Expressive Cognitive Architecture for a Curious Social Robot

MAOR ROSENBERG, Curiosity Lab, Industrial Engineering Department, Tel-Aviv University, Israel
HAE WON PARK, Personal Robots Group, MIT Media Lab, United States
RINAT ROSENBERG-KIMA, Faculty of Education in Science and Technology, Technion, Israel
SAFINAH ALI, ANASTASIA K. OSTROWSKI, and CYNTHIA BREAZEAL, Personal Robots Group, MIT Media Lab, United States
GOREN GORDON, Curiosity Lab, Industrial Engineering Department, Tel-Aviv University, Israel

Artificial curiosity, based on developmental psychology concepts wherein an agent attempts to maximize its learning progress, has gained much attention in recent years. Similarly, social robots are slowly integrating into our daily lives, in schools, factories, and in our homes. In this contribution, we integrate recent advances in artificial curiosity and social robots into a single expressive cognitive architecture. It is composed of artificial curiosity and social expressivity modules and their unique link, i.e., the robot verbally and non-verbally communicates its internally estimated learning progress, or learnability, to its human companion. We implemented this architecture in an interaction where a fully autonomous robot took turns with a child trying to select and solve tangram puzzles on a tablet. During the curious robot's turn, it selected its estimated most learnable tangram to play, communicated its selection to the child, and then attempted at solving it. We validated the implemented architecture and showed that the robot learned, estimated its learnability, and improved when its selection was based on its learnability estimation. Moreover, we ran a comparison study between curious and non-curious robots, and showed that the robot's curiosity-based behavior influenced the child's selections. Based on the artificial curiosity module of the robot, we have formulated an equation that estimates each child's moment-by-moment curiosity based on their selections. This analysis revealed an overall significant decrease in estimated curiosity during the interaction. However, this drop in estimated curiosity was significantly larger with the non-curious robot, compared to the curious one. These results suggest that the new architecture is a promising new approach to integrate state-of-the-art curiosity-based algorithms to the growing field of social robots.

CCS Concepts: • **Human-centered computing → Collaborative and social computing**; • **Computing methodologies → Cognitive robotics**;

Additional Key Words and Phrases: Artificial curiosity, child-robot interaction, social robots

12

The reviewing of this article was managed by associate editor Catherine Pelachaud.
This research was supported by the National Institutes of Health, under grant 5R01HD086899-02.
Authors' addresses: M. Rosenberg and G. Gordon, Curiosity Lab, Industrial Engineering Department, Tel-Aviv University, 50 Haim Levanon street, Tel-Aviv 6997801, Israel; emails: maor.rosenberg@cs.tau.ac.il, goren@gorengordon.com; H. W. Park, S. Ali, A. K. Ostrowski, and C. Breazeal, Personal Robots Group, MIT Media Lab, 75 Amherst Street, Cambridge, Massachusetts; emails: haewon@media.mit.edu, safinah@mit.edu, {akostrow, cynthiab}@media.mit.edu; R. Rosenberg-Kima, Faculty of Education in Science and Technology, Technion city, Haifa 3200003, Israel; email: rinatros@technion.ac.il.

2160-6455/2021/06-ART12 $15.00
https://doi.org/10.1145/3451531

ACM Reference format:
Maor Rosenberg, Hae Won Park, Rinat Rosenberg-Kima, Safinah Ali, Anastasia K. Ostrowski, Cynthia Breazeal, and Goren Gordon. 2021. Expressive Cognitive Architecture for a Curious Social Robot. *ACM Trans. Interact. Intell. Syst.* 11, 2, Article 12 (June 2021), 25 pages.
https://doi.org/10.1145/3451531

1 INTRODUCTION

Cognitive architectures attempt to capture important aspects of human cognition in generalized conceptual and computational models [39, 58]. While many architectures focus on perception, attention, memory, and learning, few have focused on social interaction with other agents [59].

Additionally, the field of artificial curiosity has gained much attention in recent years [23, 29, 37, 40, 50]. Artificial curiosity, also called intrinsic motivation models, is based on the premise that the agent can quantify how much it can learn from specific actions, and then selects those actions that maximize learnability [49, 66]. The field has become mainstream in the wake of deep learning, where numerous models incorporate intrinsic motivation modules to add intrinsic rewards in an otherwise sparse reward space [1, 29, 37]. Nevertheless, most artificial curiosity studies focus on the game or robot itself, while only few have incorporated social interaction into their scenarios [24, 55].

In this article, we present a novel cognitive architecture that incorporates an **Artificial Curiosity** component and a **Social Expressivity** component (see Figure 1). For the robot to externalize its curiosity states, i.e., help users interacting with the robot perceive and understand the rationale behind its curiosity-driven choices, the variables of the artificial curiosity module need to be socially expressed in human terms. The **Embodied Curiosity** sub-component within the **Social Expressivity** component translates artificial curiosity variables to verbal and non-verbal social expressions. This way, the robot is both curious, i.e., selects tasks that maximize learnability, and social, i.e., expresses its intrinsic motivation to other social agents (users).

We implement the expressive cognitive architecture in a social robot for children, that plays a tangram puzzle game on a tablet with a child. The robot and child take turns in selecting which puzzle to try and solve next, and then attempt to solve it. The robot implements an artificial curiosity component that estimates the learnability of each puzzle in the selection and expresses its estimation in verbal and non-verbal communication to the child, e.g., "I chose this puzzle because I can learn the most from it," with an enthusiastic smile. We validate the curiosity module and show that the robot learns, estimates the learnability, and improves when its selection is based on its learnability estimation. We then report on a user study with 92 children who played with either a curious or non-curious robot. We show that children who played with the curious robot chose significantly more unknown and challenging puzzles to solve and had a significantly smaller decrease in estimated curiosity.

The contributions of this article are as follows: (i) a novel expressive cognitive architecture that links intrinsic motivation to social expressivity; (ii) a validation scheme of an artificial curiosity module; (iii) a user study with a fully autonomous curious social robot that influences children's behaviors; and (iv) estimation of children's curiosity based on the same cognitive architecture.

The structure of the article is as follows: we first present related work, then introduce the general architecture, followed by the specific implementation and description of the user study. We then present the validation and user study analysis, followed by discussion and future work.

