Dog Sit! Domestic Dogs (*Canis familiaris*) Follow a Robot's Sit Commands

Meiying Qin
meiying.qin@yale.edu
Yale University

Yiyun Huang*
Ellen Stumph*
yiyun.huang@yale.edu
ellen.stumph@yale.edu
Yale University

Laurie Santos
laurie.santos@yale.edu
Yale University

Brian Scassellati
brian.scassellati@yale.edu
Yale University

ABSTRACT

As personal social robots become more prevalent, the need for the designs of these systems to explicitly consider pets become more apparent. However, it is not known whether dogs would interact with a social robot. In two experiments, we investigate whether dogs respond to a social robot after the robot called their names, and whether dogs follow the 'sit' commands given by the robot. We conducted a between-subjects study ($n = 34$) to compare dogs' reactions to a social robot with a loudspeaker. Results indicate that dogs gazed at the robot more often after the robot called their names than after the loudspeaker called their names. Dogs followed the 'sit' commands more often given by the robot than given by the loudspeaker. The contribution of this study is that it is the first study to provide preliminary evidence that 1) dogs showed positive behaviors to social robots and that 2) social robots could influence dog's behaviors. This study enhance the understanding of the nature of the social interactions between humans and social robots from the evolutionary approach. Possible explanations for the observed behavior might point toward dogs perceiving robots as agents, the embodiment of the robot creating pressure for socialized responses, or the multimodal (i.e., verbal and visual) cues provided by the robot being more attractive than our control condition.

KEYWORDS

Animal–Robot Interactions; Canine; Social Robot; Evolutionary Approach

ACM Reference Format:
Meiying Qin, Yiyun Huang, Ellen Stumph, Laurie Santos, and Brian Scassellati. 2020. Dog Sit! Domestic Dogs (*Canis familiaris*) Follow a Robot's Sit Commands. In *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction (HRI '20 Companion), March 23–26, 2020, Cambridge, United Kingdom.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3371382.3380734

*Both authors contributed equally to this research.

HRI '20 Companion, March 23–26, 2020, Cambridge, United Kingdom

ACM ISBN 978-1-4503-7057-8/20/03…$15.00
https://doi.org/10.1145/3371382.3380734

1 INTRODUCTION

Figure 1: A dog encounters with a robot.

As robots become popular, a pet may come in contact with robots regularly in everyday lives. One only needs to look to the near limitless archive of internet videos featuring pets and robotic vacuum cleaners for evidence of the ubiquity of these–often times entertaining–interactions. As personal social robots become more prevalent, the need for the designs of these systems to explicitly consider pets become more apparent. Unlike robot vacuum cleaners, social robots may be designed with versatile behaviors rather than simple movements and thus may attract the animals' attention. However, previous studies showed that the movements of mechanic toys could trigger stress and fear in animals [5, 13, 21]. Ideally, the social behaviors of the robots should not be seen as threatening and should not consistently trigger the 'fight-or-flight' responses. Beyond the psychological well-being of pets, it is not known whether pets would interact socially with a social robot. If pets indeed interact with robots socially, then the designs of personal social robots should consider the potential impacts of the robot behaviors may have on pets. Moreover, personal social robots could even be designed with behaviors to entertain pets, given that the pets would interact socially with the robot.

Furthermore, the study of the interactions between animals and social robots could provide insight to Human–Robot Interactions (HRI) studies. Psychologists use both developmental and evolutionary approaches to study human cognition. The developmental approach focuses on the changes at the individual level. Studies with this approach compare adults against children to understand how certain behaviors or cognitive abilities develop ontogenetically. Studies in human-robot interactions already have shown that children could interact with a social robot at an early age [26]. By comparison, the evolutionary approach studies the changes at the species level. Studies with this approach compare humans against other species to understand how certain traits develop phylogenetically. No previous studies in HRI so far took this approach. The result of this study could be one of the first studies to enhance the understanding of the nature of the social interactions between humans and social robots from the evolutionary approach.

In this study, we explored how animals would react to personal social robots by using the model animal dogs (*Canis familiaris*) (Figure 1). Dogs were chosen for two reasons. Firstly, dogs are one of the most common types of family pets, as well as one of the most common species that works with humans. Secondly, dogs excel in understanding human social behaviors when compared to other species [9]. We investigated whether dogs would respond to a social robot when calling their names and whether dogs would follow commands given by the social robot.

2 RELATED WORK

We first consider studies testing the initial responses of dogs when interacting with a social robot, as well as a series of studies measuring the effectiveness of different animated objects used as stimulus in dog social cognition research. We conclude with a review of past work exploring how dogs react to virtual agents with no embodiment.

2.1 Dogs Interacting with Social Robots

Morovitz et. al [19] studied how dogs responded to an unfamiliar social robot. They compared dogs' reactions to a humanoid robot Nao behaving in a human-like manner–walking, waving, vocalization, and head turning–and a Nao behaving mechanically. The authors minimized guardians' interactions with the dogs and with the robot during the 3-minutes interactions. Their results showed that 57.1% of the dogs took the food rewards from the Nao behaving mechanically, and none of the dogs took the food rewards from the Nao behaving socially. Dogs also showed more positive behaviors when interacting with the Nao behaving mechanically than with the Nao behaving socially.

This result might suggest dogs prefer non-social robots over social robots, at least in brief encounters. However, it is not clear whether this preference would persist given longer interactions between the dogs and the robot. Other studies [16] have found that dogs show positive responses when they were allowed sufficient time to interact with a robot that behaved socially.

2.2 Effectiveness of Animated Objects in Dog Cognition Studies

Researchers studying dog cognition have employed a wide variety of animated objects in order to make stimuli controllable, repeatable, and bias free [1, 3, 4, 16]. Many researchers have simply assumed that animated objects can trigger social responses in dogs, and that failure to trigger social responses was attributed to the limitations of social behaviors exhibited by the animated objects. Nonetheless, the results from dog cognition studies using animated stimuli can guide the design of dog–robot interactions studies. In this section, we review studies using animated objects engaging in three different kinds of behaviors: simple social behaviors, human-like behaviors, and dog-like behaviors.

