Robotic Presence: The Effects of Anthropomorphism and Robot State on Task Performance and Emotion

Lawrence H. Kim[1], Veronika Domova[2], Yuqi Yao[3], Chien-Ming Huang[4], Sean Follmer[5], and Pablo E. Paredes[1]

***Abstract*—Robots are becoming more ubiquitous in our daily lives including in our offices and homes. As such, it is necessary to understand whether and how robotic presence affects us both emotionally and cognitively. In this work, we investigate the effects of anthropomorphism and robot state of robotic presence on task performance and user emotion. In an online study, 113 participants completed three tasks within a virtual 3D environment in the presence of robots with varying anthropomorphism (non-anthropomorphic vs. anthropomorphic) and robot state (active vs. idle) in addition to alone and with a human. The study findings suggest that having an active moving robot can expedite the task performance especially for easy difficulty levels albeit with some decrease in accuracy, while anthropomorphism affects the dominance axis of affect and has interaction effects with the robot state. The robot state also had as much influence on the perceived anthropomorphism and animacy as the anthropomorphism of the robot. These results provide initial insights on the effects of robotic presence and the impact of the robot state and anthropomorphism.**

***Index Terms*—Design and Human Factors, Social HRI, Robot Companions, Human-Centered Robotics, Human Performance Augmentation**

I. INTRODUCTION

RECENT advances in robotics and artificial intelligence allow automation of a wide range of human activities in their professional and daily routines. While earlier research efforts were predominantly focused on how robots can help people complete dull, dirty, and dangerous tasks [1], more recent efforts explore whether robots could play an important role in enhancing human learning including cognitive [2], behavioral [3], emotional [4], and motor learning [5].

In addition to the scenarios where robots and users interact directly, HRI researchers started to investigate whether and how the mere presence of a robot can affect people cognitively and emotionally. In particular, researchers successfully replicated the *social facilitation* effect with the presence of an anthropomorphic robot instead of a human [6], [7], [8], the effect where the presence of an observer facilitates one's performance for easy or familiar tasks while inhibiting it for hard or unfamiliar tasks. Emotionally, robotic presence led to users reporting higher sense of being monitored than either human presence or alone [7].

However, in prior work that explores using robots for social facilitation, most researchers leverage solely anthropomorphic robots that are largely still that do not demonstrate significant body movements besides facial expressions (e.g., blinking eyes). In practice, anthropomorphic features in co-presented commercial robots, such as vacuum cleaners, are not always available. This is because the limbs and head do not necessarily help them perform their duties in domestic environments, stores, hospitals, museums, and hotels. Furthermore, in some contexts, anthropomorphic robot designs are not recommended because they can induce false expectations in users [9]. With respect to movement, many co-presented commercial robots are on the move while performing their tasks (e.g., vacuum cleaners, delivery robots, assistants). In addition, both the anthropomorphism and movement of an agent have been shown to significantly affect human perception. For instance, the Uncanny Valley show that humanoid objects that imperfectly resemble human beings provoke feelings of uneasiness and revulsion [10] while motion can induce perception of animacy even for simple geometrical figures [11]. For these reasons, HRI community will benefit to know whether anthropomorphism and motion are influential features in causing the social facilitation effect. To our knowledge, no existing study employed non-anthropomorphic moving robots in social facilitation studies or investigated how the anthropomorphism and motion of the agent influence the social facilitation effect.

Building on the exploration of robotic presence and its effect on nearby humans, we explore how different aspects of the robot, such as its anthropomorphism and its state (active vs. idle), mediate its effect on people. To study the effects of anthropomorphism and robot state on people's task performance and emotion, we ran a study where 113 participants performed three cognitive tasks within a virtual

---

Manuscript received: January, 14, 2022; Revised April, 17, 2022; Accepted June, 1, 2022.

This paper was recommended for publication by Editor Angelika Peer upon evaluation of the Associate Editor and Reviewers' comments. This work was supported by Stanford Transforming Learning Accelerator.

[1] Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, CA, USA `[lawkim,pparedes]@stanford.edu`

[2] Department of Mechanical Engineering, Stanford University, Stanford, CA, USA `vdomova@stanford.edu`

[3] School of Education, Stanford University, Stanford, CA, USA `yuqiyao2@stanford.edu`

[4] Department of Computer Science, Johns Hopkins University Baltimore, MD, USA `cmhuang@cs.jhu.edu`

[5] Department of Mechanical Engineering, Stanford University, Stanford, CA, USA `sfollmer@stanford.edu`

Digital Object Identifier (DOI): see top of this page.

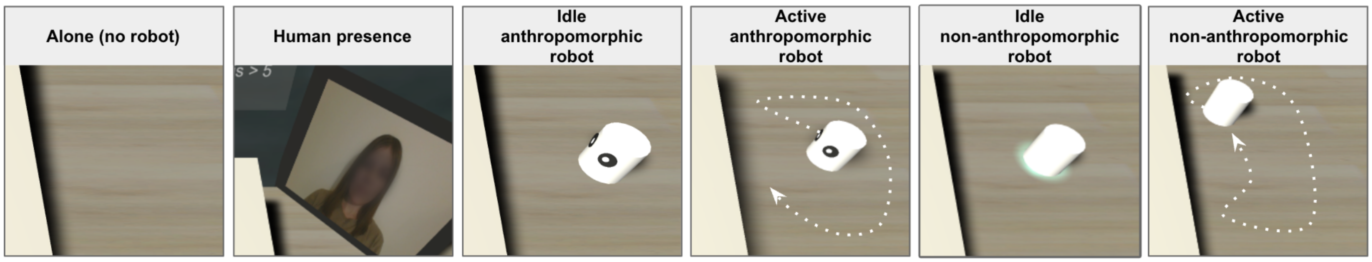

Fig. 1. Different presence conditions presented to participants to test the effects of anthropomorphism and state of a robot on social facilitation.

environment in the presence of a robot placed next to the task. We tested binary levels of anthropomorphism (non-anthropomorphic vs. anthropomorphic) and robot state (idle vs. active), and compared them with two baseline conditions (i.e., alone and human presence). In addition to the task performance-related measures, participants also self-reported their feelings, such as their emotion and sense of being judged, and their perception of the robots, such as desirability, noticeability, anthropomorphism, and animacy. The study findings suggest there are many interaction effects between different parameters, such as robot state, anthropomorphism, and task difficulty, and warrant further investigation.

