The Case of the Curious Robot: On the Social Viability of Curious Behavior in Non-Human Agents

Kevin Weatherwax
wax@ucsc.edu
University of California, Santa Cruz
USA

Doug Dooley
douglasbdooley@gmail.com
University of California, Santa Cruz
USA

Elin Carstensdottir
ecarsten@ucsc.edu
University of California, Santa Cruz
USA

Leila Takayama
takayama@hokulabs.com
University of California, Santa Cruz
USA



Figure 1: Stills from video stimuli. The robot agent is summoned to pick-up mail, leaves its charging station, travels down a cubicle hall, notices a box of stuff at which point it performs one of four levels of curious behavior focused on the box, and finally arrives at a persons desk to receive a letter.

ABSTRACT

Curiosity is increasingly being looked at to develop interactive non-human agents capable of lifelong learning, particularly in the space of Ai driven agents. However, the way a non-human agent's mental state is described could prime expectations in human counterparts as well as alter perceptions of an agent's behavior. We conducted an online experiment examining people's perceptions of four levels of curious behavior portrayed by a robot agent and whether this behavior matched people's expectations when primed to anticipate a "curious" vs. "learning" vs. "autonomous" robot agent. As curiosity increased, the ratings of the robot agent's thinking ability also increased. This was juxtaposed by a decrease in the robot's ratings as an effective worker and social agent. Further, we captured a unique trajectory of expectation matching showing that agents' *internal* mental states should be matched with analogous *external* behaviors. Our findings have implications for the design and development of non-human agents that engage in learning modeled on forms of human cognition.

CCS CONCEPTS

• **Human-centered computing → Empirical studies in interaction design**.



This work is licensed under a Creative Commons Attribution International 4.0 License.

IVA '24, September 16–19, 2024, GLASGOW, United Kingdom
© 2024 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-0625-7/24/09
https://doi.org/10.1145/3652988.3673920

KEYWORDS

Curiosity, Workplace Robots, Behavior Design

ACM Reference Format:
Kevin Weatherwax, Doug Dooley, Elin Carstensdottir, and Leila Takayama. 2024. The Case of the Curious Robot: On the Social Viability of Curious Behavior in Non-Human Agents. In *ACM International Conference on Intelligent Virtual Agents (IVA '24), September 16–19, 2024, GLASGOW, United Kingdom.* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3652988.3673920

1 INTRODUCTION

There is growing interest in creating robots that are capable of lifelong learning, enabling robots to expand beyond their preprogrammed capabilities over time [25, 47, 61, 65]. However, traditional approaches to artificial learning for nonhuman agents, while impressive in their ability to find and reproduce abstract patterns in large datasets, have struggled with operationalizing what we consider to be actual learning. This is because human learning isn't simply a product of input and output but rather it is the outcome of motivation through a complex interplay of physiology, emotion, and cognition, all situated in–and acting in response to–the world [27, 44, 45, 59, 60]. Often considered to be at the center of this motivation, driving those other facets of learning, is curiosity. For many artificial intelligence (AI) researchers, curiosity is seen as a possible missing link in developing true "thinking machines" that are inspired by various conceptions of human curiosity that drive intrinsically motivated robot learning over time [8, 11, 12, 42, 45, 66]. This is because curiosity, as a motivational manager of learning, can drive the discovery, understanding, mastery, and contextualization of "rare or deceptive rewards" [44].

While curiosity is a potentially useful construct for developing machines capable of life-long learning, the implications for human to non-human agent interactions are less clear. There is a risk that these curious non-human agents might not develop into systems that match the abilities [2] or values [10] of the people who live and work around them, especially if their responses to these robots are not taken into account. It may be especially important to reconcile internal functions of curiosity (e.g., subject of focus, interest, intent) with appropriate external representations of the same social and behavioral states, as demonstrated with expressive robots in the past [64]. Just a few of these questions include:

- Will (untrained) people recognize curiosity in a robot?
- Is there such a thing as having too much robot curiosity?
- In which situations will a curious robot be acceptable vs. unacceptable to the human stakeholders around it?
- Does a robot's internal state (mind) need to match its external behaviors (body)?

Essentially, we must also understand the external aspects of curiosity in robots; how it is experienced, perceived, interpreted in robotic form(s) and how we design for it. While research in human to non-human agent interaction–and in psychology more generally–has explored some of these questions directly, and tangentially by studying human relationships [6, 13, 14, 26, 28–30, 35, 37–39, 43, 45, 46, 51, 52], many important gaps remain. Without attention, curious robots could be poorly received, seen as intrusive, creepy, or threatening, regardless of ability or usefulness [54].

This paper explores how curious robot behaviors are perceived and interpreted by untrained end-users who might one day encounter robots in their workplaces. Here, we evaluated whether priming participants to have different expectations of a robot's cognitive model (i.e., being curious, learning, or autonomous) would alter people's perceptions of increasingly curious behavior. Additionally, we examined how well a robot's behavior matched participants' expectations based on how they were primed. From a design perspective, our study shows that curiosity can be communicated effectively through robot behaviors, using the actual degrees of freedom and joint limits of a real mobile robot. Moreover, our empirical findings regarding the boundaries of acceptable behaviors have implications (and raise more questions) for researchers developing robots, and other non-human agents, that seek to learn.

2 RELATED WORKS

To situate the current work, we present a brief overview of curiosity, then review where we stand in developing computational curiosity for learning agents, how curiosity is expressed by people, and how behavior design can be used to express robot curiosity. We explain how curiosity drives human learning and has become an area of focus for algorithmic approaches aimed at imitating the same kind of adaptive learning humans are capable of. We also discuss how curiosity, in human-centered research, is not a wholly internal process but also includes external signals that trigger social interactions which facilitate the learning process. Computational approaches to instantiating curiosity in non-human agents (e.g., robots) has, conversely, not considered these important social and behavioral facets. Finally, we discuss how curious–and other learning based–robots could be designed more thoughtfully by aligning their internal (mind) and external (body) states.

2.1 What is curiosity?

Curiosity is a complex construct and a fundamental facet of what drives human learning. The prevalent view, informed by Loewenstein, conceptualizes curiosity as a kind of motivating drive for exploration and learning in humans and animals [14, 29, 31, 44]. In particular Loewenstein framed the catalyst of curiosity as the realization of a "knowledge gap" that an agent then tries to fill. To do this, individuals exploit previously obtained knowledge/skills to engage in a mode of learning that involves an active search for, and exploration of, novel situations and information. With this in mind we can frame curiosity as a cycle whereby an agent; (1) recognizes an **information gap**, (2) develops the **motivation** to fill it with relevant information, (3) actively **seeks knowledge** in an area that makes the most sense based on previous discoveries, and (4) **rewarded** by the acquisition of new knowledge that can be exploited, seeks new avenues for exploration.

