactive drawers. As they were not prompted with instructions about how they should interact with the drawers, the participants used a variety of different physical gestures to communicate with the drawers. These include touching, waving, and performing swiping motions. Some participants even tried talking to the drawers in addition to gesturing. It is also interesting to note that participants quickly create a mental model of how to interact with the drawers and do not try different gestures or motions, even when the drawers do not immediately react.

VI. CONCLUSION

This Wizard of Oz controlled study shows that having a robot perform expressive movements greatly affects how an interaction is perceived by the user. Expressiveness allows the robot to seem more engaged, involved, and interested in the interaction as compared to a more functional non-expressive robot. Similarly, the expressive robot also appears friendlier to the participants. Proactivity seems to have no effect on how engaged the robot appears, but a reactive and expressive robot does appear to the participant to be like a good friend. We also note that a proactive robot seems to be more dominant than reactive robots, to the point where the participants feel a sense of inequality. However we find that expressiveness can help temper this feeling, as an expressive robot can seem dominating but still of a similar social standing. These findings can facilitate in the formulation of design guidelines for robots that will collaborate with human users.

ACKNOWLEDGMENT

We would like to acknowledge Nik Martelaro, Jamy Li, Daniel O’Shea, Ashley Mills, Brandon Hightower and Courtney Yang for their efforts and contributions to this study. This research is funded and supported by the Hasso Plattner Design Thinking Research Program.