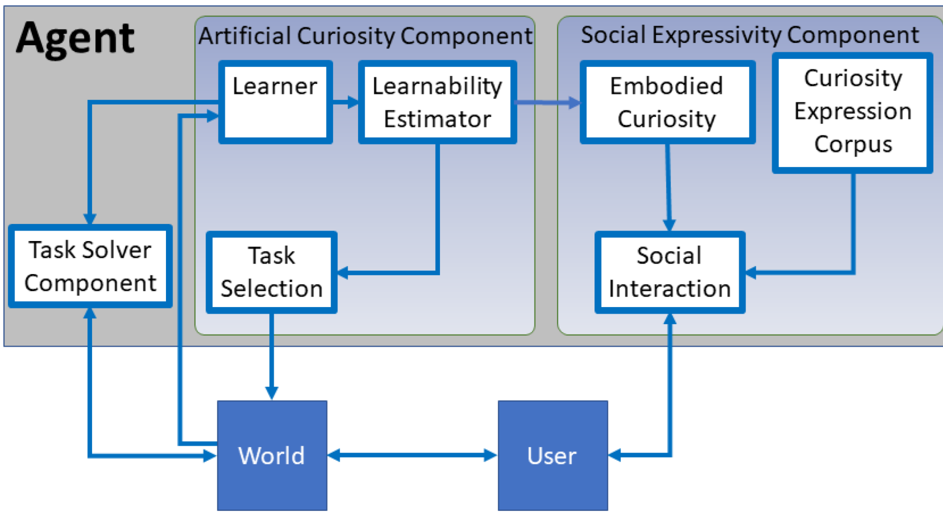


Fig. 1. Expressive cognitive architecture. The Artificial Curiosity component implements algorithms that select tasks based on their evaluated learnability. The Social Expressivity includes a Curiosity Expressions Corpus, e.g., general enthusiasm of learning, and externalizes intrinsic curiosity variables via the Embodied Curiosity sub-component. The Task Solver component solves the selected task, with prior knowledge accumulated by the Learner sub-component.

2 RELATED WORK

2.1 Cognitive Architectures

Cognitive architectures have been used for over four decades to try and capture human cognitive aspects and formalize them in conceptual and computational models [58]. Each architecture focuses on specific aspects, such as memory, perception, and learning, while others try to integrate as many aspects as possible [39].

We focused on human-robot interaction-oriented architecture, with emphasis on learning, action selection, social expressivity, and curiosity. Several other architectures also addressed these issues. The ACT-R/E [68] embodied cognitive architecture for human-robot interaction modeled many aspects of cognition and action and was evaluated on several human study-driven data. The goal of the architecture was to facilitate a better development of human-robot interaction. However, all rewards in the model were extrinsic and thus could not account for curiosity. The **CAIO** [2] (Cognitive and Affective Interaction-Oriented) architecture for social human-robot interactions was aimed to allow robots to reason on people's mental states and to act physically, emotionally, and verbally. The architecture did not include a learning mechanism and thus could not account for curiosity-driven learning and behavior.

2.2 Curious Robots

The field of artificial curiosity has started more than two decades ago [5, 25, 63], and has been gaining traction in the last five years [29, 38, 40, 48], with many applications, especially in the field of deep learning [30, 53, 65]. It is based on the premise that rewarding learning progress of one component, e.g., a deep neural network (DNN), can improve the performance of another component, e.g. a, reinforcement-learning DNN. However, most of the recent uses were within the confines of games, e.g., [15, 34], and non-social robots, e.g., [14, 15, 30], with several exceptions that linked artificial curiosity to the emergence of social behavior [24] and to interaction with humans [55].

Curious robots have been used in the context of human-robot interaction [11, 26, 55], where pre-programmed curiosity-based behaviors were shown to promote children's curiosity [26]. A robot that asks questions was shown to promote curiosity-based behaviors in adults [11]. An intrinsically rewarded social robot learned how to behave in social contexts via interaction with humans in real unstructured environments [55]. However, these studies have not directly connected intrinsic curiosity to social behavior. Here, we directly link intrinsic curiosity variables to externalized social behaviors.

2.3 Social Robot Expressivity

Since its beginning, social robotics has focused on expressions and non-verbal behaviors as an important channel of information and interaction with humans [8–10]. Social robots can express human-like non-verbal behavior or more robot-like behaviors, e.g., changing eye-colors [44, 56], where the former has been shown to be more effective.

One major aspect of non-verbal behaviors are facial expressions. Recent work has shown that reverse-engineering human facial expressions into social robots is a promising approach [12], especially for culturally sensitive scenarios [13]. Facial expressions and body language can generate more trust in robots [18, 51] and more engaging experience with robots [45].

Gestures are important non-verbal behaviors [17, 57]. It has been shown that congruent speech-gesture combination is perceived as more anthropomorphic and likeable [57], and using gestures to augment language learning has been shown to be effective [17].

In an educational setting, there has been a growing body of research that shows social robots' expressivity can impact learning gains [36]. For example, a robot performing contingent back-channeling was perceived as more attentive and children preferred to tell stories to such a robot [28].

However, previous studies have pre-designed expressivity, whereas in this contribution, we link both verbal and non-verbal expressivity to intrinsic variables.

2.4 Curiosity Assessment

Furthermore, the field of model-based curiosity assessment has started to gain attention. However, curiosity is a multi-faceted characteristic. In a pioneering work, Jirout and Klahr [33] have developed a tablet game that assesses uncertainty seeking behavior in children, based on Loewenstein's [43] information-gap theory. Another tablet game has been developed to assess free exploration behavior [26], which has later been adapted for adults [64], and has been shown to correlate with self-report of curiosity.

A graph-theory-based tablet game has recently been introduced that assesses question-asking behavior and has shown that curious children, as assessed externally by their teachers, ask *different types* of questions, than less curious children [67]. Finally, social robots have been used to assess adults' physical curiosity, using artificial curiosity algorithms [22]. Here, we develop a new quantitative model-based assessment tool for children's curiosity, that is inspired by a curious robot's algorithm.