2.2.1 Animated Objects with Simple Social Behaviors. Animated objects with simple behaviors are generally referred to as unidentified moving objects (UMOs). UMOs are self-propelled objects with no obvious control by people. Gergely et. al [4] showed that dogs engaged in longer durations of gazing at and touching a UMO which took various different routes toward a goal than a UMO that repeated the same route towards the goal. Gergely et. al [3] showed in another study that dogs followed cues for hidden food rewards provided by a UMO which had provided food rewards in previous testing sessions, but ignored the cues provided by a UMO which did not interact with them before. Abdai et. al [1] showed that UMOs which had helped to retrieve food in previous testing sessions could even alter dogs' choices such that dogs who preferred larger quantities of food chose food rewards with smaller quantities.

In this way, dogs responded positively to the simple social behaviors of UMOs, and could understand the cues correctly.

2.2.2 Animated Objects with Human-like Behaviors. Lakatos et. al [16] conducted a study to test how dogs responded to the pointing cues given by a PeopleBot with customized arms. The PeopleBot may either exhibited human-like behaviors or no social behaviors, depending on the condition. The social behaviors included shaking hands with guardians, communicating with guardians verbally, and walking alongside with guardians in the room. A dog participant observed the robot interacting with the guardian either socially or mechanically for six minutes in the interaction phase. The robot then delivered a food reward for the dog. In the subsequent testing phase, the robot pointed to one of the two buckets with hidden food rewards.

In the interaction phase, dogs gazed longer at the head of the robot and spent more time close to the robot with human-like social behaviors than to the robot behaving mechanically. In the testing phase, dogs performed better in the condition with a social robot than with a nonsocial robot. However, no evidence suggested the mean performance with the social robot was significantly different from 50%, which is the chance level in two-choice tasks. Therefore, the dogs did not consistently follow the pointing cues provided by the social robot, even though dogs in general follow human pointing cues well [10, 18]. However, dogs gazed longer at the locations pointed by the robot with human-like behaviors.

Therefore, this study showed evidence that dogs behaved more positive to a robot with human-like social behaviors, but did not suggest dogs could follow the cues provided by the robot to make the correct choices.

2.2.3 Animated Objects with Dog-like Behaviors. AIBO is a robot dog that could display dog-like behaviors. Kubinyi et. al [14, 15] studied whether AIBO could trigger social responses in dogs that are similar to responses toward their conspecifics. The dogs encountered the robot in two situations: a simple encounter with no specific task, and a competitive situation in which the robot appeared to compete with the dogs for food. Results showed that dogs displayed social responses as if they were interacting with conspecifics, but only for a few seconds. Moreover, in the competitive situation, dogs did not perceive AIBO as a competitor for food.

A key difference between this study and previous studies lies in that AIBO was disguised to mimic a dog, but the UMOs were not designed to mimic any animals and the PeopleBot was not designed to fool the dog as if it were a human. As a result, unlike the stimuli in other studies which could trigger prolonged social responses in dogs, AIBO could not trigger social responses for more than a few seconds.

2.3 Dogs Interacted with Virtual Agents

In this section, we will cover studies on how dogs responded to virtual agents. Péter et. al [20] conducted a study in which dogs were presented with life-size videos of human experimenters on a screen locating food hidden in the room. The videos were prerecorded and thus were non-interactive. Dogs were able to find the food when it was hidden in the same room where the videos were shown. However, they could not find the food if the locations of the hidden food were located in a room other than where the videos were filmed.

Pongrácz et. al [22] tested whether dogs followed commands from their guardians with various levels of embodiment. The guardians may be present in the same room as the dogs (i.e., 3D condition), or interacted with the dogs via live-stream life-size interactive videos (i.e., 2D condition), or interacted with the dogs with only their voices came out of a loudspeaker (i.e., 0D condition). Dogs followed the commands most reliably in the 3D condition. They followed the commands least consistently in the 0D condition, and their performances were between 3D and 0D condition in the 2D condition.

Resner [23] tested how professional dog trainers could train dogs who were previously clicker-trained to follow the commands from a loudspeaker which is part of the Rover@Home system. Out of four dogs tested, only one dog showed sufficient evidence of following the commands from the loudspeaker during the first training session.

3 METHODOLOGY

In the current study, we tested how dogs responded to a social robot, Nao. Dogs were evaluated with two experiments: 1) whether dogs attended to the robot when the robot called their names after brief interactions, and 2) whether dogs followed the robot's 'sit' command after extensive interactions.

We compared the dog's behaviors towards a social robot (i.e., the testing condition), with a loudspeaker (i.e., the control condition). Figure 2 summarizes the procedure of the experiment. In the control condition, the robot was simply replaced with a loudspeaker. The detailed differences between the testing condition and the control condition were described in the following sections. The experiment started with a very brief introduction between robot/loudspeaker, the dog and its guardian. In Experiment 1, the robot/loudspeaker called the dog's name. Then the robot/loudspeaker verbally communicated with the guardian by asking a series of predetermined questions. The robot/loudspeaker then provided food rewards to the dog. In Experiment 2, the robot/loudspeaker gave the 'sit' commands to the dog.

We hypothesized that dogs would 1) gaze at the robot more often than at the loudspeaker and 2) follow the robot's commands more often than the loudspeaker's commands.

3.1 Participants

42 dogs with various ages, genders and breeds participated in the study. The dogs in the testing condition had never participated in studies with robots. 8 dogs were excluded due to: experimenter errors (2), technical problems (2), showing no interest in food rewards (2), guardian errors (1), or not following the 'sit' commands from the guardian (1). As a result, 17 dogs participated in the testing condition, and 17 dogs participated in the control condition. Exclusion rates in studies with dogs [11, 12] are generally higher compared to HRI studies with human participants. The exclusion rate in the current study falls within the normal range for dog cognition studies.

