## Title
Mind Perception at Play: Exploring Agent and Action Dynamics in Real-Time Human-Robot Interaction



## Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 46(0)

## Authors
Pekçetin, Tuƒüçe Nur
Evsen, ≈ûeyda
Pekçetin, Serkan
et al.

## Publication Date
2024

Peer reviewed

# Mind Perception at Play: Exploring Agent and Action Dynamics in Real-Time Human-Robot Interaction

**Tuğçe Nur Pekçetin (tugce.bozkurt@metu.edu.tr)**
Department of Cognitive Science, Graduate School of Informatics, Middle East Technical University, Ankara, Turkey

**Şeyda Evsen (seyda.evsen@bilkent.edu.tr)**
Department of Neuroscience, Bilkent University, Ankara, Turkey

**Serkan Pekçetin (serkan.pekcetin@metu.edu.tr)**
Department of Computer Engineering, Middle East Technical University, Ankara, Turkey

**Cengiz Acarturk (cengiz.acarturk@uj.edu.pl)**
Department of Cognitive Science, Jagiellonian University, Krakow, Poland

**Burcu A. Urgen (burcu.urgen@bilkent.edu.tr)**
Department of Psychology, Department of Neuroscience, Bilkent University, UMRAM, Ankara, Turkey

## Abstract

The study of mind perception, particularly how one perceives the mental states of 'others,' has attracted considerable interest in cognitive science. The present study contributes to the investigation of mind perception in a human-robot interaction context, by testing a humanoid robot and a human and their communicative and noncommunicative actions. We examine mind perception across its two primary dimensions: Agency and Experience and in their High and Low ends. The novelty of our study lies in its real-time and implicit nature—both identified as crucial elements in current debates within the field. Our results indicate that testing physically present and active agents, as well as exposing participants to various types of live actions, influences mental capacity attributions across different capacities. Additionally, the integration of behavioral measurements alongside verbal data holds promise for a detailed interpretation of the mind perception process.

**Keywords:** mind perception; agency; experience; human-robot interaction; communicative actions; noncommunicative actions

## Introduction

A closer examination of fundamental debates in Cognitive Science, including the Chinese Room argument (Searle, 1980), Turing Test (Turing, 1950), and classical Theory of Mind discussions (Gordon, 1986), reveals a recurring theme: the persistent curiosity about whether an agent possesses a mind or cognitive capacities. This curiosity remains relevant today, especially with the growing interactions with "minds" such as chatbots or social robots in our daily lives, intensifying the significance of such inquiries to better understand human social reasoning (Broadbent, 2017). The human ability to attribute mental capacities to nonhuman entities (Epley, Waytz, et al., 2010) emphasizes the complexity of this process, necessitating thorough investigations into the causes and consequences of mind perception (Waytz, Gray, Epley, & Wegner, 2010). Artificial agents, especially social robots, with human-like features such as anthropomorphic appearances or movement capabilities are promising candidates for further research in this domain (Henschel, Laban, & Cross, 2021). Understanding whether and why we attribute mental capacities such as beliefs, desires, or intentions to robots is crucial, as humans tend to trust, collaborate with, and accept more readily those agents they perceive as having mental capacities (Epley, Waytz, & Cacioppo, 2007; Waytz et al., 2010). Besides having the potential to shape the future of human-robot interaction (HRI), these investigations also contribute to documenting the evolution of our relationships with these "extraordinary entities," shedding light on the broader aspects of human social cognition (Weisman, 2022).

Investigations into mind perception within the HRI context commonly employ comparative analyses between humans and robots as agents. While most previous research indicates a stronger tendency to attribute mental states to humans over robots, discrepancies exist since some studies note comparable tendencies between the two agents (Thellman, de Graaf, & Ziemke, 2022). However, the ecological and external validity of these studies is often questioned due to their reliance on verbal scenarios, images, or videos of agents. Evidence also suggests that anthropomorphism towards robots increases when they are physically present (Kiesler, Powers, Fussell, & Torrey, 2008; Straub, 2016; J. Li, 2015). Additionally, recent studies also indicate that besides factors such as perceived similarity (Waytz et al., 2010) and appearance or the degree of anthropomorphism (K. Gray & Wegner, 2012), behaviors and capabilities of the robots also significantly influence attributions of mental capacities. Therefore, testing physically present and actively performing robots in diverse actions could provide deeper insights into the dynamics of mind perception in HRI.

The inconsistencies observed in previous studies may stem from the reliance on self-report methods such as questionnaires and surveys. These explicit methodologies have notable limitations, including their inability to capture the underlying processes of mind perception and the potential to misrepresent actual thoughts due to factors like social desirability bias (Fazio & Olson, 2003; Nosek, Hawkins, & Frazier, 2011). Consequently, there is a pressing need to incorporate more implicit methodologies in this field (Greenwald & Banaji, 1995).



In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.* 

In this work, we build upon the existing questions and investigate how agent and action type influence mind perception processes in a real-time setting, employing an implicit testing methodology. Specifically, we: (1) compare two types of agents—a human and a robot; (2) make participants evaluate these agents as they perform both communicative and noncommunicative actions; (3) utilize an implicit task that we developed, which allows us to collect verbal and behavioral data through a real-time yet controlled methodology.