2.5 Conclusions

In contrast to previous studies and architectures, in this contribution we take a holistic approach to curious social robots' cognitive architecture. Each of the aforementioned studies have addressed one or two aspects, whereas we combine them in a single architecture. Thus, we combine (i) artificial curiosity algorithms [63] in both the robot's behavior [25] and children's curiosity assessment [22]; (ii) we couple the robot's intrinsic models [39] to its social verbal and non-vebal behaviors [8]; and (iii) we address both **human-robot interaction** (**HRI**)–related aspects [2] and educational ones [26].

3 ARCHITECTURE

The cognitive architecture introduced here is composed of several components (see Figure 1). The **Artificial Curiosity** component is composed of algorithms that govern the robot's task selection based on the maximization of learnability. The **Social Expressivity** component interprets the variables of the **Artificial Curiosity** component and generates a set of verbal and non-verbal expressions to help the robot externalize its curiosity states. The **Embodied Curiosity** sub-component externalizes the intrinsic curiosity to the physical embodied attributes of the robot [24]. The **Task Solver** component is composed of algorithms that enable the robot to solve the selected task. This component also receives input from the **Artificial Curiosity** component representing prior knowledge that could help in solving the task. The task is represented within the **World** component, which includes the space that the robot and the user share during the interaction, which is a tablet app in our specific implementation.

Thus, the architecture integrates two previously reported components, namely, artificial curiosity [50, 63] and general social curiosity [11, 26], and adds a novel one, namely, an **Embodied Curiosity** sub-component, that externalizes intrinsic variables, so that people socially interacting with the robot can perceive its inner curiosity state.

3.1 Artificial Curiosity Component

Artificial curiosity relates to task selection that aims to maximize learnability, i.e., a scalar that represents potential learning progress [50]. In other words, learnability, or *Learning Potential*, quantifies how much can be learned, if a specific choice is made [66]. From this definition, there are several attributes that must be present: (i) **Learner** sub-component: the robot attempts to learn sensorimotor representations of the world, e.g., correlations between its sensor input and actuator output, such as forward models. The **Learner** can take any form, e.g., supervised, unsupervised, or reinforcement learning. In the tangram puzzle task, the Learner took the form of a supervised learning neural network that attempts to predict the solution of the tangram based on its silhouette. (ii) **Learnability Estimator** sub-component: given a set of task options to choose from, it represents a scalar value of an estimation of how much each selection will improve the **Learner**'s ability to learn from the information that will be supplied, given the task selected. Therefore, learnability is inherently related to novelty and exploration, i.e., selecting a task that will give maximal information. In the tangram puzzle task, learnability took the form of average entropy of the Learner's predicted solution. (iii) **Task Selection** sub-component: based on the **Learnability Estimator**, the curious robot selects the task option with the highest learnability. The **Learner** component is also connected to the **Task Solver** component where prior knowledge is used when solving a new task.

To summarize, in the context of our work, learning has been used primarily in the sense of embedding past knowledge in order to both assess how much we can learn from new stimuli and to increase the performance of solving problems similar to problems we have already solved. Specifically, the **Learner** learns pairs of puzzle silhouettes and solutions, and the **Learnability Estimator** uses this for estimating how much can be learned from new puzzles. The **Task Solver** also uses the **Learner** information in order to increase the performance of solving new puzzles measured here by the decrease in number of moves to solution.

A detailed description of each subcomponent in the tangram game is presented in Section 4.2.

For an implementation of the **Artificial Curiosity** component, one must be able to validate the following criteria: (a) Does the **Learner** learn? Does more information from the environment improve the **Learner**? (b) Does the **Learnability Estimator**'s output correlate to actual learning progress, e.g., decrease with repeated learning? (c) Does selection according to the **Learnability Estimator** lead to better learning?

Table 1. Example Utterances the Robots Make to Express Their Curiosity Intent

Condition Context	Non-curious	Curious
Question asking	No question asking	"I am so curious if there are any robots like me in the outer space."
Child losses	"Better luck next time."	"You must have learned new things."
Child wins	"You solved the puzzle."	"Good job trying something new and different."

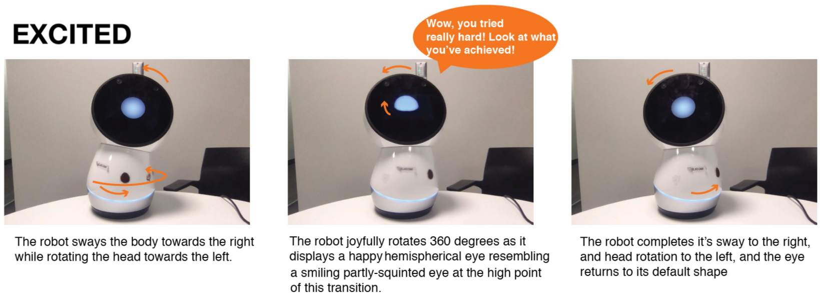
3.2 Social Expressivity Component

In psychology literature [42, 46], there is an important distinction between state curiosity and trait curiosity. Trait curiosity refers to the stable component of the person's personality, e.g., "I am a curious person." In contrast, state curiosity refers to the moment-by-moment changes in curiosity, e.g., "I am curious now about this thing." In general, a high trait-curiosity person experiences more high state-curiosity episodes.

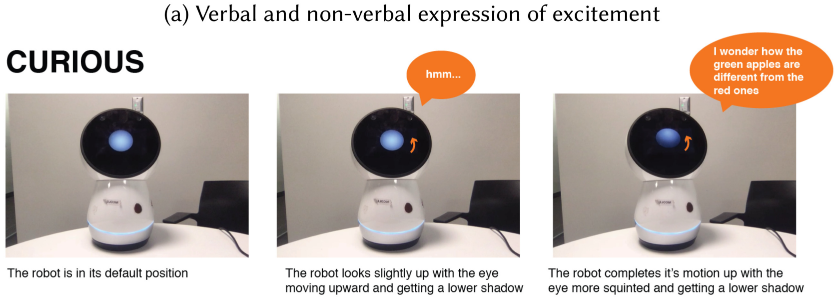
The expressive cognitive architecture implements this distinction in the Social Expressivity component. The **Curiosity Expression Corpus** sub-component represents *trait curiosity* and includes pre-coded behaviors that convey curiosity-related verbal and non-verbal expressions [11, 26]. For example, enthusiasm from learning, with no correlation to the amount of learning, is part of this component, e.g., "I love to learn" followed by a happy face. Another important behavior is rhetorical question asking that is unrelated to actual learning done by the artificial curiosity component, e.g., "I wonder where the sea shells come from?" Finally, this component is responsible for reactions to users' behaviors. Thus, if the user performs an action, the robot will react with a socially appropriate behavior, e.g., "Good job, you learned a lot." However, the behaviors are unrelated to the moment-by-moment curiosity of the robot. For more examples, see Table 1 and Figure 2.