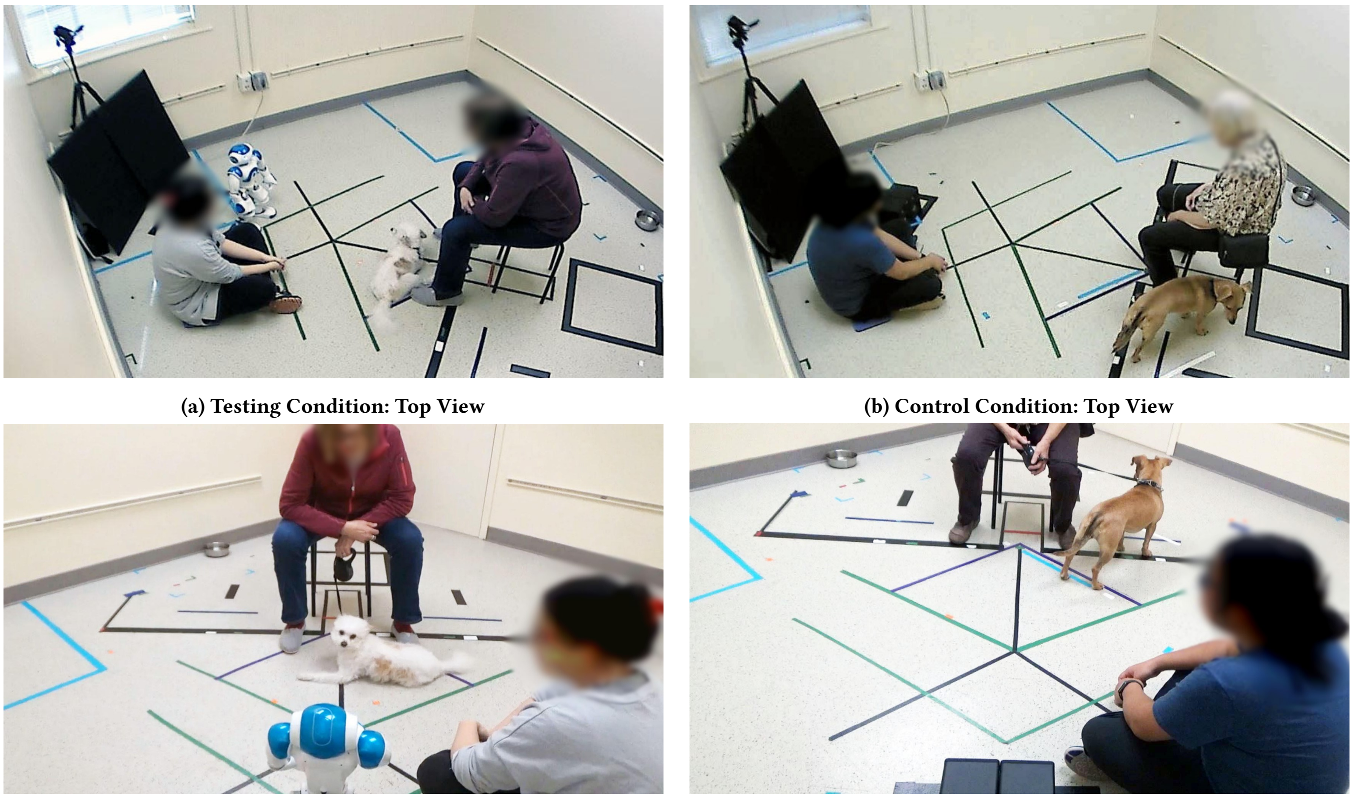
3.2 Apparatus

The robot used in this study is a commercial robot, Nao and was controlled via the Wizards-of-Oz manner (WoZ) [24]. We chose Nao because Nao is a typical choice in HRI studies, and it is known that humans socially responded to Nao [7, 8, 25, 27, 29].

All of the utterances from the robot came out of the speakers located in the head of the robot. Pre-recorded human audio clips, rather than the built-in text-to-speech (tts) module, were employed due to two reasons. First, experiment 1 tested whether a dog would respond to a social robot calling its name. Pronouncing the names precisely was crucial, but TTS could not guarantee correct pronunciations of the names, especially the names that were uncommon, made-up by the guardians, or not in the same language as the experiment conducted. Second, we focused on inducing responses from dogs over the ease of replication since few studies have demonstrated dog-robot interactions. As dogs are more sensitive to sounds than humans, we were concerned that dogs could not interpret TTS speech correctly. Even TTS that sounds natural to humans may lack the acoustic features that dogs rely upon. Moreover, dogs respond better to the exaggerated tones in infant-directed speech, which are used extensively in dog studies. Most TTS systems are limited in the tones they can provide.

We recorded two versions of the dogs' names: one with a normal pitch, and one with a raised pitch. The raised pitch helped to attract dogs' attention when giving commands. Each sentence was generated with multiple audio files containing segments of the sentence. The appropriate audio clips containing dog names were chosen based on context. The robot was also programmed with appropriate body language to accompany the utterances to make

Figure 2: General Procedure. The experiment started with a brief introduction. In Experiment 1, the robot/loudspeaker called the dog's name. Then the robot/loudspeaker verbally communicated with the guardian, and then provided food rewards to the dog. In Experiment 2, the robot/loudspeaker gave the 'sit' commands to the dog.



(a) Testing Condition: Top View

(b) Control Condition: Top View

(c) Testing Condition: Side View

(d) Control Condition: Side View

Figure 3: The Setup in the Testing room. The camera was mounted on a tripod positioned at one of the corners. The robot or loudspeaker stood in front of the occluders, with a human experimenter (E1) sitting to the right of it. The guardians sat on a stool in front of the robot. a) the setup in the testing condition from the top view; b) the setup in the control condition from the top view; c) the setup in the testing condition from the view behind the robot. d) the setup in the control condition from the view behind the loudspeaker.

the robot act more naturally. The loudspeaker used in the control condition was Alesis M1Active 330 USB. The utterances from the loudspeaker were produced by an experimenter during the study.