## Relevant Work

Acknowledging the complex nature of mind perception as a cognitive process, certain abstractions are necessary to reveal the fundamental conceptual structure. A seminal work by Gray et al. (2007) has significantly contributed to this endeavor by proposing that people's judgments about others' minds revolve around two distinct dimensions: *Agency* and *Experience*. In broad terms, Agency refers to "the capacity to do" while Experience refers to "the capacity to feel". When evaluating Agency, individuals assess the others in terms of possessing a mind capable of self-directed, intentional actions, or making choices. On the other hand, evaluations of Experience focus on having the capacity to have subjective mental states, feelings, emotions, or other experiential aspects. Further research has broadened these concepts, introducing varied agent types and dimensional structures. For instance, Weisman et al. (2017) categorized capacities into Body, Mind, and Heart, and Malle (2019) proposed divisions into Affect, Moral and Mental Regulation, and Reality Interaction. A recent validation study (Pekçetin, Barinal, Tunç, Acarturk, & Urgen, 2023) confirmed the original division into Agency and Experience while identifying additional clusters within these dimensions. These clusters particularly relate to concepts requiring a higher level of processing or those exhibiting physical or instinctual attributes.

Previous studies (H. M. Gray et al., 2007; K. Gray & Wegner, 2012; Ishii & Watanabe, 2019; Okanda, Taniguchi, & Itakura, 2019) have reported differential mental state attributions towards robots and humans, particularly noting that humans are more likely to be attributed with experience-related capacities, whereas agency-related capacities are similarly attributed to both. Additionally, previous studies focusing on how different behaviors of agents affect attributions of mental capacities to them reported stronger tendencies in attributing mental capacities to robots when they exhibit emotional (Złotowski, Strasser, & Bartneck, 2014) and social behaviors (Fraune et al., 2020; Straub, 2016), or gestures (Salem, Eyssel, Rohlfing, Kopp, & Joublin, 2011). Furthermore, some studies found that different types of actions impact various aspects of mind perception. For instance, Saltik et al. (2021) reported that the type of action performed by a robot increased the tendency to attribute higher Agency ratings, though they did not observe this effect in the Experience dimension. Złotowski et al. (2014) reported that exhibiting emotional behavior resulted in higher ratings in the Experience dimension, while exhibiting intelligent behavior resulted in higher Agency scores. These insights; however, face generalizability challenges to real-world scenarios, highlighting the need for methodologies that better capture interactive behaviors occurring in naturalistic settings.

Research on mind perception in HRI has largely utilized explicit methods like questionnaires and interviews, which, while useful in emerging or applied domains, also have limitations. Explicit measures often rely on subjective judgments that can be influenced by external factors, making them less reliable (Fazio & Olson, 2003; Nosek et al., 2011). To gain a deeper mechanistic insight into human cognition and behavior, the integration of implicit measurements is crucial (Greenwald & Banaji, 1995). Previous studies have begun to incorporate non-verbal behavioral metrics such as response times (Z. Li, Terfurth, Woller, & Wiese, 2022), anticipatory gaze (Sciutti et al., 2013; Thellman & Ziemke, 2020), and attentional cueing (Wiese, Mandell, Shaw, & Smith, 2019), revealing that these methods tend to elicit stronger attributions of mental states to robots compared to verbal assessments (Thellman et al., 2022).

## The Present Study

Building upon open questions and methodological challenges in the current state of the art, we adopt an empirical approach to explore the mind perception process in HRI, aiming to address the following key research questions: *RQ1: Do people attribute mental states to humans more than to robots?* and *RQ2: Are mental states attributed more frequently in response to communicative actions than to noncommunicative actions?* To investigate these questions, we tested a human actor and a humanoid robot, Pepper, whose extensive mobility capabilities enabled us to test a wide variety of action stimuli. We tested the agents and actions on two dimensions of mind perception—Agency and Experience—using the Real-World Implicit Association Test (RW-IAT) we developed (Pekçetin, Evsen, Pekçetin, Acarturk, & Urgen, 2024). The study design has been established upon a real-time yet controlled approach, enabling us to collect behavioral metrics besides the final verbal responses. As behavioral data, we recorded the response times and mouse trajectories of the participants while they were making their binary evaluations. Response time is the classical measure collected in classical Implicit Association Test (IAT) (Greenwald, McGhee, & Schwartz, 1998), as an indicator of the overall speed, thus the strength of the association; however, it falls short in demonstrating what happens during an IAT trial (Yu, Wang, Wang, & Bastin, 2012). So, we included mouse tracking to our RW-IAT to capture the motor trajectories which reveal the real-time course characteristics of the entire response process (Freeman & Ambady, 2010). This approach allows us to further explore: *RQ3: How are different dimensions of mind perception influenced by variations in agent and action types?* and *RQ4: Do verbal responses from participants align with behavioral metrics such as response time and mouse trajectories?*

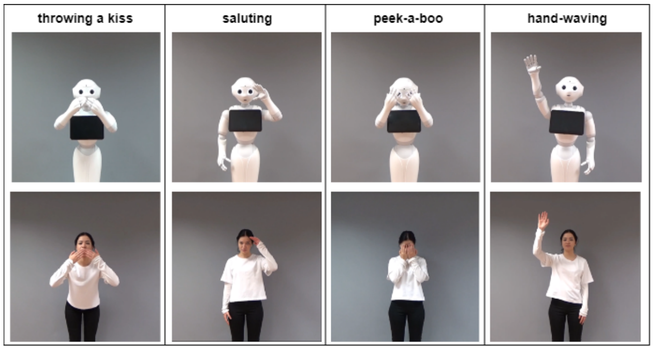


Figure 1: The communicative actions across human and robot conditions. The images are screen captures from the six-second videos filmed for the action norming study.

