From One to Another: How Robot-Robot Interaction Affects Users' Perceptions Following a Transition Between Robots

Xiang Zhi Tan
Robotics Institute
Carnegie Mellon University
Pittsburgh, USA
zhi.tan@ri.cmu.edu

Samantha Reig
Human-Computer Interaction Institute
Carnegie Mellon University
Pittsburgh, USA
sreig@cs.cmu.edu

Elizabeth J. Carter
Robotics Institute
Carnegie Mellon University
Pittsburgh, USA
ejcarter@andrew.cmu.edu

Aaron Steinfeld
Robotics Institute
Carnegie Mellon University
Pittsburgh, USA
steinfeld@cmu.edu

***Abstract*—Human-robot interactions that involve multiple robots are becoming common. It is crucial to understand how multiple robots should transfer information and transition users between them. To investigate this, we designed a 3 x 3 mixed-design study in which participants took part in a navigation task. Participants interacted with a stationary robot who summoned a functional (not explicitly social) mobile robot to guide them. Each participant experienced the three types of robot-robot interaction: *representative* (the stationary robot spoke to the participant on behalf of the mobile robot), *direct* (the stationary robot delivered the request to the mobile robot in a straightforward manner), and *social* (the stationary robot delivered the request to the mobile robot in a social manner). Each participant witnessed only one type of robot-robot communication: *silent* (the robots covertly communicated), *explicit* (the robots acknowledged that they were communicating), or *reciting* (the stationary robot said the request aloud). Our results show that it is possible to instill socialness in and improve likability of a functional robot by having a social robot interact socially with it. We also found that covertly exchanging information is less desirable than reciting information aloud.**

***Index Terms*—robots; human-robot interaction, user study; multi-robot; multi-robot-human interaction; transition; transfer**

I. Introduction

Robots in public and private settings are no longer a far-off dream but a growing reality. With deployments of service and social robots in airports [1], shopping malls [2], hotels [3], and museums [4]–[6], interacting with commercial robots such as Softbank's Pepper and Savioke's Relay will become a staple of our daily lives. Yet each of these robots is unlikely to be the only robot in a building: it is more likely that many robots will collaborate to provide services. How multiple robots should interact with humans in such scenarios, particularly when they exchange information about users, remains an open question. Humans are limited in their communication capabilities by the necessity of verbal or nonverbal conversational cues in face-to-face interaction, the need to rely on an external medium for remote interaction, and imperfect calibration of semantic meaning. However, robots are not subject to these constraints: they can communicate covertly through electronic signals, have near-perfect information transfer, and can even share the same personality across different robot embodiments [7].

This work was funded by the National Science Foundation (CBET-1317989 & SES-1734456) and the National Institute on Disability, Independent Living, and Rehabilitation Research (90DPGE0003).

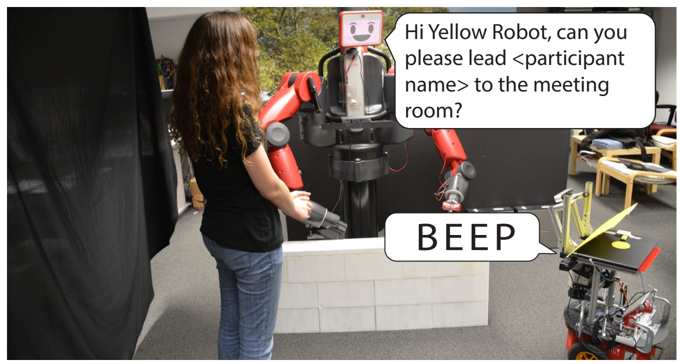

Fig. 1. Stationary robot *recites* a request to mobile robot in a *social* manner.

As an initial step, we explore the transition of a human from one robotic system to another. In human-human interaction, starting an interaction with one person and transitioning to interacting with another is common. For example, a receptionist may greet a visitor at the reception area in a building, provide them with a security badge, and then instruct them to follow a guide to their destination (and thus the visitor "transitions" from one interactant to another). For another example, consider a customer service exchange in which an employee lacks the skills or authority to complete a request and therefore asks their manager to take over. Such transitions may be planned (as in the receptionist/guide scenario), or they may emerge on the spot (as in the cashier/manager scenario).

Transitions like these will occur when humans interact with multiple robots. This is especially likely for heterogeneous robot teams in which one of the robots lacks the capabilities to complete a task. In a receptionist scenario, a stationary receptionist robot might not be able to guide a visitor to a destination because it is immobile or needs to remain at its station to fulfill other duties. In this case, the stationary robot would need to summon a mobile robot to guide the user to the destination. The other robot may also be specialized in mobility and lack social capabilities. Should robots in these situations invoke the norms and mimic the behavior of humans? Do they need to verbally communicate their intent and acknowledge users' requests? These are the questions we address in this research.

Effective communication between robots may have additional benefits beyond enhancing user experience. Equipping robots with social capabilities increases user engagement and enjoyment [8], [9]. Yet some robots, especially those that are intended to be purely functional, may not be designed to have human-like or animal-like features. By observing social behavior between a social robot and a functional (nonsocial) robot, users may attribute social intelligence, trust, and other properties normally associated with social robots to the nonsocial robot, even in the absence of its own social expressions. By being deliberate in the design of the way one robot treats another, we may be able to build up these attributions toward a functional robot that does not exhibit social qualities.

To explore how information transfer and social interactions between robots affect user perceptions, we tested three types of information transfer and three degrees of sociality in an interaction among two robots and a human participant. Our study showed that the way a stationary social robot treats a functional mobile robot with minimal social capabilities changes the way humans perceive the mobile robot.

II. Related Work

A. Multiple-Robot Interaction in HRI

Prior work explored different properties of interactions between humans and multiple robots. A variety of issues has been examined, including spatial positioning [10], effects of the number of robots and their social features [11], [12], and people's decisions to conform to robots' opinions [13]. The applications of such systems have also been explored in educational [14] and shopping [15] settings. While most of this work explored humans interacting with multiple identical robots, some has explored how being exposed to different robots influences perceptions. Fraune and colleagues [16] found that robots were perceived more negatively when presented in a group of identical robots than alone or in a diverse group of robots. One of the novel contributions of our work is that our participants did not start the interaction with multiple robots. Instead, we examined an interaction in which a person began interacting with one robot and then transitioned via the interaction to a different robot.