3.2.1 Validation of Robot's Expressions. The robot's expressions were validated using a crowdsourcing platform, **Amazon Mechanical Turk** (**MTurk**). Ethical approval was obtained from Massachusetts Institute of Technology's institutional review board. Participants' parents were recruited through MTurk under the restrictions that they were current residents of the United States (our task was only visible to US workers), had at least a 95% task approval rate, which demonstrates qualification for high data quality [54], and were parents of at least one child between the ages of 5 and 8. We obtained parental consent to enroll their children to participate in our robot expression validation study, and provided instructions to not provide any input to their children's response. We benchmarked prior work that recruited under-age participants online through their parents via MTurk [69].

The task was created and posted on MTurk under the title "Emotive motion recognition" with 0.10 USD base compensation and up to 6.00 USD bonus. The first set of screener questionnaires surveyed the workers' general demographics including average number of hours they use their computer, if the device they are currently using had a camera, how many adults they live with, and if they have one or more children (either biological or non-biological), and if so, how many in each age group (age 0–4, 5–8, 9–18). Except for the question that asked about their 5–8 year old children and the camera, the rest were filler questions so that workers would not notice which



(a) Verbal and non-verbal expression of excitement



(b) Verbal and non-verbal expression of curiosity

Fig. 2. Examples of robot's verbal and nonverbal expressivity. The robot's expressions were validated via a crowdsourcing platform with 27 children aged between 5–8.

were the qualification questions. Such screener for selecting parent workers was also used in prior work [62]. For those who indicated they have at least one child in the target age group and have a camera on their current device, we offered 6.00 USD bonus if they enrolled their child in the next set of questions which involved watching robot videos and answering what emotion it was expressing. We also disclosed that the session will be video recorded and that the parent should be with the child the entire time and help enter the answers but not answer for them. Afterwards, they were asked to review and sign a parental consent form. Each worker ID could only take the initial screener survey once. We unpublished the task after we collected 40 responses. The recorded videos were uploaded to our secure server with only participant IDs (not worker IDs) attached to the videos.

After a short question asking about the child's age and gender, each child saw 12 short robot videos, and selected one of the following options to indicate which expression she perceived from the robot's behavior: "The robot is ... (1) happy (excited), (2) sad, (3) curious, (4) surprised, (5) angry, (6) neutral (indifferent), (7) none of the above (I don't know)." The text answer options were also coupled with pictorial emojis. The questions consisted of three happy, two surprise, two sad, three curious, and two neutral expressions. There were no angry expressions. The videos in each category were randomly chosen from a corpus of robot's expressions that were used in the main study. Children who participated in this test did not partake in the main study.

After reviewing the videos, we excluded six in which parents influenced their child's answer at least once. Three were excluded because no audio was recorded and we could not determine how much the parent was influencing the child's answers. Two were excluded because the video did not shoot the child properly, and another two were excluded because of uncertain age of the child. In total, 27 children answered all 12 questions (age 6.86±0.71, female=11) without visible help from their parent. Table 2 shows the result of how well children's perceived robot expressions matched

Table 2. Intended Robot Expressions and Children's Perceived Expressions

		Perceived Expression (%)						
		happy (excited)	sad	curious	surprise	angry	neutral	N/A
Intended Expression	happy (excited)	**88.89**	0.00	2.47	4.94	0.00	3.70	0.00
	sad	0.00	**85.19**	0.00	0.00	7.41	7.41	0.00
	curious	2.47	0.00	**80.25**	12.35	0.00	4.94	0.00
	surprise	20.37	0.00	11.11	**66.67**	0.00	0.00	1.85
	neutral	**31.48**	22.22	11.11	0.00	0.00	27.78	7.41

the designed expressions. We can see that children perceived robot's expressions as we intended, especially for happy, sad, and curious expressions. Because the surprise examples were all associated with positive valence, e.g., "Wow, so many seashells on the beach!" while the robot spins around with excitement, we can see an expected split between happy and surprise. For neutral samples, e.g., "I'm going to select a puzzle," "It's your turn," children often associated them with happy and sad emotions.

We believe that neutral expressions could be perceived incorrectly because children had a hard time understanding what the terms *neutral* or *indifferent* meant in the survey. Children would often ask their parent what neutral means, and some parents helped explain it, some did not. The reason that children often associated neutral expressions as either happy or sad could be because happy and sad are the expressions that our target age children can understand better than other emotions. Lawrence et al. [41] studied how the ability to recognize different emotions develops at different ages. The authors found that children 6–7 are good at recognizing expressions of happiness, sadness, and anger, and children 8–9 years old can also recognize more complex expressions such as surprise. These findings help explain our results as well. Children were better at recognizing the intended happy and sad emotions, as compared to the surprise emotions. Because of the robot's default bright tone of voice, more children could have attributed happy emotions to neutral expressions. However, it was clear that, compared to happy and sad expressions, children pondered much more when answering the neutral expressions (in seconds, happy: 1.02 ± 0.32, neutral: 1.98 ± 0.80; $t(133) = 9.71$, $p < 0.0001$ and sad: 1.47 ± 0.62, neutral: 1.98 ± 0.80; $t(106) = 3.70$, $p < 0.001$). This observation indicates that children were able to perceive happy and sad emotions well, and did notice the difference from neutral expressions.

The validation study was conducted post the main study and could not benefit the design of expressions of the robot. In future work, we will improve the validation study design with easier to understand vocabulary, and use the results to refine the design of agent expressions.