II. RELATED WORK

A. *Influence of Robotic Presence on Task Performance*

The use of non-human agents has been explored in a variety of contexts, such as remote learning [12], physical exercises [13], remote collaboration [14], and digital gaming [15]. In educational scenarios, like knowledge acquisition and skill building, robots have been assigned or framed as different roles, such as a tutor [16], an assistant [16], a coach [17] or a guide who encourages participants to engage in the tasks [18]. Grounded in the social facilitation theory, a majority of the previous studies focused on anthropomorphic robots who mimic human facial expressions and behaviors [6], [7], [8]. Wechsung et al. investigated whether the anthropomorphism of robots influenced people's task performances and found that presence of a more human-like robot led to worse performance [6]. However, all robots were static, which lessens user's perceived animacy of the robot, and no comparisons were made with a baseline (i.e., alone or human presence). In our study, we explore the effects of the robot state and anthropomorphism of a robot and compare them with two control conditions: alone and human presence.

B. *Human Perception of Robot Motion*

Motion, especially self-produced motion, can be extremely influential in how humans perceive the object in motion (e.g., animacy) and can convey wide range of information (e.g., emotion, stories). Findings from earlier research support the idea that the brain has innate sensitivity for the detection of self-produced motion, which provides one of the most powerful cues about whether an object is animate [19]. Back in the 1940s, the famous experiment of psychologists Fritz Heider and Marianne Simmel showed how informative motion can be in terms of being able to convey stories and enhance animacy of even simple geometric shapes [11].

HRI researchers have echoed the importance of motion with a simple mobile robot. For example, experiments on different test robot platforms have shown that speed and acceleration can be used to elicit feelings along the arousal axis in the circumplex model of affect [20], [21], [22]. On the other axis for valence, smoothness, roundness and perceived stability of movement have been found to be the relevant motion features [20], [22], [23]. Despite the importance of motion due to increasing adoption of mobile robots in daily lives (e.g., vacuum and delivery robots), little to no prior studies have investigated how the presence of a mobile non-anthropomorphic robot implicitly affects a human's task performance and emotion during a task.

C. *Social Facilitation*

There are three main categories of theories that have evolved over time to explain why social facilitation occurs (see Guerin [24]): drive/arousal theories, social comparison theories, and cognitive process theories. Building upon Hull-Spence Drive Theory [25], Zajonc [26] suggested that the mere presence of others during task performance increases one's level of arousal. When an individual is performing a familiar or well-learned task, the presence of others enhances performance of the task, since the dominant response is correct performance. But when a novel or difficult task is undertaken, the presence of others hinders performance, since the dominant response is not appropriate.

Cottrell [27] further suggested that the mere presence of others may not be sufficient for increasing an individual's drive/arousal level. She proposed that increased drive/arousal level is influenced by the presence of other people who elicit evaluation apprehension. This theory is our main motivation to explore if the anthropomorphism of the co-present robot affects evaluation apprehension the participants feel and thus reduce social inhibition observed for hard tasks.

Cognitive process theories explain social facilitation in terms of physical distraction [28], attention conflict [29], [30], and restricted focus of attention [31]. For instance, Baron et al. proposed that in the presence of others, attending to the other person conflicts with attending to the task, which is potentially responsible for the arousal of the subject [30]. The attentional emphasis suggests that distraction may have effects on cognition, attitude change, and social behavior.

III. METHOD

To explore the effects of robot state and anthropomorphism on user's task performance and emotion, we designed an experiment where participants performed three cognitive tasks within a virtual 3D environment under varying conditions, either alone, with human presence, or with robots of varying anthropomorphism and robot state.

A. Hypotheses

Based on the evaluation apprehension theory [27], it is the presence of another person who elicits evaluation apprehension that leads to a heightened sense of arousal and lower performance for hard tasks. Given that the even presence of humanoid robot was able to reproduce the social facilitation effect [7], [8], we hypothesize that the anthropomorphism of the robot will affect user task performance and emotion:

H1a. presence of a non-anthropomorphic robot will lead to higher accuracy and shorter completion time than in the presence of an anthropomorphic robot for hard tasks, and

H1b. presence of a non-anthropomorphic robot will lead to lower arousal and dominance but higher valence for users than when in the presence of an anthropomorphic robot.

While Wechsung et al. suggest that higher anthropomorphism leads to higher error rate [6], they utilize static robots with varying anthropomorphism and thus may not elicit the same sensations as dynamic robots. As self-produced motion has been shown to be highly important in perceived animacy of an object [19], [11] and motion speed is correlated with perceived arousal [20], [32], exploring the effects of robot motion on user's task performance and emotion is necessary. In particular, we hypothesize that:

H2a. An active (i.e., moving) robot will lead to higher accuracy and shorter completion time than an idle (i.e., still) robot, and

H2b. An active robot will lead to higher arousal and dominance ratings than an idle robot.

B. Independent Variables

In addition to the two baseline conditions (i.e., alone and human presence), we varied three independent variables: anthropomorphism, robot state, and task difficulty.

1) Anthropomorphism: With respect to anthropomorphism, we designed two conditions: anthropomorphic robot and non-anthropomorphic robot. For this study, we grounded the robot design on simple table-top robots like the Zooids [33]. For the non-anthropomorphic version of the robot, we took their original design, a simple cylindrical shape, which is similar to simplistic design of commercially available domestic robots, such as a vaccuum robot.