B. Robot-to-Robot Communication in HRI

Previous work explored how robot-to-robot communication can be used to calibrate humans' expectations of robots' capabilities [17]. Designing for overt communication between robots has also been shown to be an engaging way to convey information [18], [19] and improve conversational coherency [20]. Some researchers have studied robot-to-robot communication to better understand the role of robots as sidekicks in interactions with and without humans [21], [22].

An early exploration of human perception of multi-robot interaction investigated how people interpreted a conversation between two robots [17]. After observing the interaction, participants were able to infer the robots' verbal and nonverbal communicative capabilities. Later, Fraune and Sabanovic [23] examined how different types of inter-robot communication (none, loud, and silent) between basic functional robots affected the attitude of a bystander. The study did not find any significant differences between conditions. The authors hypothesized that the result arose from participants assigning *groupness* rather than treating the robots as individual social entities. Our study differs from these in that participants directly interact with two robots rather than just observing them.

Our study is similar to previous work [24] that explored how different robot-to-robot communication affected user preferences. In a simulated nuclear disaster scenario, participants issued commands to the robots. The robots communicated either verbally or covertly. Interestingly, participants described covert communication between the robots as creepy. Our work differs from [24] in that (1) we create a transition rather than fostering ongoing collaboration and (2) our interaction occurs in a setting where robots work together to provide a service.

III. Method

To explore how a social, stationary robot and a functional, mobile robot should communicate when "handing off" a user, we designed a study about how Information Transfer and Stationary Robot Behavior influenced participants' preferences and perceptions. We created a navigation scenario in which a person requested assistance from a stationary robot that then summoned a mobile robot to lead the person to a destination. The study was approved by our Institutional Review Board.

A. Study Design

We designed a 3 x 3 mixed-design experiment with Information Transfer as the between-subjects manipulation and Stationary Robot Behavior as the within-subjects manipulation. Information Transfer explored different ways for robot-robot information transfer to be conveyed to the user. This was not to determine the actual best way for information to be transferred (electrical signaling is often the best option due to low noise, high reliability, and high bandwidth); instead, it was to learn how robots should indicate to their users that certain information has been shared between two robots. Our Information Transfer conditions were as follows:

- *Silent* – The stationary robot did not repeat the user's request and did not explicitly acknowledge that the request had been transferred to the mobile robot.
- *Explicit* – The stationary robot did not repeat the user's request but did explicitly acknowledge that the information had been sent to the mobile robot.
- *Reciting* – The stationary robot recited the user's request out loud to the mobile robot.

Stationary Robot Behavior describes how a stationary robot interacts with a mobile robot that lacks speech capabilities. The conditions were as follows:

- *Representative* – The stationary robot did not speak directly to the mobile robot, but instead spoke to the participant on behalf of the mobile robot.
- *Direct* – The stationary robot spoke directly to the mobile robot, delivering the participant's request in a complete sentence.
- *Social* – The stationary robot spoke directly to the mobile robot, supplementing the participant's request with social conversational behavior.

In our mixed-design study, each participant was assigned to one of the Information Transfer conditions and experienced each of the three Stationary Robot Behavior conditions. The order of the Behavior conditions was counterbalanced and the Transfer conditions were spread out equally across the 6 unique permutations of Stationary Robot Behavior.

B. Hypotheses

Our hypotheses predicted that both Information Transfer and Stationary Robot Behavior would affect participants' preference and perception of both robots.

H1. Participants will perceive the stationary robot to be more social, competent, and likable in the *social* condition than in the *direct* and *representative* conditions.

H2. Participants will perceive the mobile robot to be more social, competent, and likable in the *social* condition than in the *direct* and *representative* conditions.

H3. Information Transfer will have an effect on participants' perception of both robots.

(a) Participants will perceive the robots to be more competent in the *reciting* condition.

(b) Participants will perceive lower competence in the mobile robot and be more wary of and disturbed by the stationary robot's behavior in the *silent* condition.

H4. Participants will be more likely to see the robots as equals in the *social* condition.

H5. Participants will have a higher preference to work with the mobile robot they encounter in the *social* condition.

C. System and Study Setup

The study was conducted in a lab space on Carnegie Mellon University's Pittsburgh campus. The stationary robot was a humanoid Baxter robot by Rethink Robotics with speakers added for a clear robot voice and a camera mounted on the chest to record participant reactions. The Baxter robot wore a different name tag in each of the Behavior conditions to add to the perception that the context had been switched between each of the 3 interactions that constituted the within-subjects manipulation. The robot was controlled through the ROS-based SDK [25]. Realistic speech was generated using the Amazon AWS Polly SDK [26], and natural human responses were recognized through a pipeline consisting of Google Cloud Speech and SNIPS Natural Language Understanding Engine [27].

The mobile robot was a custom-modified Mobile Robots P3DX. The robot was controlled with ROS and could autonomously navigate to multiple waypoints throughout the study area using the ROS Navigation Stack [28]. Additional motion corrections could be made using a joystick controller if the robot moved erroneously. The robot was decorated with colorful accessories to allow participants to easily distinguish between trials and enhance the illusion that each trial had a different mobile robot. All participants saw a "blue robot", a "green robot", and a "yellow robot". There were two mappings of colors to Behavior conditions.

Because both robots are independent units that operate separately by default, we wrote a ROS package, inspired by [17], to facilitate communication between the two robots through the ROS Bridge protocol [29]. This package allowed us to send signals between the robots so that one robot could tell the other robot to execute certain actions or that an action had been completed. In our study, the stationary robot wirelessly coordinated the experience and told the mobile robot when certain behaviors were needed. To move the mobile robot, the stationary robot sent a signal with a waypoint (next to the participant, at the door to the lab, etc.) to the mobile robot. The mobile robot then autonomously planned and executed a path to the given waypoint. When the path had been completed, the mobile robot sent back a "done" signal to the stationary robot, who waited for that signal before continuing the interaction with the participant.