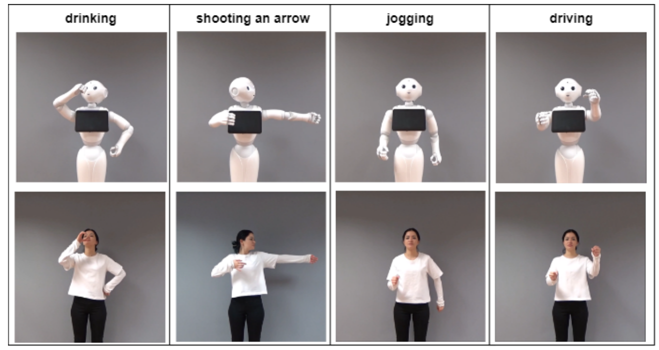


Figure 2: The noncommunicative actions across human and robot conditions. The images are screen captures from the six-second videos filmed for the action norming study.

## Methodology

The study was conducted in person in a psychology laboratory at Bilkent University. The experimental materials and protocols were approved by the Ethics Committee for Research with Human Participants at the same university. Participants were 18 or above and they provided informed consent before the study. The study took between 70 to 100 minutes. At the end of the study, the participants received compensation of 50 Turkish Liras (approximately 3 USD then).

### Participants

A total of 166 participants enrolled in the study, but data from 6 were excluded due to technical problems encountered during data collection. Specifically, sudden IP configuration changes disrupted the robot-computer connection, leading to incomplete sessions for two participants. Additionally, data from four participants were excluded due to repeated blocks caused by experiment code crashes linked to Wi-Fi connectivity losses. In the end, the data of 160 participants, comprising 98 females and 62 males, with age range 18-73 ($M$ = 44.93, $SD$ = 16.63) were included in the analyses. For participants aged 63 and above ($n$ = 40), a mini-mental state examination (Folstein, Folstein, & McHugh, 1975) was administered by a neurology specialist to assess cognitive impairments. Only the participants scoring above the cut-off score of 23-24 out of 30 for normal cognition (Thorndike, 1953), and those approved by the specialist, were recruited for the study. Within this age group, the scores ranged from a low of 27 ($n$ = 1) to a high of 30 ($n$ = 24), with a mean score of 29.43.

### Materials

Although the original IAT (Greenwald et al., 1998) was adapted to examine mind perception using images of humans and robots (Z. Li et al., 2022), it was not feasible to present the real-time actions of two actors randomly and without any delay in a single block. Inspired by the Single Category IAT (Karpinski & Steinman, 2006), we developed a tailored IAT, RW-IAT (Pekçetin, Evsen, Pekçetin, Acarturk, & Urgen, 2024) for both human and robot actors across four blocks, addressing their distinct categories. The details regarding the development of the task and lab setup are documented in our prior work (Pekçetin, Evsen, Pekçetin, Acarturk, & Urgen, 2023; Pekçetin, Evsen, Pekçetin, Karaduman, et al., 2024). RW-IAT included target and attribute concepts, similar to the structure of the classical IAT (Greenwald et al., 1998) and its modified versions (Karpinski & Steinman, 2006).

**Attribute Stimuli** We utilized Agency and Experience as attribute stimuli, and included *High Agency*, *Low Agency*, *High Experience*, and *Low Experience* to represent the extremes of both dimensions, based on a recent study (Z. Li et al., 2022). Following best IAT practices (Greenwald et al., 2022), we conducted a Lexical Training session to familiarize participants with these terms. We defined Agency as “the ability to do” and Experience as “the ability to feel,” with High and Low ends indicating varying levels of capacity within each dimension. We validated these concepts, definitions, and their sub-concepts as examples of the broader concepts in an online study (Pekçetin, Barinal, et al., 2023).

**Target Stimuli** While creating the action stimuli, we benefited from the descriptions in datasets featuring point-light actions (Manera, Schouten, Becchio, Bara, & Verfaillie, 2010; Zaini, Fawcett, White, & Newman, 2013). We selected *communicative*—conveying or exchanging information with the recipient, e.g., hand-waving and *noncommunicative* actions—object-oriented actions related to activities, not a recipient, e.g. jogging (Ekman & Friesen, 1969; McNeill, 1985). We animated these actions on the robot actor using Android Studio and Pepper SDK’s Animation Editor IDE. We standardized and filmed 40 actions for both human and robot actors, then validated them through two online studies (Pekçetin, Aşkın, et al., 2023) with 438 participants who identified each action, categorized them as communicative or noncommunicative, and rated their confidence levels. We analyzed the data based on $H$ entropy (Shannon, 1948), communicativeness score, and confidence levels using a $k$-means clustering algorithm (Hartigan & Wong, 1979; Thorndike, 1953). Figures 1 and 2 illustrate the most agreed upon and selected communicative and noncommunicative actions, respectively: *throwing a kiss*, *saluting*, *peek-a-boo*, and *hand-waving*, *drinking*, *shooting an arrow*, *jogging*, and *driving*. The action videos are online at `https://osf.io/e6mts/`. An Action Identification session was conducted to familiarize participants with these actions, serving as both a manipulation check and preparation for the implicit task.