For the anthropomorphic robot, we added eyes with blinking animation at a rate similar to that of the human's to the non-anthropomorphic robot design [34]. While Phillips et al. elicited four distinct appearance dimensions that characterize anthropomorphic robots such as Surface Look (e.g., gender, skin, eyelashes), Body-Manipulators (e.g., torso, arms, hands), Facial Features (e.g., head, eyes), and Mechanical Locomotion (i.e., wheels, treads/tracks) [35], facial features receive considerable attention for social robots because most non-verbal cues and communicative functions are mediated through the face [36], such as facilitation joint attention via gaze orientation [37]. Thus, we decided to add only the eyes because the presence of eyes is the most significant facial feature on robot heads for perception of humanness [38]. There are many similar head designs with only the eyes in many of the existing robots, such as Amazon Astro and Jibo.

2) Robot State: For the robot state, we implemented two states: idle and active. For the idle state, we placed the robot in the middle of the predefined area. When the non-anthropomorphic robot is idle, smooth glowing is added to the bottom of the robot to indicate that it is in a stand-by mode. The robot glows approximately 12 times per minute, the same rate as that of human's eye blink and that of the anthropomorphic robot's blink. Introducing such idle behaviors to the robot helps people perceive the idle robot as more animate [39]. For the active state, we made the robot randomly move within the moving area at a speed between the fast and slow reported in [32] (12 cm/s). To prevent the motion from eliciting any particular or extreme emotion, we chose the random movement, which was found to be relatively neutral [32].

3) Difficulty: Task difficulty is an important factor when studying the effect of presence on human task performance as demonstrated in the social facilitation effect. While the presence of another human could be beneficial for easy tasks, it is detrimental for hard tasks. Therefore, we included tasks with two difficulty levels. In particular, we replicated the task settings as in prior literature for numerical distance, modular arithmetic, and word recall tasks. The tasks and difficulty levels are described in detail further in this section.

In summary, participants were presented with the following six between-subjects conditions: four robotic presence conditions, namely (1) an idle anthropomorphic robot, (2) an active anthropomorphic robot, (3) an idle non-anthropomorphic robot, and (4) an active non-anthropomorphic robot, and two baseline conditions, namely (5) alone and (6) human presence.

C. Dependent Variables

1) Task Performance: To study the effects of different presence conditions on participants' task performance, it is necessary to measure how the participant performs on the given tasks under different conditions. For all three cognitive tasks, we recorded the completion time and accuracy for each trial as done in prior work [8].

2) Self-Reported Perception: We additionally gathered participants' quantitative and qualitative feedback on their experience. Specifically, we asked participants to self-report their emotion through SAM (Self-Assessment Manikin) [40] and sense of being judged on a 7-point Likert Scale. We also measured their perception of the robot on a 7-point Likert Scale in terms of anthropomorphism and animacy using the Godspeed questionnaire [41], and their desirability for the robot using the Adoption Likelihood Factors Questionnaire [42]. Qualitatively, we asked participants about the noticeability and the perceived influence of the robot.

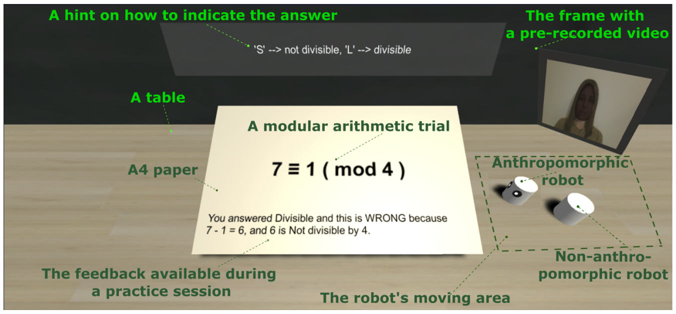

Fig. 2. The setup of the main 3D scene of the experiment consists of the task in the center, instruction on the top, and the different presence conditions on the right. The frame with a recorded video of a human, non-anthropomorphic and anthropomorphic robots are simultaneously shown here but during the actual trial, at most only one was shown.

D. Study Setup

The study was conducted in a 3D environment running in a browser. The software – developed in Unity 3D – comprises 3D scenes of the experiment, functionality for generating the tasks of various types and complexity and rendering the presence conditions, and functionality for logging the user's performance in the tasks.

The main 3D scene of the experiment, shown in Fig 2, resembles a typical context for knowledge work: a desk with an A4-sized paper. During the experiment, the study task appears on the paper. For the design of the objects in the scene, realistic neutral colors and mild lightning were chosen to avoid bright contrasts and sharp shadows. For example, we used wood material for the desk and dark gray color for the background. We kept the design of the things in the scene as simple as possible, but recognizable as everyday objects. On the right side of the paper, a robot would appear in a predefined area when applicable. For the human presence condition, a frame is shown for replaying a pre-recorded video of a human observer.

E. Tasks

Participants performed three cognitive tasks that were used in prior work: numerical distance [7], modular arithmetic [8], and word recall [43]. The task order across participants was counterbalanced using a balanced Latin square design.

1) Numerical Distance Task: The participants are shown a math inequality (e.g., '$|4-9| > 5$'), and asked to judge whether the absolute difference of the first two numbers is greater than or less than 5. The participants indicate their answer by a keypress. We constructed the difficulty level through the numerical distance effect that larger numerical distance between the two numbers will yield faster and more accurate answer [44]. Therefore, for an easy trial, we picked two numbers with the difference either 1 or 9 (e.g., '$10-1$'); for a complex trial, we picked two numbers with the difference either 4 or 6 (e.g., '$6-2$'). Easy and complex trials were presented in fully randomized order. Our setup is similar to that used by Riether et al. [7].

2) Modular Arithmetic Task: The participants are shown a math expression consisting of three numbers (e.g., '$10 \equiv 5 \pmod 2$'). Their job is to judge whether the difference of the first two numbers is divisible by a third number (i.e., the quotient is a whole number) through a key press. The difficulty of a trial was manipulated by controlling the number of digits in the first two numbers; one for an easy trial (e.g., '$7 \equiv 2$'), and two for a difficult trial ('$51 \equiv 19$'). Easy and complex trials were presented in fully randomized order. Our setup is the same as that from Park and Catrambone [8].