Both robots operated autonomously during the study. Occasionally, the experimenter adjusted the mobile robot's motion if it moved too slowly or too close to an obstacle. We have released the code behind the study online[1].

D. Procedure

After obtaining informed consent, the experimenter explained to the participant that they would take part in a scenario in which they required guidance to navigate to a room in an unfamiliar building. The experimenter told the participant that they would engage in the interaction four times—with the first being a practice run—and that they would verbally interact with a differently programmed robot during each trial.

To mitigate the impact of language processing failures, we told participants to be aware that the system was not perfect and might fail to recognize some commands, and we provided tips on what to do when failures occurred.

In the practice run, the participant interacted with the stationary robot who summoned the mobile robot to lead the participant from our lab to a door in the hallway. Once the mobile robot arrived at its destination, it beeped once to signal task completion. The experimenter then led the participant back to the room. The experimenter explained that the robot would drive itself back to the "charging station" outside of the lab. We chose to have the mobile robot park itself outside

[1]https://github.com/CMU-ARM/HRI19-MultiRobot-Transition-Study

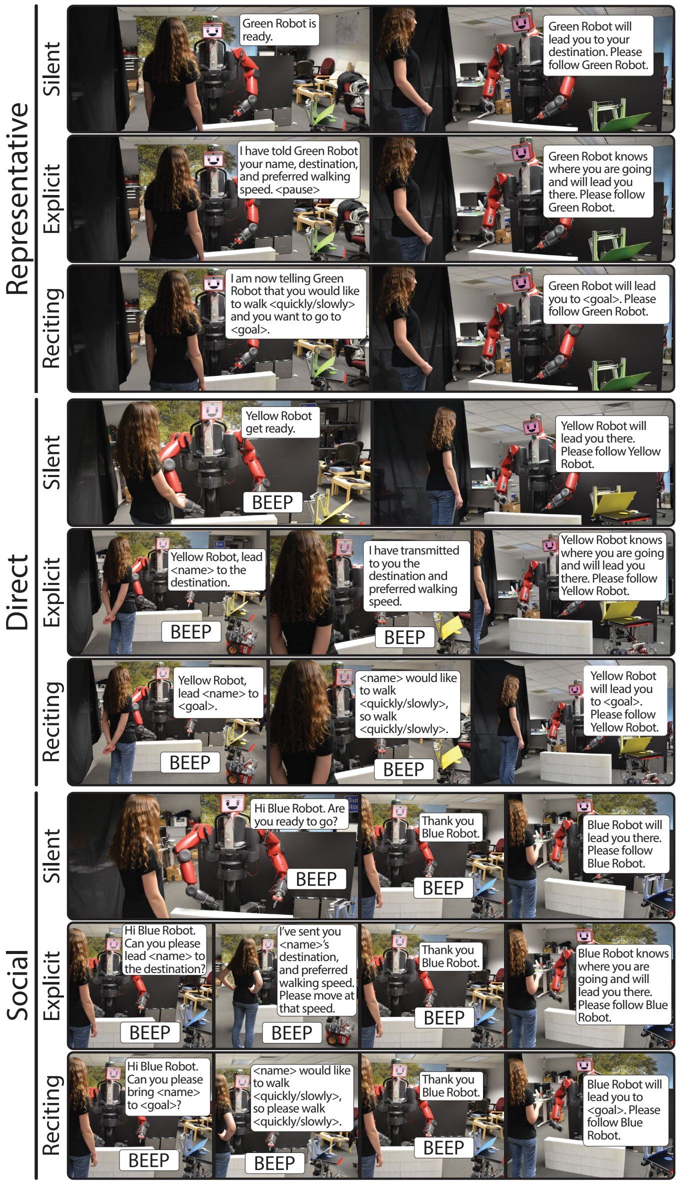

Fig. 2. Stationary and mobile robots' dialogue for each condition.

the lab (and out of sight of the participant) between trials for two reasons: (1) to prevent participants from thinking that the mobile robot gathered information by overhearing their conversation with the stationary robot rather than obtaining it directly from the stationary robot, and (2) to contribute to the illusion that there were three different mobile robot behavior "programs" by having the robot leave and return with a different appearance. When the participant re-entered the lab, they completed a demographic survey and a questionnaire (Q1) about the practice run.

Following Q1, the experimenter explained that the participant would need to follow the robot only to the door of the lab to complete each trial rather than all the way to a destination. We chose not to have the mobile robot navigate to a destination to ensure that the participant's experience and opinion of the mobile robot were based on the transition rather than its success in guiding the participant. The participant was also told that they would be prompted to select a walking speed ("quickly" or "slowly"), but that the speed selection would not determine the mobile robot's actual speed; this was to ensure that perceptions were not affected by the mobile robot's success or failure to match expectations of speed.

When the participant was ready, the experimenter pressed a button to start the next interaction. The stationary robot's face appeared on its screen. The stationary robot waited for the participant to initiate the interaction and began when it heard a prompt. The robot then greeted the participant and asked for the participant's name (which was entered in advance to ensure correctness). The robot then asked what kind of assistance the user required. Our system used the natural language pipeline to parse the request. Once the system correctly extracted the destination, it asked the participant to confirm, and then asked them for their preferred walking speed. Then, it summoned the mobile robot to a waypoint to the right of the participant. When the mobile robot arrived, it turned to the stationary robot and beeped to signal its arrival. In the *explicit* and *reciting* conditions, the stationary robot turned its arms and head towards the mobile robot. Dialogue for the condition was then executed (Figure 2), ending with the stationary robot instructing the participant to follow the mobile robot. The mobile robot moved outside of the lab, at which point the experimenter stopped the trial and administered the next questionnaire (Q2). The experimenter then exited the room, stating that they needed to switch the programs of both of the robots while the participant completed Q2. In reality, the experimenter only switched the program of the stationary robot, the name tag of the stationary robot, and the color accessories on the mobile robot. This interaction was then repeated two more times, once for each of the other within-subjects conditions. Each condition was associated with a fictional setting and a color (blue, yellow, green). To ensure that robot color was not a confound, we rotated the color order after 18 sessions (6 per between-subjects condition).