For humans, curiosity is a recognizable trait. People can easily tell when someone is curious, what they are curious about, and–perhaps most importantly–whether they can/should intervene to provide assistance. But what does a curious robot look like? In a review of robotics curiosity-based learning, Oudeyer highlighted a curious robot arm that explores its environment through a pair of joysticks. These joysticks controlled the movements of a mechanical arm confined inside a small play area with lights, sounds, and objects the robot could affect. This robot was able to discover novel effects it could have on the environment and then practice recreating these until the novelty transformed to mastery of new skills [44]. While this is a remarkable demonstration (see Oudeyer's demo: here) of the potential of intrinsically motivated learning (i.e., curiosity), to an outside observer the robot arm seems to grope about the surrounding space for long periods of time in erratic jerking movements. This specific example raises an important problem central to this study. Namely, that curiosity does not exist in a vacuum.

Human curiosity is, in large part, facilitated by interactions with other agents in the environment (e.g., teachers, peers, mentors) not simply in response to it. So, while progress has been made on developing nonhuman agents capable of curiosity-driven learning, the source of mentoring has largely been the scientists who develop it. They provide parameters, assistance, direction, and guidance, as well as tightly control the environment to facilitate success, much like a preschool teacher, to their robots because they are just that; their robots. In real world settings, this relationship may fall apart when embodied agents are not designed to take social and organizational contexts into account [40]. Goren Gordon, a leading researcher in curious robots, offers a poignant assessment on the strengths and limitations of what robots can learn in non-social vs. social settings using curiosity-based algorithms. Namely, a curious robot can learn very well in isolation, but only about itself, while social settings introduce far more complex and dynamic factors [12]. Gordon goes on to stress how the success of robot curiosity, in a social setting, is largely dependent on people's response to it.

Moreover, the only prior work that has directly examined human perceptions of a curious robot (rather than implementations of curiosity based learning for robots) identified that curious deviations during a prescribed task resulted in negative considerations of the robot [68]. This is especially interesting considering that much of the ongoing research on developing curious robots assumes curiosity to be a universally positive trait. These kinds of blanket assumptions of non-human agent usefulness have proven to be detrimental in the past. For example, Mutlu and Forlizzi's ethnographic study of a delivery robot deployed in various hospital settings found that, in a high impact trauma unit, a robot was seen as a burdensome nuisance but in a maternity ward was a welcome addition. Furthermore, these hospital robots were sabotaged, abused, and stuffed into utility closets in the trauma center but received warmly and viewed as useful in the postpartum center [40]. We can infer then that for robots to engage in lifelong learning in human-populated spaces (e.g., offices, hospitals, airports, shopping malls), it will be similarly important for us to figure out when, where, and how robots should engage in curiosity-driven exploration in a way that will be acceptable for human co-workers and bystanders. One potentially fruitful starting point for developing socially acceptable curious non-human agent behaviors is to look for inspiration in how people express curiosity.

2.2 Expressing Human Curiosity

Although there are many nuanced facets of human curiosity (e.g. Kashdan's five dimensions of curiosity [28]), exploratory and investigative behaviors are its most readily recognizable components. Internal cognitive processes of curiosity often manifest externally as focus and attention (for example, interest, specific curiosity) or wandering and seeking new experiences (e.g., exploration, divisive curiosity) [3, 4]. It is important to note that curiosity research is closely intertwined with interest research [55, 58–60], so interest is included in some of the behaviors described below.

Looking to past work on curiosity is helpful for describing what it means for someone–or some robot–to behave in a curious manner. As such, we focus on the observable traits that have been operationalized or measured in curiosity research. One of the most clear cut curious behaviors is exploration. At its core, exploration refers to investigative behaviors (e.g., seeking, searching, manipulating) relating to environmental stimuli (e.g., specific things, exploring spaces) and is often the focus of child developmental work [6, 26, 49].

Facial expressions and head tilting are also useful signals for recognizing curiosity in others. In a pair of experiments, Reeve [51, 52] captured changes in facial expressions when people were shown interesting stimuli. A more recent empirical study, aimed at developing curiosity auto-recognition, examined videos of children primed to feel curious about objects (novel vs. mundane); when participants were primed to feel curious, the researchers noticed changes in the gaze, head shifts (i.e., panning and tilting), and moving around the object to focus on different parts of it [43].

Another strong behavioral signal of curiosity, in humans, is spending more time observing and becoming familiar with novel objects [43]. This is consistent with prior research on curious behavior in children, in which reacting to new stimuli by moving toward it and spending more time examining it was indicative of high curiosity [37–39]. As each robot has a different set of affordances, these expressions of human curiosity cannot be mapped directly onto any given robot, but they can be used to inspire the design of curious robot motions.

2.3 Expressing Robot Curiosity

Robot expressive motions are a fundamental part of human-robot interaction, especially because robot motions influences how people perceive and interact with it [15, 20]. People read meaning even in the motion of abstract shapes [16]. As such, it is important to be intentional about the design of expressive robot motions rather than leaving human interpretation of robot motions up to chance. Prior work in HRI has explored using animation techniques in the design of expressive robot motion, e.g., showing forethought and responses to success or failure [64], showing preparation and follow-through in a human-robot jazz performance [21], and shared attention in a music-listening companion robot [17]. Expressiveness has been used to facilitate successful interactions and collaborations with robots. Takayama et al., demonstrated how readable robot behaviors can convey rich information about a robot's intentions and capability to human observers, as well as change how they subjectively rate the robot [64]. Similarly, Hoffman and Breazeal, showed how adding anticipatory robot motions improved the fluency of human-robot collaborative activities [18, 19]. The current study builds upon these prior works, expanding into the space of curious exploration. Algorithmic approaches to generating expressive robot motions are in development, starting with making robot motions more legible and predictable [7, 63] as well as simultaneously generating functional and expressive robot motions [67].

In the current work, we focus upon the expression of curiosity. Some social robots have been developed to express curiosity, e.g., Dragonbots designed to elicit child curiosity [13], Naos designed to play educational games with children [5], and a Wizard-of-Oz Recyclo robot that gave recycling recommendations [34]. These systems have mainly relied on *verbal* expressions of curiosity, such as asking questions or explicitly stating that they are curious about something. They are also designed for learning contexts. In contrast, the current study focuses on non-verbal, behavioral expressions of curiosity in a professional workplace context. While there are many settings where it could be useful for a robot to express curiosity (e.g., learning contexts [5, 13, 34]), it is unclear if it is always going to be a good idea for robots to express curiosity. For example, when a robot has a specific task that it is supposed to be doing, like delivering mail in an office place, will it be acceptable for the robot to pause its primary task in order to curiously explore other things along the way? Prior work has shown that curious robot behaviors can increase engagement, but can simultaneously hurt trust in the robot [34] and can make people think that the robot is less competent [68]. Additionally, since curiosity based learning robots will always be evolving–theoretically–in their interest, ability, and capability, we need to better understand how to signal that a robot is curious (i.e., capable of being curious) and when it is curious.

3 METHODS

In this online study, we examined whether priming participants to have different expectations of the robot (i.e., being curious, learning, or autonomous) would alter their perceptions of increasingly

curious behavior. We also tested how well the robot's behavior matched participants' expectations based on how they were primed. Following each treatment level, participants provided open-ended descriptions of the robots' behavior and explanations about what they liked/disliked, and why.