The food rewards were cut from natural balance beef rolls or food brought in by the guardians if the dogs had food restrictions or were not interested in the food rewards provided.

Figure 3 shows the layout of the testing room from different angles in the testing and the control condition. The testing room was 3.10 meters by 3.51 meters. The camera was mounted on a tripod positioned at one of the corners. The tripod's legs and electrical wires were covered behind black occluders. The robot or the loudspeaker stood in front of the occluders, with a human experimenter (E1) sitting to the right of it. The guardians sat on a stool approximately 1.30 meters away from the robot. The dogs were kept on a retractable leash and were allowed to move anywhere in the room other than behind the occluders. A water bowl was placed on the side so that dogs could access water at any time during the experiment.

3.3 Procedure

Before the guardian entered the testing room, they were encouraged to interact socially with the robot by engaging verbal communicates with the robot and by making eye contact with the robot. We also allowed the guardian to encourage the dog to interact with the robot, which differed from Morovitz et.al's study [19]. We believed that guardian's attitude towards the robot and their encouragement to their dogs to interact with the robot might affect how the dog perceived the robot and thus influenced the dog's behaviors. Since our goal was to test the dog's reactions to a personal social robot, allowing the guardian to interact with the robot and to encourage the dog to interact with the robot would share closer resemblance to the encounters in a natural setting.

3.3.1 Phase 1: Brief Introduction. In this phase, the guardian and the dog were led to the room by experimenter E2. The robot/loudspeaker and experimenter E1 were already situated in their designated testing locations. The robot and the guardian greeted each other in the testing condition. The loudspeaker was introduced to the guardian and the dog by E2 in the control condition.

3.3.2 Phase 2: Experiment 1–Robot/Loudspeaker Called the Dog's Name. After a brief encounter between the robot/loudspeaker and the dogs, Experiment 1 was administered. Experiment 1 consisted of only one trial to test whether the dog responded to the robot/loudspeaker calling its name. To eliminate novelty effect of sudden sound, the dog's name was embedded in the sentence 'I am so excited to play with <dog name> today.'. This is the first time in the experiment when the dog's name was pronounced with a raised pitch.

3.3.3 Phase 3: Conversation with the Guardian. In this phase, the robot/loudspeaker carried on conversation with the guardian for approximately five minutes. Previous studies showed that novel objects with sudden movements such as mechanic toys could trigger stress and fear in dogs [5, 13, 21]. Therefore, we restricted the robot to stand at the designated location where the dog first encountered it, and did not let the robot perform any sudden movements such as walking. The robot/loudspeaker asked the guardian a series of scripted closed-ended (i.e., yes-no) questions, and the robot/loudspeaker responded with scripted responses corresponding to the guardians' answers. The scripted responses included jokes to relax the guardian and make the conversions feel more natural. During this phase, the dog was allowed to explore anywhere in the room while witnessing the conversion. The purpose of this phase was to 1) provide enough time to familiarize the dog with the robot/loudspeaker to reduce the anxiety and stress in dogs induced by the novelty effect, and to 2) relieve any potential stress with the novel object (i.e., the robot or the loudspeaker) by witnessing the friendliness between the robot/loudspeaker and their guardians.

3.3.4 Phase 4: Providing Food Rewards to Dogs. In phase 4, the robot/loudspeaker interacted with the dog by providing food rewards. If the dog consumed the food rewards in two consecutive trials within a maximum of six trials, the experiment continued. If the dog failed to meet the criteria or did not consume the food rewards in the first three trials, the dog would be excluded due to lack of interest in food. In the testing condition, each trial started with the robot asking E1 for a food reward. E1 put the food reward in between the robot's fingers. The robot then called the dogs' name with raised pitch and showed the food reward to the dog. Afterwards, the robot dropped the treat on the plate in front of it. After the robot said 'okay', the dog was allowed to approach and consume the food reward.

In the control condition, each trial also started with the loudspeaker asking for the food reward from E1. E1 put the food reward on the plate in front of the loudspeaker. The dog was only allowed to approach and consume the food reward after the loudspeaker said 'okay'.

3.3.5 Phase 5: Experiment 2–Robot/Loudspeaker Gave 'Sit' Commands . Experiment 2 tested whether dogs would follow the 'sit' commands given by the robot/loudspeaker. This phase started with the guardian commanding the dog to sit to ensure that the dog could follow the 'sit' command from the guardian. Dogs who failed to follow the guardians' commands were excluded.

In the testing condition, each trial started with the robot asking E1 for a food reward and E1 placing the food reward in between the robot's fingers. The robot then called the dogs' name with raised pitch and showed the food reward to the dog. If the dog already sat, the robot would ask the guardian to let the dog stand up on all four legs. Afterwards, the robot would give the 'sit' command by saying '<dog name>, Sit!'. If the dog sat, the robot would drop the food reward on the plate in front of it and said 'okay' to allow the dog to consume the food reward. Otherwise, the robot would not drop the food reward and continue to the next trial.

There were two sections in the testing condition in this phase. In Section 1, the guardian was allowed to encourage the dog to sit or provide hints to sit. In Section 2, the guardian was not allowed to provide any encouragement or hints. In each section, if the dog could follow the 'sit' commands in two consecutive trials within a maximum of six trials in each section, the experiment moved on to the next section.

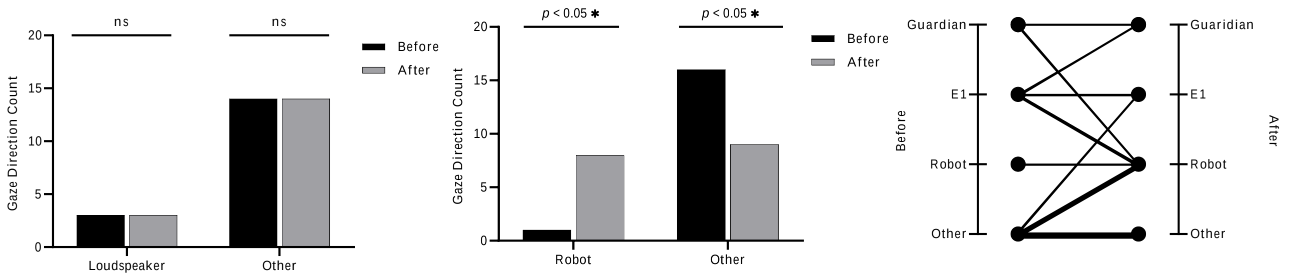
In the control condition, each trial started with the loudspeaker commanding the dog to sit by saying '<dog name>, Sit!'. The loudspeaker would praise the dog if the dog sat, but be quiet otherwise. The food reward neither appeared nor was provided to the dog when they sat.