3.2.2 Embodied Curiosity Sub-Component. The **Embodied Curiosity** sub-component represents *state curiosity* and is the main contribution of the presented architecture. In this sub-component, learnability optimization results from the Artificial Curiosity component are externalized as social expressions. More specifically, the robot externalizes its selection process. When presented with task choices, the robot not only selects the option with the highest estimated learnability, but also selects verbal and non-verbal behaviors to explain its selection, e.g., "I will choose this option because I can learn the most out of it" in an enthusiastic way. This sub-component contributes to developing an autonomous curious social robot where the robot embodies curiosity-based algorithms, i.e., selects a task with optimal learnability and communicates its selections to the user.

3.2.3 *Social Interaction Sub-Component.*

The **Social Interaction** sub-component is the actual interface the agent has with the user and is responsible for the integration of the two other sub-components according to the current context. In other words, given the current context, e.g., user responses, this sub-component selects from the **Embodied Curiosity** and **Curiosity Expression Corpus** sub-components, which social actions (verbal and non-verbal) to execute.

4 TANGRAM CHILD-ROBOT INTERACTION

In this section, we describe how the expressive cognitive architecture for the tangram puzzle interaction was designed to enable curious social robot interaction with young children.

4.1 World: Tangram Puzzle Game

The game was designed as a mobile app presented on a tablet. The game's main objective is to collect sea shells locked in treasure boxes that can be opened by solving a tangram puzzle. A puzzle is solved by fitting the given silhouette with some or all of the seven standard tangram pieces. The session began with a short introduction by the robot, and consisted of six rounds for each player, starting with the robot (see Figure 3(f)). In each round, the player is presented with three treasure boxes. The puzzles' silhouettes appear on the left and middle treasure boxes. On the right box, however, there is a question mark that we refer to as the unknown treasure box (see Figure 3(a)). Solving a puzzle is rewarded with a shell. Each round had a time limit indicated by an hourglass which was 2 minutes for the child and half a minute for the robot. At the end, the robot celebrates all the shells gathered. Note that the game is neither explicitly competitive, since there is no winner or loser, nor explicitly collaborative, since each player solves the puzzle by himself. It is rather a shared experience, where each player observes the other alternately.

We created for each player two curricula series of seven tangrams, starting with a one-piece puzzle, and adding one piece at a time up to a seven-piece puzzle. We will refer to them as "seen curriculi" and "unknown curriculi" (see Figure 3(b) and (c)). In the first round, the first two puzzles in the seen curriculum are depicted on the left and middle boxes, respectively (see Figure 3(a), left). If the player chooses the middle box, in the next round she will get the second and third puzzles in the seen curriculum on the left and middle boxes, respectively, and so on (see Figure 3(a), right, and (d)). If the player chooses the left (easier) puzzle, then in the next round she will get the same seen puzzles again on the left and middle boxes, hence not progressing in learning to solve the complex puzzle (see Figure 3(e)). If in the next round, she chooses the left box again, the seen puzzles do advance so she does not get stuck in the same "level" (see Figure 3(e)). If the player chooses the unknown treasure box, the k-th puzzle in the unknown curriculum series appears, where k is the round number. In other words, the puzzle in the unknown box advances regardless of the selection (see Figure 3(d) and (e)).

The choice of each box corresponds to different aspects of curiosity:

- Choosing the left treasure box corresponds to receiving mostly only external reward (a shell once the puzzle is solved) with little intrinsic reward since the puzzle is the easiest to solve.
- Choosing the middle treasure box corresponds to an effective curious action, since a more advanced, novel puzzle will be learned as the game advances. As before, once the puzzle is solved, the player is rewarded with a shell.
- Choosing the unknown treasure box corresponds to an uncertainty seeking action, where the player is expressing a curiosity preference for exploring the unknown where learning is less structured and incremental. Once the puzzle is solved, a shell is rewarded.

The child and the robot are given different curricula series, to prevent children learning the solution from the robot. We note that in rounds four and five of the child, we used different puzzles