After completing Q2 for the last trial, the participant completed another questionnaire (Q3) comparing all three trials. We then conducted a semi-structured interview to gain further insight into participants' impressions. The study took about 30 minutes and participants were compensated 8 USD.

E. Measures

Because we were interested in the participants' perceptions of the interactions, we relied primarily on subjective measures. Our measures included several Likert scales drawn from prior work, experiment-specific forced-choice questions, yes-or-no questions about the robots' knowledge, and open-ended questions. When drawing from validated scales, we selectively omitted less relevant items to prevent survey fatigue.

1) Perception of Social Properties: To assess both robots, we combined sections of the Robotic Social Attributes Scale (RoSAS) [30] and the Godspeed questionnaire [31]. Participants were asked to rate the robot(s) with respect to 12 words

TABLE I
PARTICIPANT DEMOGRAPHICS (36 VALID SESSIONS)

Condition	Female	Male	Other	Age (Std. Dev.)
Silent	8	4	0	29.5 (16.2)
Explicit	9	3	0	27.1 (12.5)
Reciting	8	4	0	26.3 (8.5)

from the *warmth* and *competence* RoSAS factors. We also asked them to rate 3 words from Godspeed (Likable, Mean, and Friendly) to measure perceived *likability*.

2) Trust in Guide Robot: Participants' trust in the system was measured by 6 questions, 4 of which were modified from Jian's trust scale [32] and two that were specific to the task.

3) Open Ended Questions: At the end of each trial, we asked participants to describe the relationship between the robots and what they liked or disliked about the interaction. After the third trial, we asked participants how they believed information transfer had occurred and which pieces of information had been transferred between robots.

4) Preferred Robot: The final questionnaire also included forced-choice questions asking participants which of the three mobile robots they would most want to use, which one they felt most connected to, which one they preferred the least, which one they found most likable, and which one they believed to be the most knowledgeable. We also asked in which trial the stationary robot was most likable and least preferred.

F. Other Data

We recorded each session and logged how often the robot repeated a question due to a natural language pipeline failure.

G. Participants

We recruited 44 participants from the Pittsburgh metropolitan area using an online recruitment tool. Participants were between 18 and 61 years old. Eight participants were excluded due to logistical or technical issues, resulting in 36 valid sessions (12 per between-subject condition; Table I). Participants reported using computers on a near-daily basis, $M = 6.89, SD = 0.32$, on a 7-point Likert scale that ranged from Never (1) to Daily (7). Participants also reported some familiarity with robots, $M = 3.17, SD = 1.28$ on a 7-point scale.

IV. RESULTS

Unless otherwise noted, we analyzed the results by fitting a multilevel linear model using REstricted Maximum Likelihood (REML) [33], [34] for all continuous measures with Information Transfer and Behavior as fixed effects and participant as a random effect nested within Information Transfer. The number of system mistakes (times the robot repeated a question due to language pipeline errors) was treated as a covariate and included as a fixed effect to ensure that differences in ratings were due to our manipulation and not due to system usability. All post-hoc analyses used Tukey's Honest Significant Difference (Tukey HSD). We report significant differences ($p < 0.05$) and important trends ($p < 0.1$).

A. Measure Reliability and Confounds

The RoSAS *warmth* factor was reliable for both the stationary robot (Cronbach's $\alpha = 0.92$) and the mobile robot (Cronbach's $\alpha = 0.94$). The *competence* factor was also reliable for both robots (Cronbach's $\alpha = 0.96$ and 0.92, respectively). The Likability measure was calculated by averaging participants' responses to items pertaining to the robot's *Likability* and *Friendliness*, which had item reliability of Cronbach's $\alpha = 0.90$ for the stationary robot and Cronbach's $\alpha = 0.83$ for the mobile robot. We did not include the reverse coding of *meanness* in this index because it had a low correlation with the other items. Instead, we analyzed *meanness* individually for both robots. Five of the six items that assessed trust in the mobile robot were highly correlated ($r > 0.85$). The exception was the item "I am wary of the guide robot", which was weakly correlated with the other items ($r < 0.34$). We combined the five highly correlated items (Cronbach's $\alpha = 0.98$) as a measure of trust in the mobile robot. To assess the perceived relationship between robots, we constructed a *relationship* factor composed of responses to Likert items pertaining to beliefs that the robots *knew each other well*, *ignored each other* (reverse coded), and *liked each other* (Cronbach's $\alpha = 0.76$).

To evaluate the possible effect of the mobile robot's color, we included color in a similar multi-level linear model. We did not find significant effects of color on our dependent measures.

B. Perception of Mobile Robot

We measured participants' perceptions of the mobile robot through RoSAS and other measures. For the *warmth* measure, we found that Behavior had a significant effect, $F(2, 58.22) = 5.70$, $p = .006$. Pairwise analysis showed that participants felt that the *warmth* of the mobile robot was significantly higher in the *social* condition ($M = 3.45, SE = 0.33$) than in the *representative* condition ($M = 2.67, SE = 0.31$), $p = .004$. No other pairwise difference was found. We also found a trend wherein Stationary Robot Behavior impacted perceived competence of the mobile guide robot, $p = .092$. On the mobile robot's likability, we found a significant effect of the stationary robot's Behavior, $F(2, 58.09) = 6.62$, $p = .003$. Pairwise comparison showed that when the stationary robot was *social* toward the mobile robot ($M = 5.04, SE = 0.38$), the mobile robot was more likable than when the stationary robot was *representative* ($M = 3.99, SE = 0.37$), $p = .002$. No other pairwise difference was found. We also found a trend where Information Transfer influenced the participant's wariness of the mobile robot, $p = 0.057$. In particular, participants were more wary in the silent condition ($M = 3.12$) than in *reciting* condition ($M = 1.73$). We found no significant difference among factors for participants' perceived *meanness* or trust of the guide robot.