There was only one section in the control condition. A total of four trials were administered because previous studies with the Rover@Home system [23] did not show adequate evidence that dogs follow commands from a loudspeaker. More trials in the control condition may disinterest the dog and cause the dog to disengage.

4 RESULTS

4.1 Experiment 1–Responding to Calling the Name

Experiment 1 involved a mixed 2×2 design. The independent variables were the condition and the timing variable. The condition had two levels: the testing condition with the robot and the control condition with the loudspeaker. The timing variable also had two levels: before calling dogs' names, and after calling dogs' names. The dependent variable was the gaze targets, which also had two levels: gazing at the robot, and gazing at other targets. The gaze targets before the names were called were video-coded as the gaze targets that dogs attended to immediately prior to their names being called. The gaze targets after the names were called were



(a) Gaze Targets in the Control Condition

(b) Gaze Targets in the Testing Condition

(c) Individual Gaze Targets in the Testing Condition

Figure 4: Results of Experiment 1–Responding to Calling the Names. (a) Control Condition. Dogs gaze directions and the timing variable were independent (p > .05). (b) Testing Condition. Dogs gaze directions and the timing variable were dependent (p < .05) (c) Testing Condition: The gaze targets of each individual before and after the robot called the names. The line width is proportional to the number of individuals making such gaze changes.

video-coded as the first agent that dogs gazed at within the one second after calling the names.

In the testing condition, only one dog looked at the robot before the robot called the dogs' names. After the robot called the dogs' names, eight dogs looked at the robot. Fisher's exact test showed that dog gaze directions and the timing variable were dependent in the testing condition ($p < .05$). In the control condition, three dogs looked at the loudspeaker before the loudspeaker said the dogs' name. Still three dogs looked at the loudspeaker after the loudspeaker called the dogs' names. Fisher's exact test showed that dogs gaze directions and the timing variable were independent in the control condition ($p > .05$). Figure 4a and 4b show the results in the robot and control condition, respectively. Figure 4c shows the results of further analysis of the performance of each dog in the testing condition with 4-levels of gazing targets specified: the robot, E1, the guardians and other targets. Before the robot called the dogs' name, one dog gazed at the robot, four dogs at E1, two dogs gazed at the guardian and ten dogs gazed at other targets. After the robot called the names, eight dogs gazed at the robot, two dogs gazed at E1, two dogs gazed at guardians and five dogs gazed at other targets.

4.2 Experiment 2–Responding to 'Sit' Commands: Performances

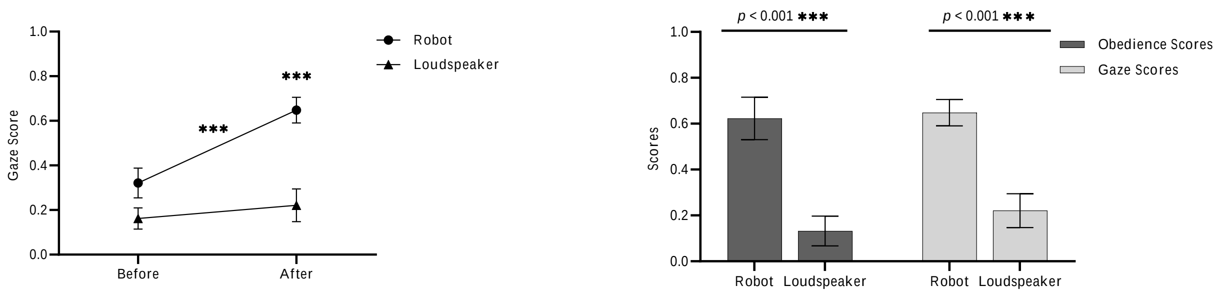
In Experiment 2, we evaluated dogs' performances and dogs' gazing targets when the robot/loudspeaker gave the 'sit' commands. Their performances were evaluated with obedience scores, following the method used in Pongrácz et. al's study [22]. The obedience score of each dog was calculated using the total number of trials in which the dog actually sat down divided by the total number of 'sit' commands given by the robot/loudspeaker. Successfully following the robot/loudspeaker's commands was defined as the dog sitting or lying down from the standing position within five seconds after the robot/loudspeaker gave the 'sit' commands. Lying down was also considered successful because pet dogs were generally not very well trained, and guardians may have rewarded the dogs for lying down after giving them the 'sit' commands at home. Trials were excluded if the dogs were already seated before the robot/loudspeaker gave the commands. On average, dogs participated 2.647 trials (SE = 0.308) in Section 1 with guardians' encouragement, and participated 2.471 trials (SE = 0.322) in Section 2 without guardian's encouragement.

A paired t-test showed that there was no significant difference between the obedience scores in Section 1 (M = 0.624, SD = 0.415) and in Section 2 (M = 0.632, SD = 0.416) ($t(16) = -0.105$, $p > 0.05$). Therefore, the results of the two sections were combined and the obedience score of each dog was recalculated in the testing condition. A t-test showed that the adjusted obedience scores in the testing condition (M = 0.623, SD = 0.381) was significantly higher than the obedience scores in the control condition (M = 0.132, SD = 0.267) ($t(28.641) = 4.348$, $p < .001$). Moreover, 65% dogs followed the robot's commands, while only 13% followed the loudspeaker's commands. In the first, second, third and fourth trials, 13%, 12%, 12% and 18% dogs followed the loudspeaker's commands, respectively.