3.1 Study Design

We used a mixed-factorial 4 (curiosity level: none, low, moderate, and high; within-subjects) x 3 (expectation priming: curious vs. learning vs. autonomous; between-subjects) experiment design to test the following hypotheses:

- H1 People hold distinct mental models for robots (how it should behave) depending on the cognition they believe it has.
- H2 People primed to expect a **curious robot** will rate the robot as more effectively matching their expectations as the level of curious behavior goes up.
- H3 People primed to expect a **learning robot** will rate it as more effectively matching their expectations at the lower levels of curious behavior but lower at the highest level.
- H4 People who expect an **autonomous robot** will rate the robot as more effectively matching their expectations when it exhibits little or no curious behavior and not when it exhibits greater amounts of curious behavior.

In particular we were interested in the hypothesized difference between a curious and a learning robot for two reasons. First, in our review of literature on computational models of the human mind, broadly, and curiosity, specifically, we found a tendency to use, interchangeably, any synonym of the same. Yet, we saw compelling arguments–often in the same works–that concepts of human thinking, learning, curiosity, etc., are quite distinct in both human function and how they are operationalized for computational purposes. This brings us to the second point: The meaning of learning is somewhat more bounded by potential outcomes (i.e., something becomes learned), whereas curiosity is, comparatively, self-sustaining because it is a motivational state (i.e., someone is curious) without a specific end point.

In Walker et al.'s study participants reported more negative views of a robot that deviated from a prescribed task to perform curiosity based exploration. We observed similar effects during piloting rounds for this study (i.e., as the robot became more curious it was seen as less competent, focused, responsible, and likable) [68]. As such we posed the following research question:

- RQ1 Will priming people to anticipate specific types of internal mental states in a robot (i.e., curious, learning, or autonomous) offset negative reactions to increasing curiosity behaviors?

A secondary aspect of this work was validating the design of curious behavior for our robot agent. As such, we also pose the following final research question:

- RQ2 When an agent expresses greater amounts of curious behavior, compared to lower amounts, will people be more likely to assign it traits relating to curiosity.

3.2 Stimuli

We developed four videos of a Mayfield Robotics Kuri operating as a delivery robot in an office setting (see: Figure 1). The Kuri was chosen because it was a widely accessible robot designed for interactive social expression that our curious behaviors could be actualized for in-person follow-up experiments. In all of the videos the robot passes by a box in a hallway between cubicles and either ignores it or examines it with various levels of interest (i.e., curiosity).

All scenes were animated with Autodesk Maya, carefully adhering to the joint limits and degrees of freedom of Kuri's hardware. Motion was restricted to realistic base movement, head pan and tilt, and eyelid opening and closing. To fit the narrative we chose, the Kuri model was altered to include a backpack to hold items.

3.2.1 Level of Curious Behavior. We worked with a professional character animator, experienced in motion design for real-world robots, to develop the curiosity behaviors for this study. Curious behaviors were created, and heightened, by altering the focus, frequency, and duration of the body, head, and eye movements. All curiosity behaviors were culled from past research on observable traits of curiosity. These included increasing head movements [43], shifts in gaze and facial expression [51, 52], moving closer to the subject of interest [37–39], and exploring the environment [6, 26, 49].

- **No Curiosity (L1)** = slight head movements and blinks while traveling through the hall, no head rotation toward box, no slowing down, or any acknowledgement box is there, except to move around it.
- **Low Curiosity (L2)** = head pans/tilts looking around while moving through hallway, head rotates toward the box, slows down while passing and head pans/tilts focused on the box, but there is no body rotation towards the box.
- **Medium Curiosity (L3)** = head pans/tilts while moving through hallway, head and body rotate toward the box; robot circles the box with 3 additional pauses, head pans/tilts quickly while looking at the box and surrounding area, before continuing its original route.
- **High Curiosity (L4)** = head pans/tilts while moving through hallway, head and body rotate toward the box; robot circles the box twice with 7 additional pauses, head pans/tilts slowly while looking at the box and surrounding area, before continuing its original route.

Because the higher levels of curiosity required more behaviors, the length of the videos were 31, 38, 56, and 126 seconds, respectively, for no, low, medium, and high curiosity. In our animation development and pilot rounds, we attempted to hold the time duration constant across conditions, but were unable to do so without introducing confounds to the experiment design (e.g., idle or erratic behaviors). In the end, we made the decision to keep the durations as close together as we could, but to prioritize the expressive curiosity, acknowledging that more curious behavior may inherently take more time to perform. Though a control for time would be ideal past research on curiosity has shown that a major facet of how people determine the intensity of curiosity in others is the amount of time they spend engaged with the subject of interest [37–39, 43].

3.3 Method

3.3.1 Participants. Respondents (N=64) were recruited from Amazon Mechanical Turk. Each participant received two U.S. dollars as a token of thanks for their time. It took approximately 10 minutes to complete the study. From an original pool of 69, five participants were removed for having a response time significantly lower than the total duration of the videos they were asked to watch and/or failing attention check via straight-lining behavior. There were 39 males, 23 females, 1 non-binary, and 1 person who preferred not to answer (self-reported) with ages ranging from 26 to 61 years (M = 38.98, SD = 8.9). Education levels of respondents included completion of high-school (n = 4), some college (n = 14), associates degrees (n = 7), bachelors degrees (n = 31), masters degrees (n = 7), and professional degrees (n = 1). In terms of reported experience with robots; 55 people had none, 3 a little, 3 a moderate amount, 2 lots, and 1 preferred not to answer.

3.3.2 Procedure. We randomly assigned participants and counterbalanced them to one of three between-subjects conditions (curious robot n = 21, learning robot n = 21, and autonomous robot n = 22). Depending on condition, we told participants they would see videos of a "curious", "learning", or "autonomous" robot working in an office. Then we requested they write record their expectations, with an open ended item, of how a curious robot, a learning robot, or an autonomous robot should behave (depending on condition).

Next, participants saw videos depicting four different curiosity levels (none, low, medium, high), randomized for order. Each video began with a five second title card reiterating that they will see a curious, learning, or autonomous robot working in an office delivering mail. They were asked to watch each video in full screen mode. After each video, participants responded to a questionnaire about the robot they just saw. Each questionnaire consisted of 19 items presented on a 7-point Likert scale of agreement.

We designed the survey items to gauge the readability of our curiosity animations as well as to assess participants' general perceptions of the robot. We chose questions focused on perceived curiosity based on common tangible/observable factors of curiosity (i.e., pertaining to curiosity as behavior, not mental state). Similarly, we picked items regarding general perception of robots in order to capture common aspects of HRI research (e.g., trust, likability, intelligence). We also included an attention check for our working state between-subjects variable by asking participants to select from four options to describe what the robot was doing when it passed the box in the hall. Next we asked participants to rate, on a single 5-point Likert-type scale, how well the robot they saw matched the expectations they recorded at the beginning of the study. Then, we asked participants to fill out two open ended items describing the robots' behavior and what they liked/disliked about the robot. The 7-point scale of agreement items, the 5-point expectation matching item, and the two open-ended responses were repeated after each video. Finally, we asked participants to share their age, gender identity, educational level, and experience with robots.