C. Perception of Stationary Robot

The stationary robot's Behavior significantly affected perceived *warmth*, $F(2, 58.19) = 8.53$, $p < .001$. Pairwise comparisons showed that participants rated the stationary robot

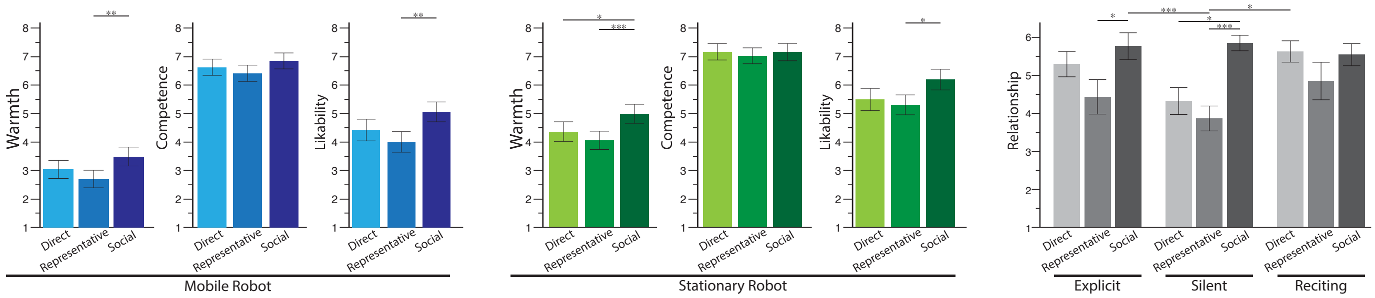

Fig. 3. Left and middle: Effect of Stationary Robot Behavior on social perceptions. Right: Interaction effect of Stationary Robot Behavior and Information Transfer on rating of robot-robot relationship. * means $p < .05$, ** means $p < .01$, and *** means $p < .001$. Error bars denote $\pm$ 1 standard error.

higher on *warmth* in the *social* ($M = 5.01, SE = 0.34$) condition than in the *direct* ($M = 4.37, SE = 0.33$) and *representative* ($M = 4.03, SE = 0.33$) conditions, $p = .028$ and $p < .001$, respectively. Though participants were treated equally by the stationary robot across conditions, their perceptions of its warmth changed when it treated the mobile robot in a nonsocial manner. We found a trend wherein Information Transfer influenced participants' *competence* ratings of the stationary robot, $p = 0.064$, such that participants viewed the stationary robot as more competent in the *reciting* condition ($M = 7.93$) than in the *silent* condition ($M = 6.36$). Participants' ratings of the stationary robot's likability were significantly affected by Information Transfer, $F(2, 31.84) = 3.90$, $p = .031$, and the stationary robot's Behavior, $F(2, 58.22) = 3.81$, $p = .028$. Pairwise comparisons revealed that participants perceived the stationary robot as more likable when it used *reciting* ($M = 6.60, SE = 0.53$) than when it was *silent* ($M = 4.55, SE = 0.52$), $p = .025$. Participants also perceived the robot to be more likable in the *social* condition ($M = 6.20, SE = 0.37$) than in the *representative* condition ($M = 5.30, SE = 0.35$), $p = 0.025$. Both Information Transfer and Behavior significantly affected *meanness* ratings for the stationary robot, $F(2, 34.47) = 3.31$, $p = .049$ and $F(2, 62.82) = 4.91$, $p = 0.011$. Again, pairwise comparisons showed that participants perceived that the robot was less mean in the *reciting* condition ($M = 1.31, SE = 0.37$) than in the *silent* condition ($M = 2.61, SE = 0.35$), $p = .038$. Participants rated the *social* ($M = 1.33, SE = 0.31$) stationary robot as significantly less mean than the *direct* ($M = 2.45, SE = 0.29$) stationary robot, $p = 0.011$. There was also a trend in which the stationary robot was perceived as less mean in the *social* condition than in the *representative* condition, $p = .053$.

D. Robot-Robot Relationship

We explored the perceived relationship between the robots via (1) a *relationship* factor and (2) an analysis of participants' responses to a free response question asking them to describe the robots' relationship after each trial. We found Behavior to have a significant effect on how participants perceived the relationship, $F(2, 59.77) = 13.43$, $p < .001$. However, we also found there was a significant interaction effect of Behavior and Information Transfer, $F(4, 59.46) = 2.56$, $p = .048$. The effect of Behavior differed depending on Information Transfer: participants perceived the mobile robot to have a better relationship with the *social* stationary robot than with the *representative* stationary robot in the *explicit*, $d = 1.377$, $p = .031$, and *silent*, $d = 2.01$, $p < .001$, Information Transfer conditions. Participants perceived the robots in the *social* condition to have a better relationship than the *direct* condition when the Transfer condition was *silent*, $d = 1.51$, $p = .013$. This interaction effect showed that the difference in robot Behavior was mainly in the *silent* condition.

For the qualitative responses about the robots' relationship, we annotated answers to the open-ended questions: "How would you describe the relationship between the receptionist robot and the guide robot?" and "What did you like and/or dislike about the interaction with the robots?" for information about the *type* and *nature* of the perceived relationship between the robots. Two of the authors inspected responses and identified 3 categories of relationship type: *equal*, *unequal*, and *no relationship*. Within the *equal* relationship type, there were 2 categories of relationship *nature*: *prescribed*, e.g., commanded by the programmer to act as equals; and *independent*, e.g., friends or coworkers. Within the *unequal* type, there were 3 categories of relationship *nature*: *positive*, e.g., teacher and student; *neutral*, e.g., boss and employee; and *negative*, e.g., master and slave. Some participants did not address relationships in their answers, and in this case, a code of "N/A" was assigned.

Two coders coded 25% of the data to calculate inter-rater reliability. For relationship type, Cohen's κ was .84, and for relationship nature, Cohen's κ was .79. One coder coded the rest of the data. The three Behavior conditions were analyzed individually within each Transfer condition. In the *explicit* Transfer condition, a Fisher's Exact test revealed an association between Behavior conditions and relationship *nature*, $p = .033$. The *representative* and *direct* conditions were more likely to merit perceptions of *unequal negative* and *unequal neutral* relationships than the *social* condition, but pairwise comparisons using a Bonferroni corrected α of .0166 did not reveal significant differences.