## Procedure

After the participants provided demographic information and consent in a separate room, they were welcomed into the main experiment room, which was divided by curtains into two sections: the participant area and the actor area. Here, participants were positioned in front of a 55" OLED screen until the experiment concluded. At the end of the session, they were introduced to the robot and shown the backstage. Before starting the implicit task, participants were briefed on the real-time nature of the study, involving both human and robot actors behind the curtains. They then completed the Lexical Training and Action Identification phases before proceeding with the implicit task.

The task consists of four blocks, with two dedicated to Agency and two to Experience, featuring both human and robot actors. The procedure is as follows: Participants first receive instructions about the dimension of the upcoming block. A fixation cross then appears on the opaque OLED screen, signaling the upcoming presentation of the stimulus (see Figure 3.1). The screen subsequently switches to its see-through mode to display live-action stimuli (see Figure 3.2). Participants observe each action for six seconds, with all actions starting and ending in a standing position to maintain consistency. After the action, the screen reverts to opaque, and the response screen appears (see Figure 3.3), where participants' responses, response times, and mouse trajectories are recorded. Participants are encouraged to respond quickly but have up to 30 seconds to evaluate each stimulus. Following their response, a fixation screen appears, signaling the next action. This sequence repeats for nine trials per block. Each block starts with a neutral (standing) action and continues with randomly presented communicative and noncommunicative actions. The robot actor is operated using a Wizard-of-Oz setup, and the human actor uses a backstage laptop to track the order of the actions. The order of the blocks is counterbalanced among participants. After each block, except the last, participants see a "Please wait!" message, and neutral background music plays while the human actor prepares the stage for the next set of actions. The actor signals readiness for the next block by waving through a security camera, prompting the experiment conductor to proceed.

## Independent and Dependent Variables

The study employs two independent variables: *Actor type*: Robot or Human and *Action type*: Neutral, Communicative, and Noncommunicative. During the analyses of the Action Identification data, we observed that over 20% of the trials labeled as 'neutral' actions were ambiguously identified (e.g., "waiting for someone" or "looking at someone"), complicating their classification as truly neutral. Consequently, our analysis now focuses exclusively on the Communicative and Noncommunicative actions, as originally intended.

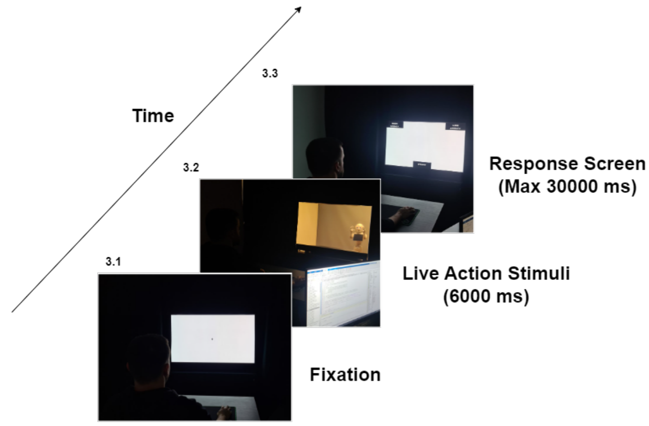


Figure 3: The sequence of a sample trial from Robot-Agency block. The task includes four blocks and there are nine trials in each block.

The primary continuous dependent variable is *Response Time (RT)*, which measures the seconds elapsed from the end of an action to a mouse click on a response option. We also measured *Maximum Deviation (MD)* and *Area Under the Curve (AUC)* to evaluate response trajectories. MD measures the largest perpendicular deviation from an idealized straight-line trajectory, indicating the maximum attraction toward the unselected choice. AUC, considered a more comprehensive index, calculates the geometric area between the actual trajectory and the idealized line, encompassing all time steps. We calculated the MD and AUC values, based on the formulas of the original work (Freeman & Ambady, 2010). High values of MD and AUC are interpreted as indicators of participant hesitation (Freeman, Dale, & Farmer, 2011; Yu et al., 2012). Lastly, the categorical dependent variable *Response* is divided into High or Low categories, with the ratio of High responses reflecting greater attributions of mental capacity in a specific dimension.

## Analyses

The initial dataset contained 5760 observations, but after excluding 6 missed and 639 neutral trials, 23 outliers with RTs over 11.83 seconds ($M$ = 2.02, $SD$ = 1.89 for RTs), and one trial with an extreme MD value, we processed a final dataset of 5091 trials and categorized data by Agency and Experience dimensions in a wide format using MATLAB 2023b. Then we used RStudio 2023.09.1 for analysis. Given the Shapiro-Wilk test indicated a significantly non-normal distribution ($p < .001$) for both actor and action types across dimensions, we applied Friedman's test (Friedman, 1937), a nonparametric alternative to repeated-measures ANOVA (Field, Miles, & Field, 2012). Following a significant result from Friedman's test, we conducted post hoc analyses using the Nemenyi test (Nemenyi, 1963) to identify specific group differences.