4.3 Experiment 2–Responding to 'Sit' Commands: Gaze Targets

When analyzing the dogs' gaze targets, we combined the data from Section 1 and Section 2 in the testing condition. Not all guardians provided encouragement, and those that did generally provided it several seconds after the robot/loudspeaker gave the 'sit' commands. In contrast, gaze targets were coded to be the targets that dogs gazed at within one second after the robot/loudspeaker gave the commands. Therefore, the encouragement later in the trial should not affect the gaze targets.

The gaze targets were also scored in a way similar to the obedience score with the numerator being the count of dogs gazing at the robot and the denominator remains the total number of 'sit' commands given by the robot/loudspeaker. In the testing condition, the gaze score before the robot gave the 'sit' commands was 0.321 (SD = 0.275), and the gaze score after the robot gaze the 'sit' command was 0.648 (SD = 0.238). In the control condition, the gaze score before the speaker gave the 'sit' commands was 0.162 (SD = 0.196), and the gaze score after the loudspeaker gave the 'sit' commands was



(a) Gaze Scores Before and After the Robot/Loudspeaker Gave the 'Sit' Commands (b) Obedience Scores and Gaze Scores After the Robot/Loudspeaker Gave the 'Sit' Commands

Figure 5: Results of Experiment 2–Responding to ‘Sit’ Commands. (a) Gaze Score. Dogs gazed at the robot significantly more after the robot gave the ‘sit’ command in the testing condition ($t(16) = -4.906$, $p < .001$), but not in the control condition ($F(1, 32) = 6.749$, $p > .05$, $\eta_p^2 = 174$). Before the robot/loudspeaker gave the commands, dogs gazed at the loudspeaker and the robot equally often ($t(32) = 1.944$, $p > .05$), and gazed more at the robot after the robot gave the commands ($t(32) = 4.561$, $p < .001$). (b) Obedience score and the gaze score after the robot/loudspeaker gave the commands. The adjusted obedience scores ($t(28.641) = 4.348$, $p < .001$) and the gaze scores ($t(32) = 4.561$, $p < .001$) in the testing condition were significantly higher than ones in the control condition. There was a positive correlation between the obedience scores and the gaze scores ($r = 0.418$, $n = 34$, $p < .05$).

0.221 ($SD = 0.305$). A two-way mixed ANOVA with repeated measurements showed that there was a significant difference between the gaze scores before and after the robot/loudspeaker provided the ‘sit’ command ($F(1, 32) = 13.967$, $p < .001$, $\eta_p^2 = .304$), there was a significant difference between conditions ($F(1, 32) = 16.933$, $p < .001$, $\eta_p^2 = .346$), and there was an interaction effect ($F(1, 32) = 6.749$, $p < .05$, $\eta_p^2 = .174$). Specifically, in the testing condition, dogs gazed at the robot significantly more after the robot gave the ‘sit’ command than before the robot gave the command ($t(16) = -4.906$, $p < .001$). There was no such effect in the control condition ($F(1, 32) = 6.749$, $p > .05$, $\eta_p^2 = 174$). Before the robot/loudspeaker gave the commands, dogs gazed at the loudspeaker and the robot equally often ($t(32) = 1.944$, $p > .05$), and gazed more at the robot after the robot gave the commands ($t(32) = 4.561$, $p < .001$). Figure 5a shows the results.

Figure 5b shows the gaze scores after the robot/loudspeaker gave the commands and the obedience scores in the testing condition and in the control condition. There was a positive correlation between the obedience scores and the gaze scores ($r = 0.418$, $n = 34$, $p < .05$).

5 DISCUSSION

The present study demonstrated that with limited exposure, dogs attended to the robot significantly more than the loudspeaker when the robot/loudspeaker called the dogs’ names. This study also showed that the dogs followed the robot’s ‘sit’ commands more than the commands from the loudspeaker after longer exposures with the robot/loudspeaker. When given the ‘sit’ commands, dogs also focused more on the robot than on the loudspeaker. There was a moderate correlation between following the ‘sit’ commands and focus on the object who gave the ‘sit’ commands in general.

This is the first study that illustrate that dogs show social responses after interacting with a social robot briefly, and follow the commands from the social robot, similar to a human, after longer interactions.

In this study, other than the significant results from statistics, some interesting behaviors are also worth mentioning. Dog D1 was generally more vocal than other dogs. It continuously barked while the robot was talking, but ceased to bark when the robot issued verbal commands. Conversely, Dog D2 turned to gaze at and began barking at the robot right after the robot gave the commands, but ultimately complied with the commands. D2 did not bark otherwise while the robot was talking. Additionally, dog D3 did not follow the robot’s ‘sit’ commands, but turned to stare at the robot right after the robot said ‘sit’. In contrast, when the robot was not issuing commands, D3 explored the room and did not attend to the robot. These behaviors were not observed in the control condition in which the dogs generally ignored the loudspeaker. These behaviors may also strengthen the conclusion that dogs could socially interact with a social robot.

The dogs in our study reacted socially to a social robot, and the robot seemed to affect the dogs’ behaviors. One potential reason is that the dogs could have perceived the robot as an agent. People have a tendency to perceive robots as agents, rather than as machines. Such tendency in humans is expressed with explicit social behaviors (e.g., gaze, verbal communication, etc.) when interacting with robots. Short et. al [28] showed increased length of gaze duration and prolonged verbal communications in humans after the robot cheated in a rock-paper-scissor game. In the present study, the social behaviors of the robot may have triggered the dogs to perceive the robot as an agent, and ultimately resulted in the dogs’ tendency to socially interact with the robot and follow the commands from the robot.

Embodiment may be another possible explanation for our results. Previous studies showed that embodied robots could increase compliance in humans [2], and an embodied tutoring robot could improve humans’ performances in problem solving [17]. The robot in this study was embodied. In contrast, the loudspeaker, which only provided verbal cues, could be seen as a virtual agent that was disembodied. To explore this possibility, further studies could compare how dogs interact with a robot using verbal cues and 2D visual cues which are projected videos, with how dogs interact with a robot that provide verbal cues and 3D visual cues which is to include a robot physically present during testing.