E. User Preference

At the end of the study, participants chose which robot they preferred and matched certain descriptive words to one of the

Fig. 4. Example F-formation. Left to right: (1) Participant faces the stationary robot, (2) Mobile robot arrives, (3) Stationary robot turns to talk to the mobile robot, (4) Participant shifts her orientation to ensure a shared space.

three mobile robots. When asked which mobile robot they preferred to lead them to their destination, participants generally preferred the robot from the *social* condition (24/36), followed by the *representative* condition (7/36) and finally the *direct* condition (5/36). A Fisher's Exact test found a significant association between choice of robots and Information Transfer condition, $p = 0.008$. Individual pairwise comparisons with a Bonferroni-corrected α of .0166 revealed a significant difference in robot preference between the *silent* and *reciting* conditions, $p = .005$: participants in the *silent* condition were more likely to prefer the mobile robot in the *social* condition. When participants reported which robot they preferred the least, 20 mentioned the robot in the *representative* condition, 12 picked the robot in the *direct* condition, and 4 chose the robot in the *social* condition. Fisher's Exact tests on all other forced-choice questions including least preferred showed no significant differences across conditions.

In Q3, we also asked participants to rate the competence of each mobile robot on a 7-point scale. We stacked the responses and analyzed them using the same multi-level linear model but without mistakes as a covariate. We found a main effect of Information Transfer on perceived competence of the mobile robot, $F(2, 87.92) = 5.26$, $p = .007$. Pairwise comparisons found that participants rated the mobile robot as more competent in the *reciting* $(M = 6.28, SE = 0.26)$ and *explicit* $(M = 6.25, SE = 0.23)$ conditions than in the *silent* condition $(M = 5.32, SE = 0.23)$, $p = .020$ and $p = .016$.

F. Other Findings

1) Human Position: We noticed shifts in participants' position and posture when the robots verbally interacted (Figure 4). The change in the human's position followed Kendon's F-formation concept [35] such that they oriented to create a shared space between them and the robots.

2) What Information Was Transferred: In the final survey, we also asked whether participants believed the mobile robot knew certain information. Nearly all the participants believed the robot knew their desired walking speed (34/36), their destination (34/36), and that they had requested help navigating (33/36). However, only 17 out of the 36 reported believing that the robot knew their name. In the open-ended questionnaire responses and interviews, multiple participants mentioned that since the mobile robot could not speak, there was no reason for it to know their name and/or confirm knowing their name.

3) How Information Was Transferred: Participants also explained how they believed the information was transferred between robots. We grouped responses by similarity and found that participants believed that information was transferred between the robots in several ways. A few participants thought that the robots communicated out loud, and that the mobile robot interpreted the stationary robot's speech and beeped back in response. For example, P301 thought that information was transferred when *"the guide robot beeped loudly... signaling to the receptionist that it was ready to begin"*. Other participants believed (accurately) that the robots were exchanging information through wireless signals. Several participants thought that communication was occurring via a combination of verbal and electronic signals. The conversation between the participant and the receptionist robot in fact depended on real-time events (though it was heavily structured), while the interaction between the receptionist robot and the guide robot was entirely predetermined. Some participants picked up on this and said that information was transferred via a combination of verbal and electronic signals: P217 said that the guide robot relied on *"wireless data transfer, speech recognition"*, P305 said *"I think the beep signal from the guide robots were some sort of signal otherwise I think through internal communication"*, and P314 thought that the information transfer happened when utterances were *"listened to and compared to a list of things the guide robot is programmed to know"*. Three participants suspected that the entire interaction had been pre-programmed (e.g., *"It could be programmed to just seem like the information was transmitted"*, P310).

V. Discussion

We found evidence partially supporting **H1** whereby Stationary Robot Behavior changed participants' overall perceptions of the stationary robot. Participants felt that the stationary robot displayed more warmth in the *social* condition than in other conditions. The stationary robot was also more likable in the *social* condition than in the *representative* condition. Participants also rated the stationary robot as meaner in the *direct* condition than the *social* condition. While we found the stationary robot to be more social and likable, the stationary robot's Behavior did not change the perceived competence of the robot, and likability of the robot was only significantly different between the *social* and *representative* conditions (and not the *direct* condition). In this study, we intended to evoke social perceptions about a nonsocial robot by changing the way it interacted with a social robot. Collectively, our results suggest that this was achieved.

H2 was partially supported as Behavior affected perceived socialness and likability of the mobile robot, but we were unable to find strong evidence that it influenced perceptions of the robot's competence and trust of the robot. We believe our inability to find an effect for both competence and trust is attributable to a ceiling effect: participants reported high confidence in the mobile robot. They reported an average trust of 6.0 $(SD = 0.98)$ on a 7-point Likert scale and rated competence at 6.6 $(SD = 1.69)$ on a 9-point scale. Their experience of successful guidance in the practice run might have also biased towards a belief that the mobile robot was capable of completing the task. There was a trend in which the mobile robot in the *social* condition was considered more competent than in the *representative* condition.

We found some support for **H3**. The type of Information Transfer changed the effect of Stationary Robot Behavior on the perceived relationship between the robots. Ratings of stationary robot competence trended higher in the *reciting* condition than in the *silent* condition, but this is not sufficient evidence to support **H3(a)**. Participants also felt that the stationary robot was less mean and more likable in the *reciting* condition than in the *silent* condition. There was also a trend where participants reported being more wary of the mobile robot in the *silent* than the *reciting* conditions. Together, these results support **H3(b)** and prior research [24] stating that people do not like covert and silent communication between robots. While we found an effect of *explicit* Information Transfer on the perceived robot relationship nature, how it affected the exact nature of the relationship was inconclusive.

While **H4** was supported in that participants felt the robots in the *social* condition had the best relationship, the effect of Stationary Robot Behavior on perceived relationship was affected by the form of the information transfer. The difference between the *social* and other Behavior conditions was more evident in the *silent* condition. We hypothesize that the absence of similar effects in the other Transfer conditions may be due to participants inferring a long-term relationship between the robots after seeing them hold a longer interaction. The open-ended responses also suggest that more participants perceived an equal relationship in the *social* Behavior condition.