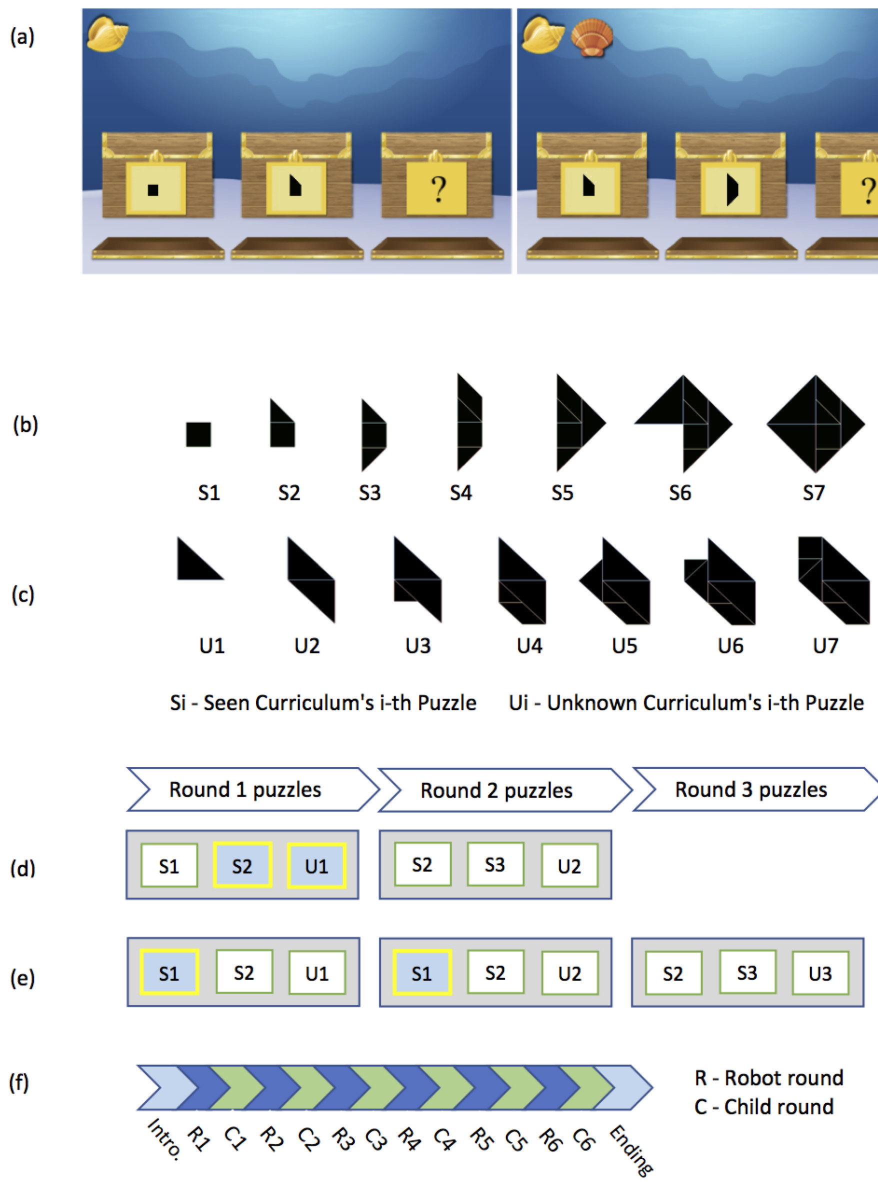


Fig. 3. (a) Selection room setup. Left: first robot round. Right: second robot round after selection of the middle puzzle in the first round. (b),(c) An example curricula series of tangram puzzles. (b) Si is the *i*-th puzzle of the robot's Seen curriculum (left and middle boxes). (c) Ui is the *i*-th puzzle of the robot's Unknown curriculum (right box). (d),(e) Puzzle progression. (d) When the player selects the middle or right box, then on the next round all puzzles advance. (e) When the player selects the left box, then on the next round the seen puzzles do not advance. When selecting left twice consecutively, then the seen puzzles do advance. The unknown puzzle advances in each case. (f) Game flow. Each player plays six rounds in turns.

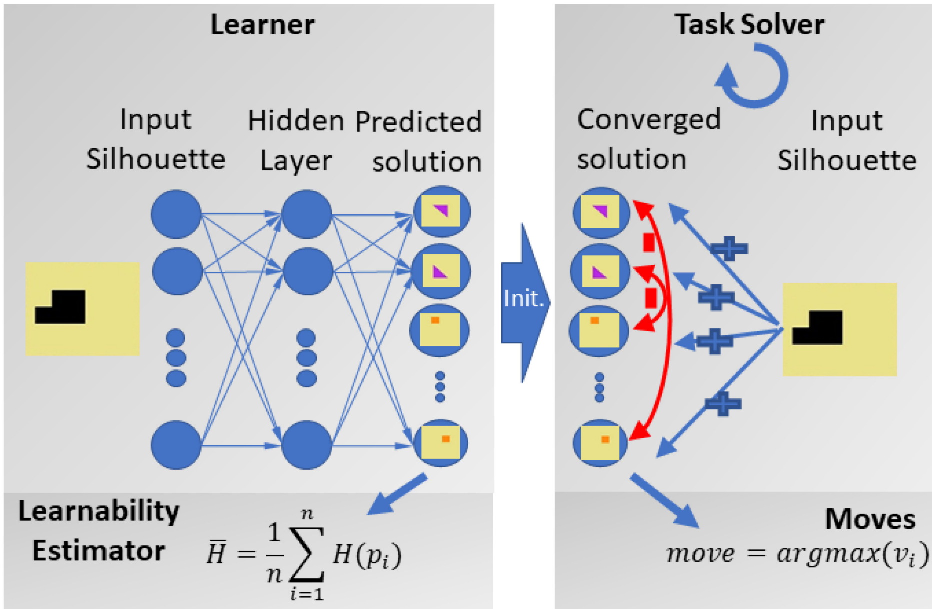


Fig. 4. Tangram Learner, Task Solver, and Learnability Estimator. The input is the silhouette of the tangram puzzle represented by binary nodes. The output is composed of nodes, each representing a single tangram piece in a single location and orientation. The Learner is a feedforward neural network with a single hidden layer the same size as the output layer, both with sigmoid activation functions. The Task Solver is a Boltzmann machine, with positive normalized edges from the silhouette to the nodes, and negative normalized edges (inhibitions) from identical-piece and overlapping-pieces nodes. The Learner prediction serves as the initialization for the Boltzmann machine Solver. The Learnability Estimator outputs the average entropy of the output layer of the Learner. The appropriate moves on the tablet are selected based on the highest-value output node at each iteration.

in order to learn more about the child's mindset [19, 52] to be reported in a future publication. Regardless, the appearance of these puzzles was identical for both conditions and does not affect the results presented here.

4.2 Artificial Curiosity Component in Tangram Game

4.2.1 Learner: Tangram Predictor. After each round, the robot tries to encode its experience. This is done with a feed-forward neural network with one hidden layer (see Figure 4). The input layer represents the two-dimensional silhouette of the tangram puzzle with $17 \times 17 = 289$ binary nodes. The output layer represents the solution of the tangram where each node corresponds to a specific piece on the board in a specific position and in a specific orientation. The number of output nodes is 312, the hidden layer size was chosen to also be 312, and both have a sigmoid activation function. At the end of each round, the neural network is trained with the silhouette of the puzzle as input and the solution as labeled output. More precisely, all the silhouettes and solutions of the previous rounds are set to be a mini batch, and the network is trained using backpropagation with 100,000 epochs.

It is worthy to note that in contrast to more common deep learning architectures, which require vast amounts of training data, the information available for the **Learner** is extremely sparse, since it learns *only* from the games it has played with the child. Thus, we implemented a simple network architecture, with long training times to enable convergence of the learning process *after each game*. However, this is but a single implementation of a **Learner**, tailored for the specific interaction we have studied. Other interactions may supply other types of information and require more elaborate networks for the **Learner** to learn.

4.2.2 Learnability Estimator: Average Entropy. Here, we define the average entropy of the output layer of the **Learner** as the estimation of learnability of a specific new tangram puzzle [7] (see Figure 4). It is computed according to the following formula: $\overline{H} = (1/n) \sum_{i=1}^{n} H(p_i)$ where p_i is the value of the i-th output node and $H(p_i)$ is the entropy of a geometric probability variable with probability p_i.