Another possible explanation for our results is the multimodal nature of the cues provided. In the control condition, only verbal cues were provided. In the testing condition, both verbal and visual cues were provided. With the combination of both types of cues, dogs might find it easier to associate the commands with the robot. In the control condition, dogs could only rely on the verbal cues to localize the source of commands. However, we observed that many dogs attended to the loudspeaker only during the first few utterances. Dogs quickly lost interest in the loudspeaker and continued to explore the room. Moreover, dogs could localize the source of the sound locations better than humans [6]. Humans without hearing impairment generally do not experience much difficulties to localize the source of the sound. Therefore, we considered this explanation is less likely compared to the explanation of agency and embodiment.

One could argue that the dogs followed the robot's commands simply due to the presence of the human experimenters and their guardians. Alternatively, they could have mistakenly recognized E1 as the agent who gave the 'sit' commands since E1 was positioned beside the robot. However, E1 and the guardians were also present in the control condition. The dogs did not follow the 'sit' command as often as the dogs did in the testing condition. Second, when the robot gave the commands, the dogs gazed at the robot more often than any other agents, including E1. Dogs did not gaze at the robot more than the loudspeaker before the commands were given. Therefore, dogs' gaze patterns implied their understanding of the robot as the source of the 'sit' commands. Their behaviors in this study were more likely to be the result of following the robot's commands rather than complying to cues from the humans or transferring the source of auditory cues of the command 'sit' to E1.

Three obvious differences presented between the control condition and the testing condition that could result in confounds. First, the non-encouragement phase was present in the testing condition but not in the control condition. The results showed that dogs in the control condition with encouragement followed the commands significantly less the dogs in the testing condition without encouragement. Such results actually strengthened the interpretation that dogs followed commands from the social robot but not the loudspeaker.

Second, the number of trials each dog participated varied in the testing condition. The dogs in a previous study did not follow the commands given by the loudspeakers [23]. Therefore, we did not expect the dogs to follow the commands in the control condition. In the control condition, the number of trials were fixed to keep the section short in order to retain the engagement of the dogs. Dogs in the control condition participated in four trials, and dogs in the testing condition experienced 5.118 trials. The effect of the extra 1.118 trials should be minuscule, and it would be unlikely to result in significant improvements in the performances.

Third, in experiment 2, dogs provided with food rewards were in the testing condition, but they were not provided with rewards in the control condition. We acknowledge that the absence of food rewards in the control condition could be a potential confound variable. However, the food rewards were not the major cause of the significant behavioral differences in the testing and control conditions. Resner [23] showed that dogs did not follow the commands from a loudspeaker with food rewards, even when professional trainers attempted to train clicker-trained dogs to use the loudspeaker system with food rewards. Moreover, our results showed that dogs followed the robot's commands considerably more than the dogs in the control condition during the first trials, when the dogs were not previously being rewarded to follow the commands. Furthermore, in the control condition, the dogs followed the commands from the loudspeaker 13%, 12%, 12% and 18% in the first, second, third and fourth trials, respectively. These results suggested the dogs were not demotivated to continue follow the loudspeaker's commands even without food rewards. Including food rewards in the control condition would have introduced more confound variables. The loudspeaker could not provide food rewards without a human experimenter. With a human participating in providing food rewards, the dog could have interpreted the commands were given under E1's control, rather than by the loudspeaker itself.

To help to design social robots to work with dogs in the future, we would also like to share our experiences from our pilot testing. Large movements from the robot (e.g., walking) or movements with increased level of noise should be avoided. They provide too much novel effect for the dogs and could startle them. It is also important to give sufficient amount of time for the dog to acclimate with the social robot. Furthermore, the guardian plays in important role to direct the attention of the dog to the social robot and to encourage the dog to interact with the robot. Future studies could explore what robot behaviors elicit dogs' responses. It is still not known yet whether the appearance of the robot (e.g., more human-like), the smell of the robot, adding the sound of heartbeat, etc., would affect the dog's responses. Future studies could also address what kind of behaviors in dogs could be elicited (e.g., social referencing).

6 CONCLUSION

The contribution of this study is that it is the first study to show direct evidence that dogs responded to a social robot, and that dogs complied to the social robot's commands. The study has practical implications for the design of personal social robots, as well as theoretical significance for HRI studies and dog cognition studies.

ACKNOWLEDGMENTS

This work was supported by the NSF-REU, award #1659085, and National Science Foundation, award #IIS-1813651, and Office of Naval Research, award #N00014-18-1-2776. We would like to thank all the dog participants and their guardians of our study for their time. We also want to thank Abigail Waugh, Caleb Kim, Skylar Regan, Chavely Calleja, Zoë Stublarec, Maisa Crispino, Nicholas Peters, Emani Brown, and Riley Martin for help with running the experiment and coding the videos, and Jake Brawer and Yuen Sang Albert Wong for help in proofreading this paper.

REFERENCES

[1] Judit Abdai, Anna Gergely, Eszter Petró, József Topál, and Ádám Miklósi. 2015. An investigation on social representations: inanimate agent can mislead dogs (Canis familiaris) in a food choice task. *PloS one* 10, 8 (2015), e0134575.

[2] Wilma A Bainbridge, Justin W Hart, Elizabeth S Kim, and Brian Scassellati. 2011. The benefits of interactions with physically present robots over video-displayed agents. *International Journal of Social Robotics* 3, 1 (2011), 41–52.

[3] Anna Gergely, Judit Abdai, Eszter Petró, András Kosztolányi, József Topál, and Ádám Miklósi. 2015. Dogs rapidly develop socially competent behaviour while interacting with a contingently responding self-propelled object. *Animal Behaviour* 108 (2015), 137–144.