When asked to choose a preferred mobile robot, most participants picked the robot in the *social* condition and found the mobile robot in the *representative* condition to be the least preferable. This provides support for **H5**.

Our work has design implications for facilitating transitions between robots. Users may appreciate acknowledgments that certain information has passed between robots. Our results suggest that whether the robots recite the information aloud or simply acknowledge that a transfer has occurred, some confirmation is better than no confirmation in terms of establishing perceptions of robot competence.

Our results also show that even though robots do not need to treat each other socially to function, there are benefits to having two robots socially interact. Users may perceive the robots as more likable and social if they observe a social interaction between the two robots. Multiple participants wished that they had seen the robots treat each other the way humans do—that is, even more cordially—but appreciated that the interaction was social. We found a preference for socialness, even for functional robots that lack social capabilities. However, a few participants also commented that the social interaction was unnecessarily long and they might prefer a streamlined interaction: P210[*explicit* + *social*] commented, "*The dialogue between the robots went on a bit longer. It didn't bother me, but I could see people get irritated with it.*" We believe the interaction can be optimized for brevity and socialness. Future work should explore this relationship.

VI. Limitations

In multiple sessions, the system had difficulty parsing the participant's speech and required them to repeat their request. Often, participants simply raised their voice in response to a delay, and that solved the issue. However, in a few cases, the experimenter had to step in and ask the participant to speaker louder or rephrase their answers. While we controlled for the potential effect of these mistakes through the inclusion of *system mistakes* as a covariate, it is likely that it still added noise to our study. Future work can explore how failures of one robot influence participants' trust and confidence in another robot. In designing the experiment, we knew that technical problems might emerge, and we considered having a human-in-the-loop or using a Wizard-of-Oz design to avoid such issues. However, we believe it is crucial for HRI studies to utilize real, operational technologies and systems: they better reflect how HRI practitioners may use findings in the wild, and they have potential to expose findings and insights that cannot be captured by other methods [36]. Lastly, we did not include gender as a variable in our analysis because we had fewer male participants. We do not believe that gender influenced the results, but future work should test for gender effects.

Although we counterbalanced the order of the within-subjects conditions and attempted to control for novelty effects and comfort level with a practice run, there was still potential for learning effects during the study. Some participants may have picked up on the phrasing that worked well with the natural language pipeline. We used this design in spite of inherent learning effects because it allowed us to account for individual differences in people's perceptions of social robots.

VII. Future Work and Conclusion

Our results pertain to a heterogeneous robot pair (social and stationary vs. functional and mobile) with an obvious size difference. Future work should explore how the perceived socialness of the robots changes if both robots are human-like (e.g., two Peppers) or otherwise similar to each other. Whether the influence of the social interaction will be overshadowed by robots' nonsocial features remains to be seen.

The emergence of F-formations when the stationary robot turned towards the mobile robot was an unexpected finding. Future work should not only explore how to detect F-formations in these settings but also investigate the possibility of influencing a human's orientation and position by changing multiple robots' orientations and positions (e.g., [37]–[39]).

Our work demonstrated that transitioning a user from one robot to another is not a trivial design problem, but an important aspect of a smooth interaction. Through the expression of information exchange and designing social interactions between robots, we can instill confidence in the robots and change how users perceive their abilities. Having robots treat each other in ways that are consistent with social expectations was the most preferred form of robot-robot interaction. However, it is clear there are subtleties that require careful designs.

REFERENCES

[1] "The KLM airport robot Care-E wants you to take the load off," https://www.gearbrain.com/airport-robots-2606831725.html, accessed: 2018-10-01.

[2] A. M. Sabelli and T. Kanda, "Robovie as a mascot: a qualitative study for long-term presence of robots in a shopping mall," *International Journal of Social Robotics*, vol. 8, no. 2, pp. 211–221, 2016.

[3] "The Relay hotel delivery robot will soon spot Wi-Fi dead zones and mingle with guests," https://www.theverge.com/2018/1/11/16879432/savioke-relay-hotel-delivery-robot-wi-fi-dead-zones-mingle-ces-2018, accessed: 2018-10-01.

[4] I. R. Nourbakhsh, C. Kunz, and T. Willeke, "The mobot museum robot installations: A five year experiment," in *Intelligent Robots and Systems, 2003.(IROS 2003). Proceedings. 2003 IEEE/RSJ International Conference on*, vol. 4. IEEE, 2003, pp. 3636–3641.

[5] D. Karreman, G. Ludden, and V. Evers, "Visiting cultural heritage with a tour guide robot: A user evaluation study in-the-wild," in *International Conference on Social Robotics*. Springer, 2015, pp. 317–326.

[6] "The Smithsonian is using a swahili-speaking robot to break down language barriers," https://www.smithsonianmag.com/smithsonian-institution/pepper-swahili-speaking-robot-breaks-down-language-barriers-african-art-exhibit-180968902/, accessed: 2018-09-30.

[7] S. Reig, J. Forlizzi, and A. Steinfeld, "Leveraging robot embodiment to facilitate trust and smoothness," in *Companion of the 2019 ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 2019.

[8] R. Kirby, J. Forlizzi, and R. Simmons, "Affective social robots," *Robotics and Autonomous Systems*, vol. 58, no. 3, pp. 322–332, 2010.

[9] I. Leite, C. Martinho, and A. Paiva, "Social robots for long-term interaction: a survey," *International Journal of Social Robotics*, vol. 5, no. 2, pp. 291–308, 2013.

[10] T. Matsumoto, M. Goto, R. Ishii, T. Watanabe, T. Yamada, and M. Imai, "Where should robots talk?: Spatial arrangement study from a participant workload perspective," in *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 2018, pp. 270–278.

[11] M. R. Fraune, S. Sherrin, S. Sabanović, and E. R. Smith, "Rabble of robots effects: Number and type of robots modulates attitudes, emotions, and stereotypes," in *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 2015, pp. 109–116.

[12] W.-L. Chang, J. P. White, J. Park, A. Holm, and S. Šabanović, "The effect of group size on people's attitudes and cooperative behaviors toward robots in interactive gameplay," in *RO-MAN, 2012 IEEE*. IEEE, 2012, pp. 845–850.