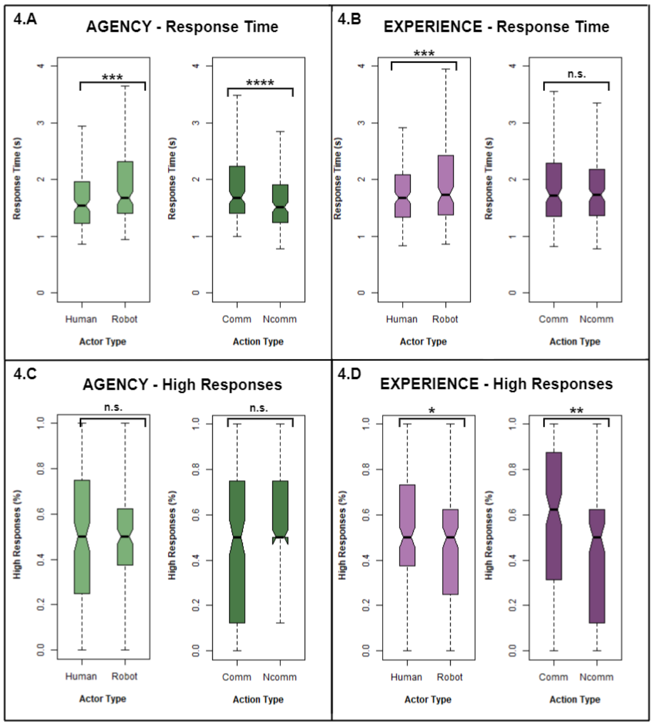


Figure 4: Boxplots of Response time (top row) and Response (bottom row) based on actor and action type. The boxes represent the 25th to 75th percentiles, and the line inside indicates the median.

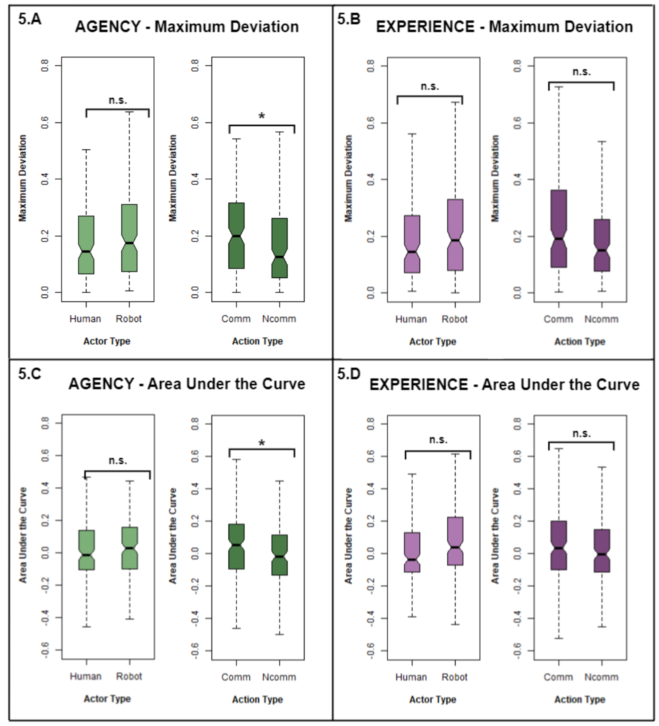


Figure 5: Boxplots of Maximum Deviation (top row) and Area Under the Curve (bottom row) based on actor and action type. The boxes represent the 25th to 75th percentiles, and the line inside indicates the median.

## Results

All participants evaluated each actor and action combination across both Agency and Experience dimensions. Consequently, we have organized the results separately for these two dimensions. In the subsequent section, Figure 4 displays boxplots for Response Time and the Ratio of High Responses, while Figure 5 showcases the boxplots for Maximum Deviation and Area Under the Curve.

### Agency Dimension

**Actor Type** The maximum deviation ($\chi^2(1) = 3.60$ , $p = .058$), the area under the curve values ($\chi^2(1) = 1.60$, $p = .206$) and the ratio of High responses ($\chi^2(1) = 1.33$, $p = .248$) showed no significant changes across actor types. However, the response times significantly varied by actor type, $\chi^2(1) = 11.02$, $p < .001$ (see Actor Type in Figure 4.A). Pairwise comparisons indicated that the response times for the robot actor were significantly higher than for the human actor, $p < .001$.

**Action Type** There was a main effect of action type on the response times, $\chi^2(1) = 16.90$, $p < .0001$ (see Action Type in Figure 4.A); maximum deviations, $\chi^2(1) = 5.62$, $p = .018$ (see Action Type in Figure 5.A); and areas under the curve, $\chi^2(1) = 6.40$, $p = .011$ (see Action Type in Figure 5.C). However, the action type did not significantly affect the ratio of the High responses ($\chi^2(1) = 3.46$, $p = .063$). Pairwise comparisons revealed that the RT, $p < .0001$; MD, $p = .018$; and AUC values, $p = .011$ were significantly longer, higher, and larger, respectively, for communicative actions compared to noncommunicative actions.

### Experience Dimension

**Actor Type** The type of the actor significantly affected the response times, $\chi^2(1) = 11.02$, $p < .001$ (see Actor Type in Figure 4.B) and the ratios of the High responses, $\chi^2(1) = 5.95$, $p = .015$ (see Actor Type in Figure 4.D). Pairwise comparisons revealed that the robot actor elicited significantly longer RT, $p < .001$, and the ratio of High Experience responses for the robot was significantly lower than the ratio for the human actor, $p = .048$. The maximum deviations, ($\chi^2(1) = 3.02$, $p = .082$) and areas under the curve, ($\chi^2(1) = 3.60$, $p = .058$) showed no significant changes across actors.