[4] Anna Gergely, Eszter Petró, József Topál, and Ádám Miklósi. 2013. What are you or who are you? The emergence of social interaction between dog and an unidentified moving object (UMO). *PloS one* 8, 8 (2013), e72727.

[5] ME Goddard and RG Beilharz. 1984. A factor analysis of fearfulness in potential guide dogs. *Applied Animal Behaviour Science* 12, 3 (1984), 253–265.

[6] Jay M Goldberg and Paul B Brown. 1969. Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: some physiological mechanisms of sound localization. *Journal of neurophysiology* 32, 4 (1969), 613–636.

[7] Elena Corina Grigore, Andre Pereira, Jie Jessica Yang, Ian Zhou, David Wang, and Brian Scassellati. 2016. Comparing ways to trigger migration between a robot and a virtually embodied character. In *International Conference on Social Robotics*. Springer, 839–849.

[8] Elena Corina Grigore, Andre Pereira, Ian Zhou, David Wang, and Brian Scassellati. 2016. Talk to me: Verbal communication improves perceptions of friendship and social presence in human-robot interaction. In *International conference on intelligent virtual agents*. Springer, 51–63.

[9] Brian Hare, Michelle Brown, Christina Williamson, and Michael Tomasello. 2002. The domestication of social cognition in dogs. *Science* 298, 5598 (2002), 1634–1636.

[10] Brian Hare, Josep Call, and Michael Tomasello. 1998. Communication of food location between human and dog (Canis familiaris). *Evolution of communication* 2, 1 (1998), 137–159.

[11] Angie M Johnston, Paul C Holden, and Laurie R Santos. 2017. Exploring the evolutionary origins of overimitation: a comparison across domesticated and non-domesticated canids. *Developmental science* 20, 4 (2017), e12460.

[12] Angie M Johnston, Yiyun Huang, and Laurie R Santos. 2018. Dogs do not demonstrate a human-like bias to defer to communicative cues. *Learning & behavior* 46, 4 (2018), 449–461.

[13] Tammy King, Paul H Hemsworth, and Grahame J Coleman. 2003. Fear of novel and startling stimuli in domestic dogs. *Applied Animal Behaviour Science* 82, 1 (2003), 45–64.

[14] Enikö Kubinyi, Ádám Miklósi, Frédéric Kaplan, Márta Gácsi, József Topál, and Vilmos Csányi. 2002. Can a dog tell the difference? dogs encounter AIBO, an animal-like robot in two social situations. In *Proceedings of the seventh international conference on simulation of adaptive behavior on From animals to animats*. MIT Press, 403–404.

[15] Enikő Kubinyi, Ádám Miklósi, Frédéric Kaplan, Márta Gácsi, József Topál, and Vilmos Csányi. 2004. Social behaviour of dogs encountering AIBO, an animal-like robot in a neutral and in a feeding situation. *Behavioural processes* 65, 3 (2004), 231–239.

[16] Gabriella Lakatos, Mariusz Janiak, Lukasz Malek, Robert Muszynski, Veronika Konok, Krzysztof Tchon, and Á Miklósi. 2014. Sensing sociality in dogs: what may make an interactive robot social? *Animal cognition* 17, 2 (2014), 387–397.

[17] Daniel Leyzberg, Samuel Spaulding, Mariya Toneva, and Brian Scassellati. 2012. The physical presence of a robot tutor increases cognitive learning gains. In *Proceedings of the annual meeting of the cognitive science society*, Vol. 34.

[18] Á Miklösi, Rob Polgárdi, J Topál, and V Csányi. 1998. Use of experimenter-given cues in dogs. *Animal cognition* 1, 2 (1998), 113–121.

[19] Maretta Morovitz, Megan Mueller, and Matthias Scheutz. [n. d.]. Animal-Robot Interaction: The Role of Human Likeness on the Success of Dog-Robot Interactions. In *1st International Workshop on Vocal Interactivity in-and-between Humans, Animals and Robots*.

[20] András Péter, Ádám Miklósi, and Péter Pongrácz. 2013. Domestic dogs'(Canis familiaris) understanding of projected video images of a human demonstrator in an object-choice task. *Ethology* 119, 10 (2013), 898–906.

[21] Robert Plutchik. 1971. Individual and breed differences in approach and withdrawal in dogs. *Behaviour* 40, 3-4 (1971), 302–311.

[22] Péter Pongrácz, Ádám Miklósi, Antal Dóka, and Vilmos Csányi. 2003. Successful application of video-projected human images for signalling to dogs. *Ethology* 109, 10 (2003), 809–821.

[23] Benjamin Ishak Resner. 2001. *Rover@ Home: Computer mediated remote interaction between humans and dogs*. Ph.D. Dissertation. Massachusetts Institute of Technology.

[24] Laurel D Riek. 2012. Wizard of oz studies in hri: a systematic review and new reporting guidelines. *Journal of Human-Robot Interaction* 1, 1 (2012), 119–136.

[25] Nicole Salomons, Michael van der Linden, Sarah Strohkorb Sebo, and Brian Scassellati. 2018. Humans conform to robots: Disambiguating trust, truth, and conformity. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 187–195.

[26] Brian Scassellati, Jake Brawer, Katherine Tsui, Setareh Nasihati Gilani, Melissa Malzkuhn, Barbara Manini, Adam Stone, Geo Kartheiser, Arcangelo Merla, Ari Shapiro, et al. 2018. Teaching language to deaf infants with a robot and a virtual human. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 553.

[27] Sarah Strohkorb Sebo, Priyanka Krishnamurthi, and Brian Scassellati. 2019. âĂIJI Don't Believe YouâĂİ: Investigating the Effects of Robot Trust Violation and Repair. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 57–65.

[28] Elaine Short, Justin Hart, Michelle Vu, and Brian Scassellati. 2010. No fair!! an interaction with a cheating robot. In *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 219–226.

[29] Sarah Strohkorb Sebo, Margaret Traeger, Malte Jung, and Brian Scassellati. 2018. The ripple effects of vulnerability: The effects of a robot's vulnerable behavior on trust in human-robot teams. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 178–186.