[13] N. Salomons, M. van der Linden, S. Strohkorb Sebo, and B. Scassellati, "Humans conform to robots: Disambiguating trust, truth, and conformity," in *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 2018, pp. 187–195.

[14] I. Leite, M. McCoy, M. Lohani, D. Ullman, N. Salomons, C. Stokes, S. Rivers, and B. Scassellati, "Emotional storytelling in the classroom: Individual versus group interaction between children and robots," in *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 2015, pp. 75–82.

[15] M. Shiomi, T. Kanda, D. F. Glas, S. Satake, H. Ishiguro, and N. Hagita, "Field trial of networked social robots in a shopping mall," in *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*. St. Louis, MO, USA: IEEE, 2009, pp. 2846–2853. [Online]. Available: http://ieeexplore.ieee.org/document/5354242/

[16] M. R. Fraune, Y. Nishiwaki, S. Sabanović, E. R. Smith, and M. Okada, "Threatening flocks and mindful snowflakes: How group entitativity affects perceptions of robots," in *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 2017, pp. 205–213.

[17] T. Kanda, H. Ishiguro, T. Ono, M. Imai, and K. Mase, "Multi-robot cooperation for human-robot communication," in *Robot and Human Interactive Communication, 2002. Proceedings. 11th IEEE International Workshop on*. IEEE, 2002, pp. 271–276.

[18] K. Hayashi, D. Sakamoto, T. Kanda, M. Shiomi, S. Koizumi, H. Ishiguro, T. Ogasawara, and N. Hagita, "Humanoid robots as a passive-social medium-a field experiment at a train station," in *Human-Robot Interaction (HRI), 2007 2nd ACM/IEEE International Conference on*. IEEE, 2007, pp. 137–144.

[19] K. Hayashi, T. Kanda, T. Miyashita, H. Ishiguro, and N. Hagita, "Robot manzai: Robot conversation as a passive–social medium," *International Journal of Humanoid Robotics*, vol. 5, no. 01, pp. 67–86, 2008.

[20] T. Iio, Y. Yoshikawa, and H. Ishiguro, "Pre-scheduled turn-taking between robots to make conversation coherent," in *Proceedings of the Fourth International Conference on Human Agent Interaction*. ACM, 2016, pp. 19–25.

[21] M. Vázquez, A. Steinfeld, S. E. Hudson, and J. Forlizzi, "Spatial and other social engagement cues in a child-robot interaction: Effects of a sidekick," in *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*. ACM, 2014, pp. 391–398.

[22] M. Luria, "Designing robot personality based on fictional sidekick characters," in *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 2018, pp. 307–308.

[23] M. R. Fraune and S. Sabanovic, "Negative attitudes toward minimalistic robots with intragroup communication styles." in *RO-MAN*, 2014, pp. 1116–1121.

[24] T. Williams, P. Briggs, and M. Scheutz, "Covert robot-robot communication: Human perceptions and implications for human-robot interaction," *Journal of Human-Robot Interaction*, vol. 4, no. 2, pp. 24–49, 2015.

[25] "Hello Baxter SDK Wiki," http://sdk.rethinkrobotics.com/wiki/, accessed: 2018-09-30.

[26] "Amazon AWS Polly SDK," https://aws.amazon.com/polly/developers/, accessed: 2018-09-30.

[27] "SNIPS Natural Language Understanding Engine," https://github.com/snipsco/snips-nlu, accessed: 2018-09-30.

[28] E. Marder-Eppstein, E. Berger, T. Foote, B. Gerkey, and K. Konolige, "The office marathon: Robust navigation in an indoor office environment," in *Robotics and Automation (ICRA), 2010 IEEE International Conference on*. IEEE, 2010, pp. 300–307.

[29] R. Toris, J. Kammerl, D. V. Lu, J. Lee, O. C. Jenkins, S. Osentoski, M. Wills, and S. Chernova, "Robot web tools: Efficient messaging for cloud robotics." in *IROS*, 2015, pp. 4530–4537.

[30] C. M. Carpinella, A. B. Wyman, M. A. Perez, and S. J. Stroessner, "The robotic social attributes scale (rosas): Development and validation," in *Proceedings of the 2017 ACM/IEEE International Conference on human-robot interaction*. ACM, 2017, pp. 254–262.

[31] 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.

[32] J.-Y. Jian, A. M. Bisantz, and C. G. Drury, "Foundations for an empirically determined scale of trust in automated systems," *International Journal of Cognitive Ergonomics*, vol. 4, no. 1, pp. 53–71, 2000.

[33] H. Patterson, "Maximum likelihood estimation of components of variance," in *Proceeding Eight International Biometric Conference, 1975*. Biometric Soc., 1975.

[34] W. W. Stroup, *Generalized linear mixed models: modern concepts, methods and applications*. CRC Press, 2016.

[35] A. Kendon, *Conducting interaction: Patterns of behavior in focused encounters*. CUP Archive, 1990, vol. 7.

[36] A. Steinfeld, O. C. Jenkins, and B. Scassellati, "The Oz of Wizard: Simulating the human for interaction research," in *Proceedings of the 4th ACM/IEEE International Conference on Human Robot Interaction*, ser. HRI '09. New York, NY, USA: ACM, 2009, pp. 101–108.

[37] M. Vázquez, E. J. Carter, B. McDorman, J. Forlizzi, A. Steinfeld, and S. E. Hudson, "Towards robot autonomy in group conversations: Understanding the effects of body orientation and gaze," in *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, ser. HRI '17. New York, NY, USA: ACM, 2017, pp. 42–52.

[38] C. Shi, M. Shimada, T. Kanda, H. Ishiguro, and N. Hagita, "Spatial formation model for initiating conversation," in *Proceedings of Robotics: Science and Systems*, Los Angeles, CA, USA, June 2011.

[39] C. Shi, M. Shiomi, T. Kanda, H. Ishiguro, and N. Hagita, "Measuring communication participation to initiate conversation in human–robot interaction," *International Journal of Social Robotics*, vol. 7, no. 5, pp. 889–910, 2015.