4 RESULTS

4.1 Order Effects and Manipulation Checks

Our presentation of stimuli (i.e., curiosity level) was randomized between participants. However, since curiosity level was a within-subjects independent variable, there was a risk that we would see order effects in the data (i.e., having seen the first condition would affect responses to the subsequent conditions). To address this concern, we restructured the data in terms of the first, second, third, and fourth curiosity condition presented to each participant. We then ran a mixed ANOVA, using order of curiosity stimuli presented (first, second, third, fourth) as a within-subjects independent variable and working state as a between-subjects independent variable. In running analyses for all dependent variables, we did not find significant effects of the order of presentation of the stimuli.

4.2 Quantitative Data Analysis

4.2.1 Pre-Processing. We began our analysis by pre-processing our data. We performed principal component analysis (PCA), scale reliability testing, and component construction at each of the four levels of curious robot behavior. We identified three components which loaded together, at all four levels of our repeated measure, with acceptable loading scores greater than .40 [48, 62]. Then we constructed three new components with the mean scores of related items from our Likert-type scales of agreement. The first was the robot as a "thinking agent" which was built from items pertaining to the robot; desiring knowledge, learning on its own, being independent, intelligent, and curious. Second, was the robot as a "working agent" that was constructed from items of the robot being; focused, competent, obedient, unobtrusive, acceptable, and following instructions. Third, was the robot as a "social agent" which was assembled from items relating to the robot being; trustworthy, likable, responsible, and approachable. Next, we conducted reliability testing on each of the three component sets, which all returned acceptable Cronbach's alphas ($\alpha > 0.7$) at each level of our repeated measure (for details see: Appendix 1). Four individual items that were strongly independent of all others during PCA. These are related to the robot; being autonomous, adaptable, exploring on its own, and taking initiative to perform new tasks.

4.2.2 Results. We ran a mixed ANOVA, using the level of curiosity as a within-subjects independent variable and priming expectations of a robot's cognitive state (Curious vs. Learning vs. Autonomous) as a between-subjects IV to predict how well the robots' behaviors matched expectations. We found a significant main effect of robot curiosity behaviors, $F(3,183)=7.03$, $p<.001$, $\eta_p^2=.10$, on how well the robot behaviors matched the expectations of the participants.

To identify where this occurred specifically, we performed repeated contrasts on each of the four levels of curiosity behaviors. Here we found significant difference between levels 1 vs. 2, $F(1,61)=23.22$, $p<.001$, $\eta_p^2=.28$, and levels 3 vs. 4, $F(1,61)=17.20$, $p<.001$, $\eta_p^2=.22$. Our repeated contrasts did not reveal a difference between levels 2 vs. 3. Depending on the level of curiosity behavior the robot exhibited, people felt it matched their expectations more or less, particularly between levels 1 vs. 2 and 3 vs. 4.

We also observed a main effect of the anticipated cognitive state (Curious vs. Learning vs. Autonomous), $F(2,61)=2.15$, $p<.05$, $\eta_p^2=.12$, on expectation matching. Furthermore, we captured a significant interaction effect between robot behavioral curiosity and anticipated cognitive state, $F(6,183)=15.93$, $p<.001$, $\eta_p^2=.10$, and the robot matching participant expectations. These contrasts revealed a significant difference at all four levels of curiosity: Levels 1 vs. 2, $F(2,61)=8.48$,

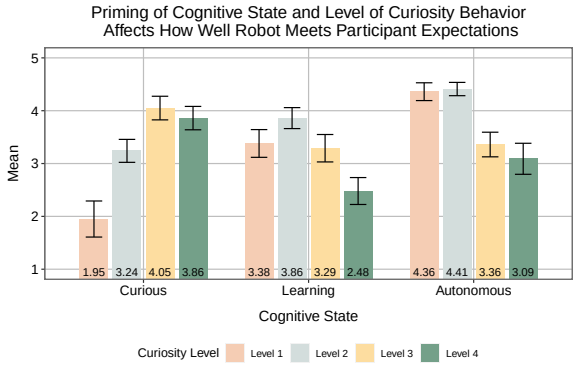


Figure 2: Robot's curiosity and anticipated cognitive state (Curious vs. Learning vs. Autonomous) impacted how well the robot met participant expectations.

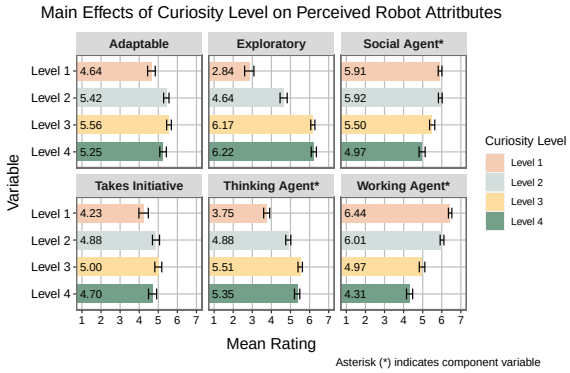


Figure 3: Impact of curiosity behavior on rating the robot as adaptable, exploratory, and taking initiative as well as a social agent, thinking agent, and working agent,.

p=.001, η_p^2=.22; levels 2 vs. 3, F(2,61)=19.90, p<.001, η_p^2=.31; and levels 3 vs. 4, F(2,61)=3.56, p<.05, η_p^2=.11. See Figure 2.

We ran mixed ANOVAs using level of curiosity as a within-subjects independent variable and priming expectations of a robot's cognitive state (Curious vs. Learning vs. Autonomous) as a between-subjects independent variable to predict ratings of the robot on each of our three constructed components (i.e.,thinking agent, working agent, and social agent) separately. We found a significant main effect of a robot's curiosity behavior, F(3,183)=50.09, p<.001, η_p^2=.45, on the "thinking agent" component. To identify where the change in ratings of the robot as a thinking agent occurred, we ran repeated contrasts on each of the four levels of curiosity behaviors. Here we found significant difference between levels 1 vs. 2, F(1,61)=62.47, p<.001, η_p^2=.51, and levels 2 vs. 3, F(1,61)=20.60, p<.001, η_p^2=.25. We did not observe a significant difference between levels 3 vs. 4.

Additionally, we captured a main effect of the curiosity behaviors, F(3,183)=96.49, p<.001, η_p^2=.61, on rating the robot as a "working agent". Again, we ran repeated contrasts on each level of curiosity behaviors. We observed significant difference between levels 1 vs. 2, F(1,61)=30.33, p<.001, η_p^2=.33, levels 2 vs. 3, F(1,61)=55.84, p<.001, η_p^2=.48, and levels 3 vs. 4, F(1,61)=28.86, p<.001, η_p^2=.32. We also found a significant main effect of the robots' curiosity behaviors, F(3,183)=27.68, p<.001, η_p^2=.31, on ratings of the "social agent" component. To identify where the change in ratings of the robot as a social agent occurred, we again ran repeated contrasts on each of the four levels of curiosity behaviors. Here we found significant difference between levels 2 vs. 3, F(1,61)=12.63, p=.001, η_p^2=.17, and levels 3 vs. 4, F(1,61)=26.98, p<.001, η_p^2=.31, but not between levels 1 vs. 2. We did not observe a main effect of robot cognitive state primes or interaction effects for participant ratings of the robot as a thinking, working, or social agent. See: Figure 3. For additional explanation of these findings see: Section 4.4.