For example, if the values of the output layer are only ones or zeroes, then the level of assurance is very high and indeed the average entropy will be zero, which means that we estimate that nothing can be learned from the puzzle. Higher average entropy means there is uncertainty about the solution of the presented silhouette, thus if the **Learner** will be presented with the solution to this puzzle, it will learn the most.

However, it is important to note that this measure can range between 0 and log(2). Due to the specific curriculum used in the setup, i.e., an increase in a single tangram piece, once a puzzle solution is learned, the uncertainty in the next presented puzzles will be in the order of *a single output node*. In other words, the range of estimated learnability will be between 0 and $\log(2)/312 \approx 0.001$.

4.2.3 Task Selection: Most Learnable Tangram. (Table 3) The **Task Selection** component receives the estimated learnability of all options and chooses the most learnable one. In the current scenario, the silhouettes of the left and middle boxes are set as input to the Learner's neural network and their estimated learnability is computed, $\overline{H}_{\text{left}}$ and $\overline{H}_{\text{middle}}$, respectively. However, there is also an unknown (right) tangram. We thus define a learnability threshold, H_{th}^{C}, for the curious robot, such that if both known tangrams', i.e., left and middle, estimated learnabilities are less than this threshold, meaning they are too "easy," the robot selects the unknown tangram; otherwise, it selects the one with the highest learnability. For the non-curious robot, if both known tangrams' estimated learnabilities are above the threshold H_{th}^{NC}, meaning they are too "hard," it selects the unknown one. H_{th}^{C} and H_{th}^{NC} are initialized to 0.1 and 0.5, respectively, and are updated according to the following schedule after each round in order to achieve better scaling to current average entropy values: $H_{\text{th}}^{C} \leftarrow max(\overline{H}_{\text{left}}, \overline{H}_{\text{middle}}) \times 0.8$ and $H_{\text{th}}^{NC} \leftarrow min(\overline{H}_{\text{left}}, \overline{H}_{\text{middle}}) \times 1.2$.

4.3 Task Solver Component: Tangram Solver

The robot's algorithm solves each tangram puzzle using a Boltzmann machine relaxation search algorithm [47]. The Boltzmann machine is composed of nodes and edges (see Figure 4), where each node corresponds to a specific tangram piece in a specific location on the board and with a specific orientation (same as the **Learner**'s output layer). Each edge between two nodes has a negative value (inhibition) if the corresponding pieces overlap or are the same piece. Each assignment of the nodes with Boolean values corresponds to a collection of pieces on the board. On the correct assignment, there will be no inhibition since there are no overlapping pieces, nor the use of the same piece more than once.

The nodes are excited according to the area of overlap between the silhouette of the puzzle and the corresponding piece of each node. At each iteration of the algorithm, the nodes' values v_i are updated based on their input: if excitation is smaller than 1, the node is set to 0; else if inhibition is larger than -0.01, the node is set to 1; otherwise, the value is set to 1 with a temperature-dependent

Table 3. Cognitive Expressive Architecture Implementation

Condition	Learnability Estimator	Intrinsic Curiosity State	Task Selection	Embodied Curiosity
Curious	$\overline{H}_{\text{left}}, \overline{H}_{\text{middle}}$	if $max(\overline{H}_{\text{left}}, \overline{H}_{\text{middle}}) < H_{\text{th}}^{C}$	right/?	"I know how to solve the other ones so I will try the mystery box." Happy face.
		else if $\overline{H}_{\text{left}} > \overline{H}_{\text{middle}}$	left	"This one is easy, but I'm curious if there are other ways to solve this puzzle." Curious face.
		else	middle	"I want to try a new puzzle, so I will select the box in the middle." Happy face.
Non-Curious	$\overline{H}_{\text{left}}, \overline{H}_{\text{middle}}$	if $min(\overline{H}_{\text{left}}, \overline{H}_{\text{middle}}) > H_{\text{th}}^{NC}$	right/?	"I should know how to solve it. This time, I will choose this one." Nod.
		else if $\overline{H}_{\text{left}} > \overline{H}_{\text{middle}}$	middle	"I think I know how to solve the next level, so I will choose this middle box." Nod.
		else	left	"I want to win easily, so I will choose what I know I can solve, like this one!" Nod.

Boltzmann probability, where the temperature starts high ($T = 1$) and cools down exponentially at a rate of 0.99; else it is set to 0 [47].

The initial values of the nodes are set by the **Learner**'s output layer. The **Learner**'s input is the silhouette of the puzzle and its output is the predicted solution based on the previous puzzles it encountered. Hence, the **Learner** has the potential of influencing the performance of the **Task Solver**. We measure the performance of the **Task Solver** by the number of steps, or moves, to solution. A move is a change of the location and/or orientation of a puzzle piece on the playing board.

We note that while in the validation of the architecture, we use the measure of the number of moves to reach a solution. In the actual game, each round has a time limit, hence the robot will not necessarily succeed in solving the puzzle within that time limit.

Moreover, since at each iteration, the assignment of several nodes can be different from the previous one, it is not clear how to choose the moves that the robot performs on the actual playing board. The solution we used is choosing to move, at each iteration, only one piece corresponding to the node with the highest value, unless it was already moved to its place at a previous iteration. Another criteria is that the value of the node is positive and that the total energy of the solution is negative. This was done in order to select more significant moves, such that the sequence of moves will not be too long.

Nevertheless, it was still possible at the end of this process that some pieces on the board will not be at their correct position, even though the **Task Solver** has converged to the solution of the puzzle. In other words, the solution at the outer layer of the **Task Solver** represents the solution to the puzzle, but not all the pieces have been moved to their correct and final positions. To overcome this, the robot performed moves that removed all the puzzle pieces that were not in the final solution from the playing board, one by one, and then moved the rest of the missing pieces, i.e., pieces that were in the solution but are not yet on the board, one by one. This results in an increase of the number of moves, but limited to seven, due to the maximal number of pieces in the final solution.