**Action Type** The response times ($\chi^2(1) = 0$ , $p = 1$), maximum deviations ($\chi^2(1) = 3.60$, $p = .058$), and areas under the curves ($\chi^2(1) = 1.60$, $p = .21$) did not significantly change across action types. However, there was a significant effect of the action type on the ratio of High responses of the participants, $\chi^2(1) = 10.96$, $p < .001$ (see Action Type in Figure 4.D). Pairwise comparisons revealed that the ratio of High Experience responses for the communicative actions was significantly higher compared to the ratio of the High scores given for the noncommunicative actions, $p < .01$.

## Discussion

In the current study, we aimed to examine the interplay between agent and action dynamics in mind perception during real-time human-robot interactions. We tested both a human actor and a humanoid robot, employing a range of systematically selected communicative and noncommunicative actions alongside validated conceptual stimuli. Our novel experimental setup facilitated the simultaneous presentation of live-action stimuli and collection of behavioral data in a controlled environment. Our findings for RQ1, RQ2, and RQ3 indicate that both agent and action types significantly influence mental capacity attributions, with distinct patterns observed in the Agency and Experience dimensions. Regarding RQ4, our results higlight the value of behavioral metrics alongside verbal measures, as they can reveal the duration of the evaluation process and hesitations during this process, offering insights that extend beyond what final verbal responses alone can provide. Additionally, we emphasize the potential impact of robots' physical presence and activity on mental capacity attributions.

Regarding agency dynamics, we observed similar Agency capacity attributions for both human and nonhuman agents, with the human agent receiving higher attributions for Experience capacities. This is consistent with prior research that used explicit measurements, which found comparable Agency attributions between humans and robots but higher Experience attributions for humans (H. M. Gray et al., 2007; K. Gray & Wegner, 2012; Ishii & Watanabe, 2019; Okanda et al., 2019). Our data also showed that longer response times in Agency evaluations were associated with a preference for nonhuman agents, aligning with previous findings (Levin, Killingsworth, Saylor, Gordon, & Kawamura, 2013). For Experience evaluations, despite longer decision times for nonhuman agents, the human agent was still attributed with higher Experience capacities. This difference in Agency and Experience attributions suggests that nonhuman agents' mimicry of human behavior may elevate their Agency scores, narrowing the gap in mental state attributions (Abubshait & Wiese, 2017). The physical presence of the robot in our study might have enhanced this tendency.

We observed the impact of action type on the behavioral metrics within the Agency dimension and on the attribution levels in the Experience dimension. Notably, communicative actions required longer evaluation times and elicited more hesitations regarding Agency. Despite these differences in evaluation time and hesitations, communicative actions resulted in Agency attribution levels similar to those of noncommunicative actions. This result differs from earlier research that tested various action videos implicitly and found that communicative actions received higher Agency ratings (Saltik et al., 2021). We interpret this as potentially due to participants observing all actions in real-time, which may simplify the attribution of Agency scores across both action types. In the Experience dimension, communicative actions led to higher Experience ratings. This finding aligns with a previous study involving a physically present robot, which found that while intelligent behavior did not significantly alter Agency ratings, the Experience ratings were affected by the robot's display of emotionality (Złotowski et al., 2014). This highlights the importance of examining mind perception across different types of behaviors.

### Limitations and Future Work

In our implicit task, live-action stimuli necessitate a six-second duration to accommodate the robot actor's initiation, performance, and return to a standing posture. Despite potential concerns about the implicit nature of the task, this timeframe is the most viable option within current technological constraints. Furthermore, our participant pool represented four generations. Although exploring generational differences in mind perception is beyond the scope of this paper, we anticipate age-related differences in mental capacity attributions in further analyses.

Our study primarily examined the impact of action types on mind perception using a unidirectional setup. However, our flexible experimental design can be adapted to investigate bidirectional interactions in a realistic yet controlled environment. Future directions could enhance realism by incorporating factors such as facial expressions, eye contact, and linguistic cues, and by enabling agents to exchange feedback with participants. Future research could also involve comparing the outcomes of this real-time study with those of an online equivalent to precisely assess the impact of physical presence.

## Concluding Remarks

Since our study is unique in terms of examining action dynamics with physically present non-human agents through implicit measurements, we would like to conclude with several insights. Implicit measurements are crucial for deeply exploring the determinants and implications of mind perception, though their applications in real-life scenarios pose challenges. We encourage adherence to the best experimental practices while customizing the experimental setup and stimuli to minimize confounds, thus ensuring the reliability of measurements. Furthermore, the inclusion of physically present and active robots, or more broadly, the use of naturalistic stimuli, has the potential to enhance the ecological and external validity of the research. Integrating implicit measurements with verbal assessments also provides deeper insights into the underlying processes of mind perception. Beyond experimental considerations, our findings that the nature of robot behaviors influences human perceptions have significant implications for designers and engineers. Specifically, the design of social robots should ensure their capability to interact with humans through various gestures and emotional behaviors, in addition to performing manipulative or locomotive actions. We believe that continued research in mind perception within HRI will both advance robot design and deepen our understanding of perceptual, cognitive, and motor processes during interactions with other minds.