3) Word Recall Task: The participants are shown a set of words one by one. Their task is to remember the words and type them after a 10-second pause, similar to the setup by Berger et al. [43]. For the answer, only the accurate spelling is important. Relying on the concept of familiar and unfamiliar words [43], we selected 24 familiar words that were used for easy trials (e.g., "baker", "money", "pecan") and 24 unfamiliar words that were used for hard trials (e.g., "bosun", "xilos", "kalab"). During the task, each word would appear at least once, but at most two times.

F. Participants

We recruited 132 participants - approximately 20 participants for each condition - through Amazon Mechanical Turk for a mixed factorial design study. For quality control, only participants that satisfy the following requirements were included in the analysis: 1) located in the US, 2) the HIT approval rate is greater than 95, 3) the number of HITs approved greater than 1000, and 4) they have completed both the 3 tasks and the post-study questionnaire.

Requirements 1-3 were enforced through Amazon Mechanical Turk. For requirement 4, we removed 19 participants who did not complete all the tasks and/or the post-study questionnaire. To filter out the outliers, we removed trials within each task based on the completion time (i.e., $< Q3+1.5(Q3-Q1)$ and $> Q1-1.5(Q3-Q1)$). Then, for each task of every participant, we included their data only if they had a certain number of valid non-outlier trials. For numerical distance and modular arithmetic tasks, the threshold was 15 out of 20, while the threshold was 2 out of 2 for the word recall task.

For the analysis, 113 participants (72 men, 41 women) with a mean age of 37.7 years (SD = 11.2) were included. 17%, 63% and 20% of participants reported education levels of middle/high school, college and advanced degrees, respectively. On average, participants spent 7 minutes (SD = 4) on the actual task and 4 minutes (SD = 3.2) on the questionnaire for a total of 11 minutes; they were compensated at a rate of $15 per hour. None of the participants had neurological disorders, impaired vision, headache, fatigue, or any other conditions that may have affected their performance. The study was approved by the University's Institutional Review Board with subjects providing informed consent.

G. Procedure

The study was conducted using Amazon Mechanical Turk and Qualtrics. From MTurk - the entry point to the study - the participants proceeded to Qualtrics to access the study link with a unique parameter combination: a participant id and the condition type; the conditions were distributed as equally between the participants as possible by Qualtrics. After the study, the participants were directed back to Qualtrics to complete a post-study questionnaire.

The study link opens a welcoming screen with an introduction to the study. The participants are asked to a) enter the full screen mode, b) listen to the white noise plaid in the main study scene, c) avoid using external help, d) answer with maximum accuracy, and e) rest between the sessions and the tasks. Next, the participant receives the first task description. The task starts with a practice session (three trials for the math tasks and one trial with three words for the word recall task) with feedback. Next comes the actual session, where the participant performs trials of hard and easy difficulty (20 trials for the math tasks and 2 trials with 6 words each for the word recall task) under their condition. After each task, participants answer a questionnaire consisting of the self-assessment manikin (SAM) for measuring emotion (mild-intense, pleasant-unpleasant, dominated-dominant) on a 7-point scale [40] and a 7-point Likert scale on how much they felt like being judged/observed/evaluated. The same process repeats for the three tasks.

After finishing the experiment, the participant was redirected to Qualtrics for a post-study questionnaire to collect qualitative feedback and self-reported perception; the questions depended on the study condition. In particular, the participants were asked about the robots' a) noticeability, i.e., whether the user noticed the robots and how they perceived them, b) affective influence, i.e., how the presence of the robots affected the user's emotion, and c) desirability, i.e., whether the user could imagine using such robots daily.

H. Data Analysis

To examine the effects of the independent variables including interaction, a Levene's Test of Equality of Error Variances and a mixed-design ANOVA were performed for each task performance measure. To compare across robot states and anthropomorphism within the robot presence conditions, we used a 2 x 2 x 2 mixed-design ANOVA with two between-subject factors (i.e., robot states and anthropomorphism) and one within-subject factor (i.e., task difficulty). To compare across all conditions including alone and human presence conditions, we used a 2 x 6 mixed-design ANOVA with one between-subject factors (i.e., presence conditions) and one within-subject factor (i.e., task difficulty). If any independent variable or combinations had statistically significant effects (p ¡ 0.05), Bonferroni-corrected post-hoc tests were used to determine which pairs were significantly different.

Similarly, we ran both a 2 x 2 ANOVA with two between-subject factors (i.e., robot states and anthropomorphism) and a 6-way ANOVA with one between-subject factors (i.e., presence conditions) for the self-reported perception ratings. Then, we performed the Bonferroni-corrected post-hoc tests on the statistically significant effects.

IV. RESULTS

We summarize findings from our study in terms of task performance in Table I, and user perception (e.g., emotion, perceived anthropomorphism, perceived animacy, and adoption) from the self-reported questionnaire.

TABLE I
TASK PERFORMANCE (M: MEAN, SD: STANDARD DEVIATION) FOR NUMERICAL DISTANCE, MODULAR ARITHMETIC, AND WORD RECALL TASKS. THE BLUE AND BOLDED NUMBERS HIGHLIGHT THE BEST SCORE FOR EACH DIFFICULTY.