Finally, we ran mixed ANOVAs using level of curiosity as a within-subjects independent variable and priming expectations of a robot's cognitive state (Curious vs. Learning vs. Autonomous) as a between-subjects independent variable to predict ratings of the robot on our four remaining individual items (i.e.,. adaptable, likely to explore, taking initiative on new tasks, and autonomous). We found a significant main effect of the robots' curiosity behaviors, F(3,183)=7.33, p<.001, η_p^2=.11, on ratings of the robots' adaptability. To identify where the change in perceived robot exploration occurred, we ran repeated contrasts on each of the four levels of curiosity behaviors. Here we found significant difference between levels 1 vs. 2, F(1,61)=17.05, p=.001, η_p^2=.22, and levels 3 vs. 4, F(1,61)=4.53, p<.05, η_p^2=.07, but not between levels 2 vs. 3. We also captured a significant main effect of the robots' curiosity behaviors, F(3,183)=82.17, p<.001, η_p^2=.57, on ratings of robot exploration. Here, contrasts revealed significant difference between levels 1 vs. 2, F(1,61)=50.25, p<.001, η_p^2=.45, and levels 2 vs. 3, F(1,61)=52.76, p=.001, η_p^2=.46, but not between levels 3 vs. 4. We found a significant main effect of the robots' curiosity behaviors, F(3,183)=3.25, p<.05, η_p^2=.05, on ratings of the robot being likely to take initiative on new tasks. Here we only observed a significant difference between levels 1 vs. 2, F(1,61)=5.98, p<.05, η_p^2=.09. There was no main effect for robot cognitive state or interaction effect observed relating to any of the aforementioned individual survey items. No effects were found between our IVs and ratings of the robot as autonomous.

4.3 Sentiment Analysis

We examined user sentiment in order to look for emotional changes in participants' written feedback across conditions. We conducted sentiment analysis on participants' open ended responses using VADER (see: VADER for open source access). VADER generates a positive, negative, or neutral valence score (1.0 to -1.0). This provides a reliable metric for emotional tone in short text entries [24]. VADER was developed on tweets with a max length of 140 characters [24] making it suitable for examining our open ended responses which had a character length of M = 74.84 (SD = 55.91).

We ran a mixed ANOVA using level of curiosity as a within-subjects independent variable and priming expectations of a robot's cognitive state (Curious vs. Learning vs. Autonomous) as our between-subjects IV to predict the emotional tone of participants' written responses. We captured a significant main effect a robot's curiosity, F(3,183)=9.38, p<.001, η_p^2=.13, on the emotional tone of participant responses. There was no main effect for a robot's cognitive state or interaction effect observed. We also ran repeated

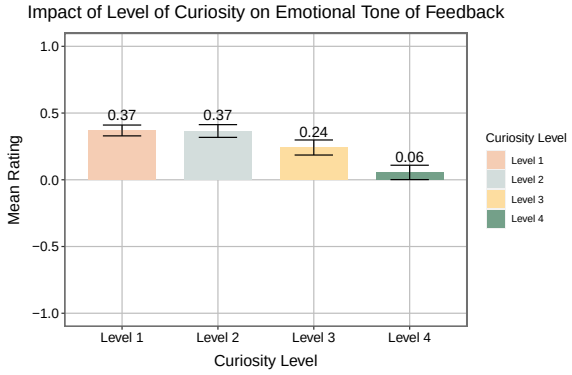


Figure 4: Robot curiosity behavior impacted emotional tone of participant open-ended responses. Significant change occurred at the highest level of curiosity.

contrasts on each of the four levels of curiosity behaviors. Here we observed a significant difference between levels 3 vs. 4, $F(1,61)=8.19$, $p<.01$, $\eta_p^2=.12$, but none between levels 1 vs. 2 or 2 vs. 3. See Fig. 4.

4.4 Summary of Findings

Our results provide evidence that intentionally designed robot behaviors can appropriately convey analogous internal states. We tested hypotheses about how well increasingly curious robot behaviors would match people's expectations when they were primed to anticipate a curious robot, a learning robot, or an autonomous robot. We found support for *H2*, when primed for a curious robot peoples' expectations were more effectively matched as the level of curious behavior went up. We also found support for *H3*, when primed for a learning robot people's expectations were most effectively matched in the low level of curiosity (Level 2) and then began to fall off at each subsequent level. We also found support for *H4*, when primed for an autonomous robot participant expectations would be most effectively matched when the robot showed little or no curiosity. Overall though in support of *H1*, people held unique expectations of external behaviors in the same robot based solely on single word primes about its internal mental state.

Unfortunately, priming participants to expect specific cognitive qualities from a robot did not offset negative changes to general perceptions of the robot (RQ1) as a social or thinking agent. Indeed, we observed near identical downward trends from our piloting rounds in ratings of the robot as both a social and working agent while ratings of its thinking ability went up, regardless of priming conditions (see: Figure 3). Our analysis of sentiment scores in open-ended response questions paralleled the survey data and we saw an overall negative trend in the emotional tone of people's feedback as curiosity increased. See: Figure 4. However, priming people to anticipate a robot possessing specific cognitive profiles (autonomous vs. learning vs. curious) and the level of curious behavior did impact how well its behaviors met their expectations (see: Fig. 2).

5 DISCUSSION

In this paper, we presented findings from an online experiment showing that there may indeed be such a thing as too much curiosity in non-human agents. As a social agent, our robot became significantly less likable at the highest level of curious expression. Similarly, it was perceived as being less competent as a working agent. Digging deeper, we observed negative differences between medium curiosity (L3) and high curiosity (L4). However, there were no negative differences between no curiosity (L1) and low curiosity (L2). Since there was a significant increase in the perceived cognitive ability of the robot at each stage, it may be that more subtle forms of curious expression will be more socially acceptable.

We found that when a person was primed to anticipate a curious robot, the more curious behavior it exhibited, the more the robot matched their expectations. An autonomous robot, on the other hand, met expectations when the robot showed no or very little curious behavior and quickly dropped off after that. Since learning and curiosity are often treated as interchangeable constructs, we were particularly interested in whether the learning condition would follow the same trajectory as the curious robot (i.e., as curious behavior increased so would expectation matching). We found that expectations of a learning robot followed its own unique trajectory peaking early and then steadily dropping off as the robot exhibited increasing amounts of curious behavior.

This suggests that when designing the internal cognitive models of a robot (i.e., how the robot thinks/learns), the relevant external qualities (i.e., how the robot acts/moves) should be simultaneously considered with the same level of thoughtfulness. In the future, when artificial learning processes for a robot are based on specific models of human cognition, then the robot should be carefully outfitted with the ability to express appropriate behavioral analogs. This may allow human understanding of, and interaction with, a robot to be intuitively informed. Furthermore, making sure to match human expectation with the robot's actual behaviors and abilities can help prevent expectation violations on the part of the human. What we captured in this study suggests that alignment of both mind and body is necessary for a harmonious robot.