## Acknowledgments

This study was funded by The Scientific and Technological Research Council of Türkiye, TÜBİTAK (Project Number: 120K913). We thank Tuvana Dilan Karaduman for her acting and contributions to the design of the robot actions and the project, Sabiha Tezcan Aydemir for conducting online neurological tests on participants over the age of 63 as a neurology specialist, Gaye Aşkın, Aslı Eroğlu, and Deniz Erdil for their theoretical and practical contributions to this project. We also thank our research assistants Badel Barinal, Jana Tunç, Gizem Çınar, Zeynep Koçak, Duru Naz Han, Melis Akdeniz, and İdil Ece Karaosman for their collaboration and efforts in data collection for the survey part of the study.

## References

Abubshait, A., & Wiese, E. (2017). You look human, but act like a machine: Agent appearance and behavior modulate different aspects of human–robot interaction. *Frontiers in Psychology*, *8*. doi: 10.3389/fpsyg.2017.01393

Broadbent, E. (2017). Interactions with robots: The truths we reveal about ourselves. *Annual Review of Psychology*, *68*, 627–652.

Ekman, P., & Friesen, W. V. (1969). The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica*, *1*(1), 49–98.

Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: a three-factor theory of anthropomorphism. *Psychological Review*, *114*(4), 864.

Epley, N., Waytz, A., et al. (2010). Mind perception. *Handbook of Social Psychology*, *1*(5), 498–541.

Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and use. *Annual Review of Psychology*, *54*(1), 297-327. (PMID: 12172003) doi: 10.1146/annurev.psych.54.101601.145225

Field, A., Miles, J., & Field, Z. (2012). *Discovering Statistics Using R*. Sage Publications.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "mini-mental state": a practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*(3), 189–198.

Fraune, M. R., Oisted, B. C., Sembrowski, C. E., Gates, K. A., Krupp, M. M., & Šabanović, S. (2020). Effects of robot-human versus robot-robot behavior and entitativity on anthropomorphism and willingness to interact. *Computers in Human Behavior*, *105*, 106220. doi: https://doi.org/10.1016/j.chb.2019.106220

Freeman, J. B., & Ambady, N. (2010). Mousetracker: Software for studying real-time mental processing using a computer mouse-tracking method. *Behavior Research Methods*, *42*(1), 226–241.

Freeman, J. B., Dale, R., & Farmer, T. (2011). Hand in motion reveals mind in motion. *Frontiers in Psychology*, *2*. doi: 10.3389/fpsyg.2011.00059

Friedman, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the American Statistical Association*, *32*(200), 675–701.

Gordon, R. M. (1986). Folk psychology as simulation. *Mind & Language*, *1*(2), 158–171.

Gray, H. M., Gray, K., & Wegner, D. M. (2007). Dimensions of mind perception. *Science*, *315*(5812), 619-619. doi: 10.1126/science.1134475

Gray, K., & Wegner, D. M. (2012). Feeling robots and human zombies: Mind perception and the uncanny valley. *Cognition*, *125*(1), 125–130.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, *102*(1), 4.

Greenwald, A. G., Brendl, M., Cai, H., Cvencek, D., Dovidio, J. F., Friese, M., … others (2022). Best research practices for using the implicit association test. *Behavior Research Methods*, *54*(3), 1161–1180.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, *74*(6), 1464.

Hartigan, J. A., & Wong, M. A. (1979). Algorithm as 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, *28*(1), 100–108. Retrieved 2023-12-07, from `http://www.jstor.org/stable/2346830`

Henschel, A., Laban, G., & Cross, E. S. (2021). What makes a robot social? a review of social robots from science fiction to a home or hospital near you. *Current Robotics Reports*, *2*, 9–19.

Ishii, T., & Watanabe, K. (2019). How people attribute minds to non-living entities. In *2019 11th International Conference on Knowledge and Smart Technology (KST)* (pp. 213–217).

Karpinski, A., & Steinman, R. B. (2006). The single category implicit association test as a measure of implicit social cognition. *Journal of Personality and Social Psychology*, *91*(1), 16-32.

Kiesler, S., Powers, A., Fussell, S. R., & Torrey, C. (2008). Anthropomorphic interactions with a robot and robot–like agent. *Social Cognition*, *26*(2), 169-181. doi: 10.1521/soco.2008.26.2.169

Levin, D. T., Killingsworth, S. S., Saylor, M. M., Gordon, S. M., & Kawamura, K. (2013). Tests of concepts about different kinds of minds: Predictions about the behavior of computers, robots, and people. *Human–Computer Interaction*, *28*(2), 161-191. doi: 10.1080/07370024.2012.697007

Li, J. (2015). The benefit of being physically present: A survey of experimental works comparing copresent robots, telepresent robots and virtual agents. *International Journal of Human-Computer Studies*, *77*, 23–37.

Li, Z., Terfurth, L., Woller, J. P., & Wiese, E. (2022). Mind the machines: Applying implicit measures of mind perception to social robotics. In *2022 17th ACM/IEEE In-*

*ternational Conference on Human-Robot Interaction (HRI)* (p. 236-245). doi: 10.1109/HRI53351.2022.9889356

Malle, B. F. (2019). How many dimensions of mind perception really are there? In *Cogsci* (pp. 2268–2274).

Manera, V., Schouten, B., Becchio, C., Bara, B. G., & Verfaillie, K. (2010). Inferring intentions from biological motion: a stimulus set of point-light communicative interactions. *Behavior Research Methods*, *42*(1), 168–178.