Numerical Distance		Time (s)				Accuracy (%)			
		Easy		Hard		Easy		Hard	
Condition	N	M	SD	M	SD	M	SD	M	SD
Alone	15	**1.47**	0.53	1.61	0.61	77	28	76	27
Human	14	1.50	0.45	1.73	0.55	**90**	20	**86**	22
Idle Non-Ant* Robot	16	1.87	0.42	1.97	0.32	84	24	83	22
Active Non-Ant* Robot	15	1.52	0.48	**1.46**	0.51	82	21	83	20
Idle Ant* Robot	10	1.62	0.46	1.7	0.43	81	26	70	29
Active Ant* Robot	16	1.66	0.67	1.76	0.7	74	33	77	33

Modular Arithmetics		Time (s)				Accuracy (%)			
		Easy		Hard		Easy		Hard	
Condition	N	M	SD	M	SD	M	SD	M	SD
Alone	15	**1.91**	0.93	**2.21**	1.19	67	19	63	21
Human	13	2.16	0.91	2.63	1.35	66	19	68	23
Idle Non-Ant* Robot	11	2.18	0.72	2.64	0.96	71	24	66	17
Active Non-Ant* Robot	14	1.97	0.84	2.67	1.31	**73**	23	63	27
Idle Ant* Robot	12	2.43	1.14	2.85	1.40	65	21	64	18
Active Ant* Robot	16	2.38	0.97	2.94	1.43	70	22	**71**	23

Word Recall		Time (s)				Correct Words (%)			
		Easy		Hard		Easy		Hard	
Condition	N	M	SD	M	SD	M	SD	M	SD
Alone	13	26.82	12.86	22.20	11.26	**78**	24	47	33
Human	13	27.46	11.54	24.20	11.26	68	28	46	26
Idle Non-Ant* Robot	17	24.50	8.35	**20.65**	8.10	68	15	51	24
Active Non-Ant* Robot	15	21.44	8.41	24.69	14.69	74	26	52	23
Idle Ant* Robot	16	29.91	13.14	30.56	13.69	77	20	57	24
Active Ant* Robot	15	**21.14**	10.69	24.90	12.82	**78**	21	**60**	27

*Anthropomorphic

A. Task Performance

1) Numerical Distance Task: In terms of the completion time, the *active non-anthropomorphic robot* condition had the fastest time for hard difficulty level while the *alone* condition yielded the fastest time for easy difficulty level. For accuracy, *human* condition led to the highest accuracy for both easy and hard difficulty levels.

Difficulty had statistically significant effects on completion time ($F(1,80) = 6.76, p = .011, \eta^2 = .078$), where the harder difficulty yielded longer completion time as expected.

There were statistically significant interaction effects between difficulty and robot state on accuracy ($F(1,53) = 4.47, p = .039, \eta^2 = .078$). As shown in Fig. 3A, there were statistically significant differences between the two difficulty levels for idle robots ($p = .036$) but not for active robots.

A 2x2x2 mixed-design ANOVA revealed close to statistically significant interaction effects between anthropomorphism and robot state on completion time ($F(1,53) = 3.36, p = .073, \eta^2 = .06$). Bonferroni-corrected post hoc tests show that there were statistically significant differences in that for completion time, the active non-anthropomorphic robot had lower completion time ($p = .017$) than the idle non-anthropomorphic robot as shown in Fig. 3B.

2) Modular Arithmetic Task: In terms of the completion time, the *alone* condition had the fastest time for both difficulties. For accuracy, the *active non-anthropomorphic robot* had the highest accuracy on easy difficulty while the *active anthropomorphic robot* condition had the highest accuracy on hard difficulty.

Difficulty had statistically significant effects on completion time ($F(1,75) = 24.0, p < .001, \eta^2 = .242$), where the harder difficulty level yielded longer completion time as expected. No other statistically significant effects were found.

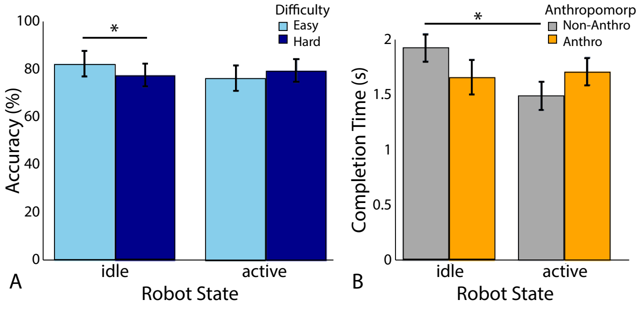

Fig. 3. Interaction effects for the numerical distance task: A) between the robot state and difficulty on accuracy and B) between the robot state and anthropomorphism on completion time.

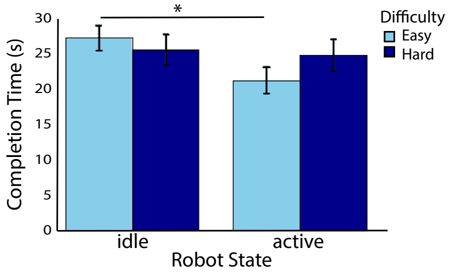

Fig. 4. Interaction effects between the robot state and difficulty for the word recall task on completion time.

3) Word Recall Task: Table I shows the means and standard deviations of completion time and percentage of correctly recalled words within 1 letter for easy and hard difficulties.In terms of the completion time, the *active anthropomorphic robot* condition had the fastest time for easy difficulty level while the *idle non-anthropomorphic robot* condition had the fastest time under hard difficulty. For accuracy, the *active non-anthropomorphic robot* and *alone* conditions had the highest accuracies on easy difficulty while the *active anthropomorphic robot* condition had the highest accuracy on hard difficulty.

Difficulty had statistically significant effects on accuracy (i.e., # of correct words within 1 letter) ($F(1, 83) = 95.8, p < .001, \eta^2 = .536$), where the harder difficulty level yielded lower accuracy as expected.

A 2x2x2 mixed-design ANOVA revealed close to statistically significant interaction effects between difficulty and robot state on completion time ($F(1, 59) = 3.37, p = .071, \eta^2 = .054$). Bonferroni-corrected post hoc tests demonstrate that there were statistically significant differences for completion time. For the easy difficulty, the active robots led to lower completion time ($p = .027$) than the idle robots, but not for the hard difficulty as shown in Fig. 4.