This empirical work sheds light onto how different levels of curious robots' behaviors will affect people's perceptions of robots in the workplace. These results can be used to inform the design of robots that engage in curiosity-driven learning. In optimizing for lifelong learning, it will be important to include robot learning goals and considerations of how much curious exploration will be acceptable by the people in the robot's environment.

5.1 Designing Alignment of Body and Mind

While character animation techniques have long been used in the animation of 3-D computer graphics [33], the translation of animation techniques to robot motion design is less straightforward. Additionally, other researchers have suggested models of design approach that includes animation techniques as a central pillar [53]. Unlike on-screen characters, robots do not have infinite degrees of freedom, and real world robots need to abide by the laws of physics. While the HRI community is making headway in programmatically producing more human-readable motion trajectories [7, 19, 23, 32], there are still many challenges ahead as robots not only work with human collaborators, but also encounter bystanders in the workplace who are not necessarily directly interacting with them (e.g., staff who do not directly benefit from using the robot). These robots must be socially acceptable by everyone the workplace, including

those indirect stakeholders [9, 10], if they are going to stand a chance at being adopted for long-term use.

The professional animator who collaborated on this work used some formalizable guidelines for curious behavior design, like using gaze at a target to indicate minimal curiosity, adding body rotation toward the target to indicate more curiosity, and adding full-body pausing and circles around a target to indicate even more curiosity (see Sec 3.2.1 for additional details). However, there are other animation techniques employed that might be harder to formalize with existing approaches to motion planning for robot expressivity (e.g., deciding how to glance vs. look vs. stare). At least two approaches to incorporating those other techniques include: (1) bringing character animators onto robot motion design teams as full collaborators and (2) gathering enough examples of robot motions designed by animators to train machine learning algorithms to produce more clearly expressive robot motion trajectories.

5.2 Limitations & Future Work

A limitation of this study is that it was run online instead of in-person so we do not yet know exactly how these findings will generalize to real world robot deployments. Prior research comparing in-person to on-screen robotic agents has shown that the direction of effects is often similar between the two (e.g., [50, 69]), but effects can vary by robot interaction scenario type (e.g., task-oriented vs. conversation-oriented [22]. Furthermore, people tend to be more engaged with [1, 50] and feel more empathy for [57] a collocated physical robot than for an on-screen one. Because this study was conducted in the spring of 2021, we were limited in our ability to conduct in-person studies. We chose to use an on-screen prototype to study these curious robot movements (as recommended for movement-centric robot designs [20]. We made sure to produce behaviors that worked within the degrees of freedom and joint limits of the real Kuri hardware platform so that it will be possible for us to run follow-up studies, using the same curious behaviors, using the real world Kuri robots.

Another limitation is that the robot was only shown performing one specific task in one setting. During our initial study design we considered comparing other settings that a curious robot might act in but it quickly became too many variables to control before establishing the validity of our curious behavior designs. It is entirely conceivable that people don't want a delivery robot to be curious while it is working but that in other contexts it might be acceptable. Additionally, our displayed curiosity was–by design–mundane and without clear benefit. What if the curious exploration had yielded some tangible benefit (e.g., notice and address a safety hazard) for the robot and/or its human counterpart? Further work would benefit from exploring other settings, tasks, and outcomes.

Similarly, we only tried one intervention (i.e., priming expectations of the robot) in an attempt to offset negative views of robot curiosity. Future work should examine other social interventions. Would people be more generous evaluating curiosity at a time when it is beneficial to be especially inquisitive (e.g., first day on a job)?

Participants in our experiments also viewed the robot from a bystander perspective (i.e., third person perspective) rather than from a direct interactant perspective. While this could be a limitation, it is actually the perspective that most people will have of these robots and indirect stakeholders (e.g., bystanders) [10] are often important for considering in the design of real interactive systems.

As we mentioned in section 3.1, we were unable to hold video time constant. The robot simply needed more time to display its curiosity at the higher levels. In real world deployments, robots that do more curious exploration will likely take more time to explore than robots that do less (or no) curious exploration.

As with many studies focused on nonhuman agents, we only used one type of robot in one type of setting with only English-speaking participants. To assess the generalizability of these results, it will be important to replicate this study with other hardware platforms (e.g., robot arms) in other settings (e.g., classrooms, homes, hospitals) and with broader user groups (e.g., with other cultural groups). One of the strengths of online studies is that they are easier to replicate and conduct follow-ups. To that end, we have shared our materials so that others can run their own experiments with a similar framework may use the stimuli provided in Appendix 5.

The short-term nature of this study may make the results subject to novelty effect biases. If we had found extremely positive responses to the extremely curious robot behaviors, we would have been skeptical of these results simply showing how engaging a brief animation of a robot could be. However, most of these findings showed negative reactions to the curious robot behaviors. It is possible that people might come to see the benefit of curious robots over time as their performance improves or they become more attractive over time. As such, we are currently planning longer-term, real world deployments on this hardware platform.

As we develop curiosity in nonhuman agents, it will be important to take the perspectives of coworkers and bystanders into account. If agents, are going to succeed in performing their tasks, especially their learning tasks, they may need instruction, guidance, and feedback from the people in their work environments. In particular there is a growing call for look for ways to create harmonious interactions wherein both the robot and human agents' needs are met, as opposed to a paradigm where a robot achieving its goal without overtly bothering the human is considered successful [56].

Fortunately, there are creative ways to approach this from both an algorithmic and interaction design space. For instance, in a recent study, a robotic agent–in an online virtual setting–incorporated a model of individual human helpfulness to outperform other, traditional, ML approaches by asking human counterparts for help less often but at more optimal times [41]. The same study found that people had more negative social reactions to the robots that didn't adapt to individual and contextual factors when asking for help [41]. It seems that before we try and have lay-people offer guidance to non-human agents, we may need to better understand how expressing internal mechanisms of learning (e.g., curiosity) is perceived by human stakeholders–both direct and indirect.

REFERENCES

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

[2] Liam J Bannon. 1995. From human factors to human actors: The role of psychology and human-computer interaction studies in system design. In *Readings in human–computer interaction*. Elsevier, 205–214.

[3] Daniel Ellis Berlyne. 1954. A theory of human curiosity. *British Journal of Psychology. General Section* 45, 3 (1954), 180–191.

[4] D. E. Berlyne. 1966. Curiosity and Exploration. *Science* 153, 3731 (jul 1966), 25–33. https://doi.org/10.1126/science.153.3731.25

[5] Jessy Ceha, Nalin Chhibber, Joslin Goh, Corina McDonald, Pierre-Yves Oudeyer, Dana Kulić, and Edith Law. 2019. Expression of Curiosity in Social Robots: Design, Perception, and Effects on Behaviour. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–12.

[6] Amy Chak. 2007. Teachers' and parents' conceptions of children's curiosity and exploration. *International Journal of Early Years Education* 15, 2 (2007), 141–159.

[7] Anca D Dragan, Kenton CT Lee, and Siddhartha S Srinivasa. 2013. Legibility and predictability of robot motion. In *2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 301–308.