McNeill, D. (1985). So you think gestures are nonverbal? *Psychological Review*, *92*(3), 350.

Nemenyi, P. B. (1963). *Distribution-free multiple comparisons.* Princeton University.

Nosek, B. A., Hawkins, C. B., & Frazier, R. S. (2011). Implicit social cognition: From measures to mechanisms. *Trends in Cognitive Sciences*, *15*(4), 152–159.

Okanda, M., Taniguchi, K., & Itakura, S. (2019). The role of animism tendencies and empathy in adult evaluations of robot. In *Proceedings of the 7th International Conference on Human-Agent Interaction* (pp. 51–58).

Pekçetin, T. N., Aşkın, G., Evsen, Ş., Karaduman, T. D., Eroglu, A., Barinal, B., … Urgen, B. A. (2023). A Set of Communicative and Noncommunicative Action Video Stimuli for Human-Robot Interaction Research. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 45).

Pekçetin, T. N., Barinal, B., Tunç, J., Acarturk, C., & Urgen, B. A. (2023). Studying Mind Perception in Social Robotics Implicitly: The Need for Validation and Norming. In *Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction* (pp. 202–210).

Pekçetin, T. N., Evsen, Ş., Pekçetin, S., Acarturk, C., & Urgen, B. A. (2023). A Naturalistic Setup for Presenting Real People and Live Actions in Experimental Psychology and Cognitive Neuroscience Studies. *JoVE (Journal of Visualized Experiments)*(198), e65436.

Pekçetin, T. N., Evsen, S., Pekçetin, S., Acarturk, C., & Urgen, B. A. (2024). Real-World Implicit Association Task for Studying Mind Perception: Insights for Social Robotics. In *Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction* (pp. 837–841).

Pekçetin, T. N., Evsen, Ş., Pekçetin, S., Karaduman, T. D., Acarturk, C., & Urgen, B. A. (2024). A Naturalistic Laboratory Setup for Real-World HRI Studies. In *Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction.* doi: https://doi.org/10.1145/3610978.3641113

Salem, M., Eyssel, F., Rohlfing, K., Kopp, S., & Joublin, F. (2011). Effects of Gesture on the Perception of Psychological Anthropomorphism: A Case Study with a Humanoid Robot. In *Social Robotics: Third International Conference, ICSR 2011, Amsterdam, The Netherlands, November 24-25, 2011. Proceedings 3* (pp. 31–41).

Saltik, I., Erdil, D., & Urgen, B. A. (2021). Mind perception and social robots: The role of agent appearance and action types. In *Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction* (pp. 210–214).

Sciutti, A., Bisio, A., Nori, F., Metta, G., Fadiga, L., & Sandini, G. (2013). Robots can be perceived as goal-oriented agents. *Interaction Studies*, *14*(3), 329–350.

Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and brain sciences*, *3*(3), 417–424.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, *27*(3), 379-423.

Straub, I. (2016). ‘It looks like a human!’ The interrelation of social presence, interaction and agency ascription: a case study about the effects of an android robot on social agency ascription. *AI & Society*, *31*, 553–571.

Thellman, S., de Graaf, M., & Ziemke, T. (2022, sep). Mental State Attribution to Robots: A Systematic Review of Conceptions, Methods, and Findings. *J. Hum.-Robot Interact.*, *11*(4), 1–51. doi: 10.1145/3526112

Thellman, S., & Ziemke, T. (2020). Do you see what I see? Tracking the perceptual beliefs of robots. *Iscience*, *23*(10).

Thorndike, R. L. (1953). Who belongs in the family. *Psychometrika*, 267–276. doi: 10.1007/BF02289263

Turing, A. M. (1950, 10). Computing Machinery and Intelligence. *Mind*, *59*(236), 433-460. doi: 10.1093/mind/LIX.236.433

Waytz, A., Gray, K., Epley, N., & Wegner, D. M. (2010). Causes and consequences of mind perception. *Trends in Cognitive Sciences*, *14*(8), 383-388. doi: https://doi.org/10.1016/j.tics.2010.05.006

Weisman, K. (2022). Extraordinary entities: Insights into folk ontology from studies of lay people’s beliefs about robots. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 44).

Weisman, K., Dweck, C. S., & Markman, E. M. (2017). Rethinking people’s conceptions of mental life. *Proceedings of the National Academy of Sciences*, *114*(43), 11374-11379. doi: 10.1073/pnas.1704347114

Wiese, E., Mandell, A., Shaw, T., & Smith, M. (2019). Implicit mind perception alters vigilance performance because of cognitive conflict processing. *Journal of experimental psychology: applied*, *25*(1), 25.

Yu, Z., Wang, F., Wang, D., & Bastin, M. (2012). Beyond reaction times: Incorporating mouse-tracking measures into the implicit association test to examine its underlying process. *Social Cognition*, *30*(3), 289-306. doi: 10.1521/soco.2012.30.3.289

Zaini, H., Fawcett, J. M., White, N. C., & Newman, A. J. (2013). Communicative and noncommunicative point-light actions featuring high-resolution representation of the hands and fingers. *Behavior Research Methods*, *45*(2), 319–328.

Złotowski, J., Strasser, E., & Bartneck, C. (2014). Dimensions of anthropomorphism: from humanness to humanlikeness. In *Proceedings of the 2014 ACM/IEEE International Conference on Human-robot Interaction* (pp. 66–73).