B. Self-Reported Perception

1) Emotion and Sense of Being Evaluated: A 2x2 mixed-design ANOVA revealed that anthropomorphism had statistically significant effects on dominance axis of emotion for the modular arithmetic task ($F(1, 39) = 5.84, p = .02, \eta^2 = .13$) and the word recall task ($F(1, 39) = 6.72, p = .013, \eta^2 = .147$). For both tasks, the participants reported feeling more dominant with the presence of an anthropomorphic robot than with the presence of a non-anthropomorphic robot as shown in Fig. 5A.

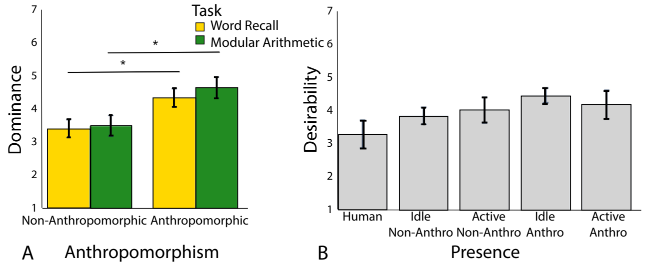

Fig. 5. (A) Effects of anthropomorphism on participants' self-reported dominance ratings and (B) Participant's self-reported desirability for different conditions.

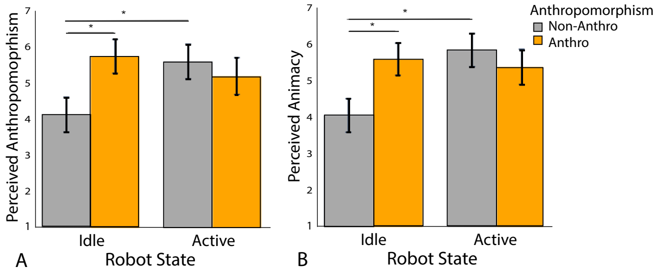

Fig. 6. Interaction effects between the robot state and anthropomorphism on the participants' perceived anthropomorphism and animacy ratings.

2) Perception of the Robot and Desirability: For both the perceived anthropomorphism and the perceived animacy of the robots, there were statistically significant interaction effects between the robot state and anthropomorphism ($F(1, 71) = 4.31, p = .041, \eta^2 = .057$, and $F(1, 71) = 4.74, p = .033, \eta^2 = .063$ respectively). Specifically, an idle non-anthropomorphic robot had lower perceived anthropomorphism and animacy ratings than an idle anthropomorphic robot ($p = .019$ and $p = .018$ respectively) and an active non-anthropomorphic robot ($p = .034$ and $p = .007$ respectively) as shown in Fig. 6.

For the participant's desirability for the robot or human presence, there were no statistically significant differences across the five presence conditions (excluding alone condition), which are shown in Fig. 5B, and no statistically significant effects of the robot state or anthropomorphism.

3) Noticeability and Perceived Influence of the Robot: Eighteen people out of 93 (93 out of 113 were those participants whose condition involved a robot) assumed that the robot was a distraction introduced to the task scene on purpose. However, the majority of the respondents did not notice the presence of the robot, glanced at it but did not pay much attention to it, or tried to avoid looking at it while focusing on the task. Twenty two people reported feeling like being evaluated when the robot was presented, while seventy one reported that the presence of the robot did not made them feel like they were being evaluated.

V. DISCUSSION & FUTURE WORK

Anthropomorphism had a significant effect on the dominance axis of affect for 2 out of 3 tasks. Interestingly, participants felt more dominant with the presence of an anthropomorphic robot compared to the presence of a non-anthropomorphic robot rejecting H1b. Because high dominance is associated with the subjects having maximum control in the situation [40], the collected data suggests that the users felt more in control when an anthropomorphic robot was presented in the scene. This observation might be related to the fact that anthropomorphizing automated entities can provide a sense of control over them and make the environment appear more predictable [45]. In terms of task performance, anthropomorphism did not lead to statistically significant effects either alone or with difficulty thus rejecting H1a. This suggests that anthropomorphism alone may not elicit evaluation apprehension at least with the two anthropomorphisms used for this study, and it may more complicated than adding anthropomorphic features to elicit evaluation apprehension among participants. Further investigation is needed to uncover what aspects lead to evaluation apprehension to better understand how to design a robot for social facilitation and other applications.

The robot state had statistically significant main and interaction effects with difficulty and anthropomorphism on task performance for both numerical distance and word recall tasks. For the numerical distance task, an idle robot led to higher accuracy for easy difficulty compared to hard difficulty. On the other hand, for both numerical distance and word recall tasks, an active robot led to lower completion time than an idle robot for easy difficulty levels. This partially confirms our hypothesis H2a as we observed lower completion time with active robots compared to idle robots albeit with sometimes lower accuracy, which may be due to the distraction from the robot motion.

In terms of the perceived anthropomorphism and animacy, the robot state turned out to be as influential as the anthropomorphism. For the non-anthropomorphic robot, changing its robot state from idle to active made its perceived anthropomorphism and animacy ratings as high as those of the anthropomorphic robots either idle or active. While quite unexpected, this finding is aligned with the findings of prior literature on self-produced motion and how it is a powerful cue for animacy and agency [19]. We also observed similar effects on completion time for the numerical distance task. Specifically, there were statistically significant differences between the idle and active non-anthropomorphic robots, while there were not any statistically significant differences between the idle and active anthropomorphic robots.

In terms of the human presence, we expected to observe the social facilitation effect, which entails social facilitation in simple tasks and social inhibition in complex tasks. However, the study results are inconsistent across the three tasks and do not demonstrate the whole effect. This is consistent with the findings from a review in social facilitation in virtual environments with virtual observers [46], where 12 out of 13 analyzed studies failed to show the whole effect. As the social realism has been identified as being an important factor for producing social facilitation [47], the implementation of our human presence may need to be more socially realistic to observe the social facilitation effect.

Despite the statistically significant effects of the robot presences on the task performance and self-reported emotion of the participants, most participants reported either not noticing the robot, not paying much attention, or trying to not look at it to focus on the tasks. This suggests that the presence of the robot had unconsciously affected the participants and both in a positive or negative manner depending on the task, task difficulty, robot state, and anthropomorphism.