[8] Sébastien Forestier, Yoan Mollard, and Pierre-Yves Oudeyer. 2017. Intrinsically motivated goal exploration processes with automatic curriculum learning. *arXiv preprint arXiv:1708.02190* (2017).

[9] Batya Friedman and David G Hendry. 2019. *Value sensitive design: Shaping technology with moral imagination*. MIT Press.

[10] Batya Friedman, Peter Kahn, and Alan Borning. 2002. Value sensitive design: Theory and methods. *University of Washington technical report* 2-12 (2002).

[11] Yogesh Girdhar and Gregory Dudek. 2016. Modeling curiosity in a mobile robot for long-term autonomous exploration and monitoring. *Autonomous Robots* 40, 7 (2016), 1267–1278.

[12] Goren Gordon. 2019. Social behaviour as an emergent property of embodied curiosity: a robotics perspective. *Philosophical Transactions of the Royal Society B* 374, 1771 (2019), 20180029.

[13] Goren Gordon, Cynthia Breazeal, and Susan Engel. 2015. Can children catch curiosity from a social robot?. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*. 91–98.

[14] Emily M Grossnickle. 2016. Disentangling curiosity: Dimensionality, definitions, and distinctions from interest in educational contexts. *Educational Psychology Review* 28, 1 (2016), 23–60.

[15] John Harris and Ehud Sharlin. 2011. Exploring the affect of abstract motion in social human-robot interaction. In *2011 Ro-Man*. IEEE, 441–448.

[16] Fritz Heider and Marianne Simmel. 1944. An experimental study of apparent behavior. *The American Journal of Psychology* 57, 2 (1944), 243–259.

[17] Guy Hoffman. 2012. Dumb robots, smart phones: A case study of music listening companionship. In *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*. IEEE, 358–363.

[18] Guy Hoffman and Cynthia Breazeal. 2007. Cost-based anticipatory action selection for human–robot fluency. *IEEE transactions on robotics* 23, 5 (2007), 952–961.

[19] Guy Hoffman and Cynthia Breazeal. 2007. Effects of anticipatory action on human-robot teamwork efficiency, fluency, and perception of team. In *Proceedings of the ACM/IEEE international conference on Human-robot interaction*. 1–8.

[20] Guy Hoffman and Wendy Ju. 2014. Designing robots with movement in mind. *Journal of Human-Robot Interaction* 3, 1 (2014), 91–122.

[21] Guy Hoffman and Gil Weinberg. 2010. Shimon: an interactive improvisational robotic marimba player. In *CHI'10 Extended Abstracts on Human Factors in Computing Systems*. 3097–3102.

[22] Laura Hoffmann and Nicole C Krämer. 2013. Investigating the effects of physical and virtual embodiment in task-oriented and conversational contexts. *International Journal of Human-Computer Studies* 71, 7-8 (2013), 763–774.

[23] Chien-Ming Huang and Bilge Mutlu. 2016. Anticipatory robot control for efficient human-robot collaboration. In *2016 11th ACM/IEEE international conference on human-robot interaction (HRI)*. IEEE, 83–90.

[24] Clayton Hutto and Eric Gilbert. 2014. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 8.

[25] Serena Ivaldi, Natalia Lyubova, Alain Droniou, Damien Gerardeaux-Viret, David Filliat, Vincent Padois, Olivier Sigaud, Pierre-Yves Oudeyer, et al. 2013. Learning to recognize objects through curiosity-driven manipulation with the iCub humanoid robot. In *2013 IEEE Third Joint International Conference on Development and Learning and Epigenetic Robotics (ICDL)*. IEEE, 1–8.

[26] Jamie Jirout and David Klahr. 2012. Children's scientific curiosity: In search of an operational definition of an elusive concept. *Developmental review* 32, 2 (2012), 125–160.

[27] Frederic Kaplan. 2007. In search of the neural circuits of intrinsic motivation. *Frontiers in Neuroscience* 1, 1 (nov 2007), 225–236. https://doi.org/10.3389/neuro.01.1.1.017.2007

[28] Todd B Kashdan, David J Disabato, Fallon R Goodman, and Patrick E McKnight. 2020. The Five-Dimensional Curiosity Scale Revised (5DCR): Briefer subscales while separating overt and covert social curiosity. *Personality and Individual Differences* 157 (2020), 109836.

[29] Todd B Kashdan and Paul J Silvia. 2009. Curiosity and interest: The benefits of thriving on novelty and challenge. *Oxford handbook of positive psychology* 2 (2009), 367–374.

[30] Todd B Kashdan, Melissa C Stiksma, David J Disabato, Patrick E McKnight, John Bekier, Joel Kaji, and Rachel Lazarus. 2018. The five-dimensional curiosity scale: Capturing the bandwidth of curiosity and identifying four unique subgroups of curious people. *Journal of Research in Personality* 73 (2018), 130–149.

[31] Celeste Kidd and Benjamin Y. Hayden. 2015. The Psychology and Neuroscience of Curiosity. *Neuron* 88, 3 (nov 2015), 449–460. https://doi.org/10.1016/j.neuron.2015.09.010

[32] Przemyslaw A Lasota and Julie A Shah. 2015. Analyzing the effects of human-aware motion planning on close-proximity human–robot collaboration. *Human factors* 57, 1 (2015), 21–33.

[33] John Lasseter. 1987. Principles of traditional animation applied to 3D computer animation. In *Proceedings of the 14th annual conference on Computer graphics and interactive techniques*. 35–44.

[34] Edith Law, Vicky Cai, Qi Feng Liu, Sajin Sasy, Joslin Goh, Alex Blidaru, and Dana Kulić. 2017. A Wizard-of-Oz study of curiosity in human-robot interaction. In *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 607–614.

[35] Jordan Litman. 2005. Curiosity and the pleasures of learning: Wanting and liking new information. *Cognition & emotion* 19, 6 (2005), 793–814.

[36] George Loewenstein. 1994. The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin* 116, 1 (1994), 75–98. https://doi.org/10.1037/0033-2909.116.1.75

[37] Wallace H Maw and A Jon Magoon. 1971. The curiosity dimension of fifth-grade children: A factorial discriminant analysis. *Child development* (1971), 2023–2031.

[38] Wallace H Maw and Ethel W Maw. 1961. Establishing criterion groups for evaluating measures of curiosity. *The Journal of Experimental Education* 29, 3 (1961), 299–305.

[39] Wallace H Maw and Ethel W Maw. 1966. Children's curiosity and parental attitudes. *Journal of Marriage and the Family* (1966), 343–345.

[40] Bilge Mutlu and Jodi Forlizzi. 2008. Robots in organizations: the role of workflow, social, and environmental factors in human-robot interaction. In *2008 3rd ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 287–294.

[41] Amal Nanavati, Christoforos I Mavrogiannis, Kevin Weatherwax, Leila Takayama, Maya Cakmak, and Siddhartha S Srinivasa. 2021. Modeling Human Helpfulness with Individual and Contextual Factors for Robot Planning.. In *Robotics: Science and Systems*.

[42] Hung Ngo, Matthew Luciw, Alexander Forster, and Juergen Schmidhuber. 2012. Learning skills from play: artificial curiosity on a katana robot arm. In *The 2012 international joint conference on neural networks (IJCNN)*. IEEE, 1–8.