One limitation of this study is that the experiment was conducted in an online virtual environment. While it serves as a useful design probe to understand user reaction and perception [48], [49], an online virtual study does not always capture all the nuances of an in-person study and limits the presence effects. We plan to conduct the same experiment in-person to verify the findings in the future.

According to the ABOT database [35], both robots used in this study rank low (< 5) in terms of their human-likeness score. On the other hand, anthropomorphism of a robot is a full continuum ranging from a non-anthropomorphic [33] to a human-like android robot [50]. As such, a wide range of anthropomorphism should be evaluated to see its effect on user's task performance and emotion. With a human-like Android, we may see an extreme level of social facilitation similar to trend observed in the Uncanny Valley [10].

VI. CONCLUSION

With robots becoming more ubiquitous, it is necessary to understand the effects of robotic presence. Given the prevalence of non-humanoid mobile robots, we investigated how the anthropomorphism and state of the robot affects user's task performance and emotion. Our virtual study findings suggest that the robot state plays a significant role as much if not more than the anthropomorphism in that an active robot can lead to shorter completion time albeit with some reduction in accuracy in easy difficulty levels. The robot state changed how people perceived the robot in terms of anthropomorphism and animacy in that an active non-anthropomorphic robot was perceived similarly to an anthropomorphic robot. Future investigation should study the complex interaction between the task, task difficulty, robot state, and anthropomorphism in an in-the-wild setting.

ACKNOWLEDGMENT

We thank all the participants for their time and effort.

REFERENCES

[1] L. Takayama, W. Ju, and C. Nass, "Beyond dirty, dangerous and dull: what everyday people think robots should do," in *2008 3rd ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2008, pp. 25–32.

[2] A. Ramachandran, C.-M. Huang, E. Gartland, and B. Scassellati, "Thinking aloud with a tutoring robot to enhance learning," in *Proceedings of the 2018 ACM/IEEE international conference on human-robot interaction*, 2018, pp. 59–68.

[3] B. Scassellati, L. Boccanfuso, C.-M. Huang, M. Mademtzi, M. Qin, N. Salomons, P. Ventola, and F. Shic, "Improving social skills in children with asd using a long-term, in-home social robot," *Science Robotics*, vol. 3, no. 21, 2018.

[4] W.-L. Chang, S. Šabanovic, and L. Huber, "Use of seal-like robot paro in sensory group therapy for older adults with dementia," in *2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2013, pp. 101–102.

[5] M. J. Matarić, J. Eriksson, D. J. Feil-Seifer, and C. J. Winstein, “Socially assistive robotics for post-stroke rehabilitation,” *Journal of NeuroEngineering and Rehabilitation*, vol. 4, no. 1, pp. 1–9, 2007.

[6] I. Wechsung, P. Ehrenbrink, R. Schleicher, and S. Möller, “Investigating the social facilitation effect in human–robot interaction,” in *Natural Interaction with Robots, Knowbots and Smartphones*. Springer, 2014, pp. 167–177.

[7] N. Riether, F. Hegel, B. Wrede, and G. Horstmann, “Social facilitation with social robots?” in *2012 7th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2012, pp. 41–47.

[8] S. Park and R. Catrambone, “Social facilitation effects of virtual humans,” *Human factors*, vol. 49, no. 6, pp. 1054–1060, 2007.

[9] J. Złotowski, D. Proudfoot, K. Yogeeswaran, and C. Bartneck, “Anthropomorphism: opportunities and challenges in human–robot interaction,” *International journal of social robotics*, vol. 7, no. 3, pp. 347–360, 2015.

[10] M. Mori, K. F. MacDorman, and N. Kageki, “The uncanny valley [from the field],” *IEEE Robotics & Automation Magazine*, vol. 19, no. 2, pp. 98–100, 2012.

[11] F. Heider and M. Simmel, “An experimental study of apparent behavior,” *The American journal of psychology*, vol. 57, no. 2, pp. 243–259, 1944.

[12] S. W. Chae, K. C. Lee, and Y. W. Seo, “Exploring the effect of avatar trust on learners’ perceived participation intentions in an e-learning environment,” *International Journal of Human–Computer Interaction*, vol. 32, no. 5, pp. 373–393, 2016.

[13] C. Anderson-Hanley, A. L. Snyder, J. P. Nimon, and P. J. Arciero, “Social facilitation in virtual reality-enhanced exercise: competitiveness moderates exercise effort of older adults,” *Clinical interventions in aging*, vol. 6, p. 275, 2011.

[14] B. Yoon, H.-i. Kim, G. A. Lee, M. Billinghurst, and W. Woo, “The effect of avatar appearance on social presence in an augmented reality remote collaboration,” in *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 2019, pp. 547–556.

[15] K. Emmerich and M. Masuch, “Watch me play: does social facilitation apply to digital games?” in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 2018, pp. 1–12.

[16] I. Howley, T. Kanda, K. Hayashi, and C. Rosé, “Effects of social presence and social role on help-seeking and learning,” in *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*, 2014, pp. 415–422.

[17] A. Baylor, “Beyond butlers: Intelligent agents as mentors,” *Journal of educational computing research*, vol. 22, no. 4, pp. 373–382, 2000.

[18] P. H. Kahn, T. Kanda, H. Ishiguro, B. T. Gill, S. Shen, J. H. Ruckert, and H. E. Gary, “Human creativity can be facilitated through interacting with a social robot,” in *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2016, pp. 173–180.

[19] E. Mascalzoni, L. Regolin, and G. Vallortigara, “Innate sensitivity for self-propelled causal agency in newly hatched chicks,” *Proceedings of the National Academy of Sciences*, vol. 107, no. 9, pp. 4483–4485, 2010.

[20] J.-H. Lee, J.-Y. Park, and T.-J. Nam, “Emotional interaction through physical movement,” in *International Conference on Human-Computer Interaction*. Springer, 2007, pp. 401–410.