[43] Behnaz Nojavanasghari, Tadas Baltrusaitis, Charles E Hughes, and Louis-Philippe Morency. 2016. The Future Belongs to the Curious: Towards Automatic Understanding and Recognition of Curiosity in Children.. In *WOCCI*. 16–22.

[44] Pierre-Yves Oudeyer. 2018. Computational theories of curiosity-driven learning. *arXiv preprint arXiv:1802.10546* (2018).

[45] Pierre Yves Oudeyer and Frederic Kaplan. 2009. What Is Intrinsic Motivation? A Typology of Computational Approaches. *Frontiers in Neurorobotics* (2009). https://doi.org/10.3389/neuro.12.006.2007

[46] Pierre-Yves Oudeyer, Frédéric Kaplan, Verena V Hafner, and Andrew Whyte. 2005. The playground experiment: Task-independent development of a curious robot. In *Proceedings of the AAAI Spring Symposium on Developmental Robotics*. Stanford, California, 42–47.

[47] German I Parisi, Ronald Kemker, Jose L Part, Christopher Kanan, and Stefan Wermter. 2019. Continual lifelong learning with neural networks: A review. *Neural Networks* 113 (2019), 54–71.

[48] Hee Sun Park, René Dailey, and Daisy Lemus. 2002. The use of exploratory factor analysis and principal components analysis in communication research. *Human Communication Research* 28, 4 (2002), 562–577.

[49] Maddie Pelz and Celeste Kidd. 2020. The elaboration of exploratory play. *Philosophical Transactions of the Royal Society B* 375, 1803 (2020), 20190503.

[50] Aaron Powers, Sara Kiesler, Susan Fussell, and Cristen Torrey. 2007. Comparing a computer agent with a humanoid robot. In *Proceedings of the ACM/IEEE international conference on Human-robot interaction*. 145–152.

[51] Johnmarshall Reeve. 1993. The face of interest. *Motivation and Emotion* 17, 4 (1993), 353–375.

[52] Johnmarshall Reeve and Glen Nix. 1997. Expressing intrinsic motivation through acts of exploration and facial displays of interest. *Motivation and Emotion* 21, 3 (1997), 237–250.

[53] Tiago Ribeiro, André Pereira, Eugenio Di Tullio, and Ana Paiva. 2016. The sera ecosystem: Socially expressive robotics architecture for autonomous human-robot interaction. In *2016 AAAI Spring Symposium Series*.

[54] Pericle Salvini, Cecilia Laschi, and Paolo Dario. 2010. Design for acceptability: improving robots' coexistence in human society. *International journal of social robotics* 2, 4 (2010), 451–460.

[55] Henk Schmidt and Jerome Rotgans. 2020. Epistemic Curiosity and Situational Interest: Distant Cousins or Identical Twins? *Educational Psychology Review* (2020).

[56] Katie Seaborn, Norihisa P Miyake, Peter Pennefather, and Mihoko Otake-Matsuura. 2021. Voice in Human–Agent Interaction: A Survey. *ACM Computing Surveys (CSUR)* 54, 4 (2021), 1–43.

[57] Stela H Seo, Denise Geiskkovitch, Masayuki Nakane, Corey King, and James E Young. 2015. Poor thing! Would you feel sorry for a simulated robot? A comparison of empathy toward a physical and a simulated robot. In *2015 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 125–132.

[58] Dajung Diane Shin and Sung-il Kim. 2019. Homo curious: Curious or interested? *Educational Psychology Review* (2019), 1–22.

[59] Paul J Silvia. 2008. Interest—The curious emotion. *Current directions in psychological science* 17, 1 (2008), 57–60.

[60] Paul J Silvia. 2012. Curiosity and motivation. *The Oxford handbook of human motivation* (2012), 157–166.

[61] Bharat Singh, Rajesh Kumar, and Vinay Pratap Singh. 2022. Reinforcement learning in robotic applications: a comprehensive survey. *Artificial Intelligence Review* 55, 2 (2022), 945–990.

[62] James P Stevens. 2012. *Applied multivariate statistics for the social sciences*. Routledge.

[63] Daniel Szafir, Bilge Mutlu, and Terrence Fong. 2015. Communicating directionality in flying robots. In *2015 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 19–26.

[64] Leila Takayama, Doug Dooley, and Wendy Ju. 2011. Expressing thought: improving robot readability with animation principles. In *Proceedings of the 6th international conference on Human-robot interaction*. 69–76.

[65] Sebastian Thrun and Tom M Mitchell. 1995. Lifelong robot learning. *Robotics and autonomous systems* 15, 1-2 (1995), 25–46.

[66] Emre Ugur, Mehmet R Dogar, Maya Cakmak, and Erol Sahin. 2007. Curiosity-driven learning of traversability affordance on a mobile robot. In *2007 IEEE 6th International Conference on Development and Learning*. IEEE, 13–18.

[67] Gentiane Venture and Dana Kulić. 2019. Robot expressive motions: a survey of generation and evaluation methods. *ACM Transactions on Human-Robot Interaction (THRI)* 8, 4 (2019), 1–17.

[68] Nick Walker, Kevin Weatherwax, Julian Allchin, Leila Takayama, and Maya Cakmak. 2020. Human Perceptions of a Curious Robot that Performs Off-Task Actions. In *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*. 529–538.

[69] Sarah N Woods, Michael L Walters, Kheng Lee Koay, and Kerstin Dautenhahn. 2006. Methodological issues in HRI: A comparison of live and video-based methods in robot to human approach direction trials. In *ROMAN 2006-the 15th IEEE international symposium on robot and human interactive communication*. IEEE, 51–58.

Component	Thinking Agent	Working Agent	Social Agent
Items (In order)	Explores Surroundings, Desires New Knowledge, Learns On Own, Independent, Curious	Focused*, Competent*, Obedient*, Unobtrusive*, Acceptable, Follows Instructions	Trustworthy, Likable, Responsible, Approachable
Loading Level 1	0.89, 0.91, 0.71, 0.57, 0.86	0.76, 0.74, 0.87, 0.80, 0.75, 0.80	0.81, 0.73, 0.72, 0.69
Loading Level 2	0.83, 0.85, 0.61, 0.66, 0.76	0.64, 0.71, 0.80, 0.66, 0.74, 0.78	0.81, 0.85, 0.84, 0.76
Loading Level 3	0.68, 0.88, 0.52, 0.71, 0.83	0.55, 0.81, 0.88, 0.65, 0.86, 0.90	0.94, 0.90, 0.91, 0.88
Loading Level 4	0.50, 0.86, 0.68, 0.79, 9.86	0.64, 0.81, 0.93, 0.51, 0.83, 0.88	0.91, 0.89, 0.91, 0.74
Reliability	α = 0.85, 0.80, 0.77, 0.79	α = 0.82, 0.78, 0.86, 0.85	α = 0.70, 0.83, 0.92, 0.89

Table 1: Principal Component Analysis.



Figure 5: Access at: https://osf.io/47w36/ or click on the image above to view the video stimuli on OSF.

.