[21] J. Posner, J. A. Russell, and B. S. Peterson, “The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology,” *Development and psychopathology*, vol. 17, no. 3, pp. 715–734, 2005.

[22] M. Saerbeck and C. Bartneck, “Perception of affect elicited by robot motion,” in *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2010, pp. 53–60.

[23] M. Pavlova, A. A. Sokolov, and A. Sokolov, “Perceived dynamics of static images enables emotional attribution,” *Perception*, vol. 34, no. 9, pp. 1107–1116, 2005.

[24] B. Guerin, “Social facilitation,” *The Corsini encyclopedia of psychology*, pp. 1–2, 2010.

[25] K. W. Spence, “Behavior theory and conditioning.” 1956.

[26] R. B. Zajonc, “Social facilitation,” *Science*, vol. 149, no. 3681, pp. 269–274, 1965.

[27] N. B. Cottrell, “Social facilitation,” *Experimental social psychology*, vol. 185, p. 236, 1972.

[28] T. Kushnir and K. Duncan, “An analysis of social facilitation effects in terms of signal detection theory,” *The Psychological Record*, vol. 28, no. 4, pp. 535–541, 1978.

[29] E. Jones and H. Gerard, “Foundations of social psychology. john whiley & sons,” 1967.

[30] R. S. Baron, D. Moore, and G. S. Sanders, “Distraction as a source of drive in social facilitation research.” *Journal of personality and social psychology*, vol. 36, no. 8, p. 816, 1978.

[31] J. A. Easterbrook, “The effect of emotion on cue utilization and the organization of behavior.” *Psychological review*, vol. 66, no. 3, p. 183, 1959.

[32] L. H. Kim and S. Follmer, “Ubiswarm: Ubiquitous robotic interfaces and investigation of abstract motion as a display,” *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 1, no. 3, pp. 1–20, 2017.

[33] M. Le Goc, L. H. Kim, A. Parsaei, J.-D. Fekete, P. Dragicevic, and S. Follmer, “Zooids: Building blocks for swarm user interfaces,” in *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, 2016, pp. 97–109.

[34] K.-A. Kwon, R. J. Shipley, M. Edirisinghe, D. G. Ezra, G. Rose, S. M. Best, and R. E. Cameron, “High-speed camera characterization of voluntary eye blinking kinematics,” *Journal of the Royal Society Interface*, vol. 10, no. 85, p. 20130227, 2013.

[35] E. Phillips, X. Zhao, D. Ullman, and B. F. Malle, “What is human-like?: Decomposing robots’ human-like appearance using the anthropomorphic robot (abot) database,” in *2018 13th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2018, pp. 105–113.

[36] J. Fink, “Anthropomorphism and human likeness in the design of robots and human-robot interaction,” in *International Conference on Social Robotics*. Springer, 2012, pp. 199–208.

[37] L. Onnasch, E. Kostadinova, and P. Schweidler, “Humans can’t resist robot eyes–reflexive cueing with pseudo-social stimuli,” *Frontiers in Robotics and AI*, p. 72.

[38] C. F. DiSalvo, F. Gemperle, J. Forlizzi, and S. Kiesler, “All robots are not created equal: the design and perception of humanoid robot heads,” in *Proceedings of the 4th conference on Designing interactive systems: processes, practices, methods, and techniques*, 2002, pp. 321–326.

[39] T. Asselborn, W. Johal, and P. Dillenbourg, “Keep on moving! exploring anthropomorphic effects of motion during idle moments,” in *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 2017, pp. 897–902.

[40] M. M. Bradley and P. J. Lang, “Measuring emotion: the self-assessment manikin and the semantic differential,” *Journal of behavior therapy and experimental psychiatry*, vol. 25, no. 1, pp. 49–59, 1994.

[41] C. Bartneck, D. Kulić, E. Croft, and S. Zoghbi, “Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots,” *International journal of social robotics*, vol. 1, no. 1, pp. 71–81, 2009.

[42] TryMyUI, “The alfq and usability testing,” 2021. [Online]. Available: https://trymyui.com/alfq

[43] S. M. Berger, K. L. Hampton, L. L. Carli, P. S. Grandmaison, J. S. Sadow, C. H. Donath, and L. R. Herschlag, “Audience-induced inhibition of overt practice during learning.” *Journal of Personality and Social Psychology*, vol. 40, no. 3, p. 479, 1981.

[44] R. S. Moyer and T. K. Landauer, “Time required for judgements of numerical inequality,” *Nature*, vol. 215, no. 5109, pp. 1519–1520, 1967.

[45] A. Waytz, C. K. Morewedge, N. Epley, G. Monteleone, J.-H. Gao, and J. T. Cacioppo, “Making sense by making sentient: effectance motivation increases anthropomorphism.” *Journal of personality and social psychology*, vol. 99, no. 3, p. 410, 2010.

[46] R. Sterna, P. Strojny, and K. Rebilas, “Can virtual observers affect our behavior?” *Social Psychological Bulletin*, vol. 14, no. 3, pp. 1–18, 2019.

[47] P. M. Strojny, N. Dużmańska-Misiarczyk, N. Lipp, and A. Strojny, “Moderators of social facilitation effect in virtual reality: Co-presence and realism of virtual agents,” *Frontiers in psychology*, vol. 11, p. 1252, 2020.

[48] E. W. Pedersen, S. Subramanian, and K. Hornbæk, “Is my phone alive? a large-scale study of shape change in handheld devices using videos,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2014, pp. 2579–2588.

[49] L. Takayama, D. Dooley, and W. Ju, “Expressing thought: improving robot readability with animation principles,” in *Proceedings of the 6th international conference on Human-robot interaction*, 2011, pp. 69–76.

[50] D. Sakamoto, T. Kanda, T. Ono, H. Ishiguro, and N. Hagita, “Android as a telecommunication medium with a human-like presence,” in *2007 2nd ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2007, pp. 193–200.