4.4 Social Expressivity Component in Tangram Game

4.4.1 Curiosity Expression Corpus: Tangram Questions. We pre-programmed the curious robot to ask sporadic rhetorical questions [11, 20, 32]. We also pre-programmed the robot to react and respond to the child's actions. The curious robot expressed support and promotion of learning, expressing trait curiosity, whereas the non-curious robot expressed neutral task-related facts and responses. Some examples are provided in Table 1 and Figure 2.

4.4.2 Embodied Curiosity: Explain Selection. The curious robot not only selected the tangram, but also verbalized and used non-verbal gestures to convey the **Learnability Estimator** calculations (Table 3). Thus, it communicated why it chose a specific tangram over others. This is a manifestation of the state-curiosity aspect of the architecture.

The choices of facial expressions were made with the rationale that curiosity correlates with the joy of learning and positive feelings, hence happy and curious face in the curious condition. The facial expressions in the non-curious were also chosen to be positive, yet more neutral, such as a nod to express the neutrality to learning. The robot movement was still very animative in the non-curious condition. Also, the robot's joyful greeting and small chitchat at the beginning of the session and a farewell at the end were the same in the curious and non-curious conditions. The responses were randomly picked from a larger database of responses that included variations of the text and the expressions. This was done in order to make the responses more diverse and less predicted.

4.4.3 Social Interaction: Robot's Behavior. This sub-component selects, according to the context, which of the two other sub-components, namely, **Curiosity Expression Corpus** and **Embodied Curiosity**, dictates the robot's behavior. In the context of puzzle selection, the robot responds according to the **Embodies Curiosity** expressions and explains its selection. In other contexts, it behaves according to the **Curiosity Expression Corpus**.

4.5 User: Children's Selections

We recorded children's puzzle selections for all six rounds. We then used the Artificial Curiosity component algorithm on these selections to project the curiosity level of each child (Table 3). In other words, we simulated the **Learner** for each child, where the training sets were the successfully solved tangram puzzles from previous rounds. We then simulated the **Learnability Estimator** for each child in each round for both left and middle puzzles. One option to simulate children's curiosity was to use the **Task Selection** component algorithm to test whether the child's specific selection conformed to the curious condition or the non-curious condition. However, this approach may result in an ill-defined state, namely, the parameters $\overline{H}_{\text{left}}$ and $\overline{H}_{\text{middle}}$ may not fit into any of the conditions.

Hence, instead of a binary state, i.e., curious or non-curious, we used the estimated learnability parameters themselves to calculate a continuous variable that represents children's curiosity based on their choices. The goal is for the variable to represent both direction of curiosity (curious and non-curious) and magnitude (highly curious vs. moderately curious). For this, we formulated a new variable, denoted by C_{child}, to be positive for a curious selection and negative for a non-curious selection (based on Table 3). Thus, it was set to be

$$C_{\text{child}} = \frac{\overline{H}_{\text{selection}} - H_{\text{th}}}{max(\overline{H}_{\text{left}}, \overline{H}_{\text{middle}}) - min(\overline{H}_{\text{left}}, \overline{H}_{\text{middle}})}, \tag{1}$$

where $\overline{H}_{\text{selection}}$ is equal to $\overline{H}_{\text{left}}$ if the child selected the left puzzle, $\overline{H}_{\text{middle}}$ if the child selected the middle puzzle, and $(\overline{H}_{\text{left}} + \overline{H}_{\text{middle}})/2$ if the child selected the right (unknown) puzzle. The values of C_{child} were truncated at ± 1 to reduce outliers.

Furthermore, we defined the threshold, H_{th}, to be updated in a continuous manner, based on the child's estimated curiosity:

$$H_{\text{th}} = \frac{1 + C_{\text{child}}}{2} H_{\text{th}}^{C} + \frac{1 - C_{\text{child}}}{2} H_{\text{th}}^{NC}. \tag{2}$$

Hence, the threshold varied continuously between H_{th}^{C} for $C_{\text{child}} = 1$ and H_{th}^{NC} for $C_{\text{child}} = -1$. For the first round, $C_{\text{child}} = 0$, resulting in $H_{\text{th}} = 0.3$.

This definition of C_{child} guarantees that it is positive if the **Task Selection** had resulted in a curious condition, negative if the **Task Selection** had resulted in a non-curious condition, and has a continuous value. C_{child} and H_{th} were calculated after each child's selection, i.e., for each round. Hence, with this variable, we were able to project the curiosity level of our participants throughout the entire interaction.

Reformulating the robot's algorithm (Table 3) to a continuous regime, resulted in a simplified equation for the child's estimated curiosity (Equation 1). The rationale behind this equation states that the threshold entropy (H_{th}) dictates the transition between a curious and non-curious selection. Thus, if the child selects a task for which the entropy is higher than this threshold, it means that the child's goal is to maximize entropy and vice versa. The denominator and truncation create a normalized parameter.

It is important to clarify that the child's curiosity estimation equation is *based* on the robot's algorithm, but is not identical. The robot's algorithm maps two curiosity states (curious, non-curious) to one of three decisions (left, middle, right). The child's estimated curiosity equation maps one of three decisions (left, middle, right) to a continuous curiosity state ($C_{\text{child}} \in [-1, 1]$). These are not identical, but had the robot's decisions been estimated using the child's curiosity equation, the correct robot condition would have been acquired.

5 EXPERIMENTAL SETUP

5.1 Robot Platform

We used the social robot Jibo to carry out our study (see Figure 5). It is an appealing, expressive, child-friendly robot (12" tall, 6" diameter cylindrical body). It was designed for long-term deployments in various settings such as homes, schools, and therapeutic centers. Jibo is a sleek, white robot that has three degrees of freedom (neck, waist, and base). It has a screen that graphically displays its animated face. Jibo interacts with children as a peer-like learning companion. Accordingly, it is designed with a personality to convey peer-like attributes with a warm, friendly demeanor such as child-like expressions and movements.

5.2 Conditions

The study involved two conditions: curious and non-curious robot. The two conditions differed in expressing trait curiosity throughout the interaction such as asking rhetorical questions about the task (curious) and reacting to the child's puzzle solving. They also differed in the expression of state curiosity during the puzzle selection interaction, e.g., seeking maximum (curious) or minimum (non-curious) learnability, and in the explanation of their moves. Tables 1 and 3 provide some

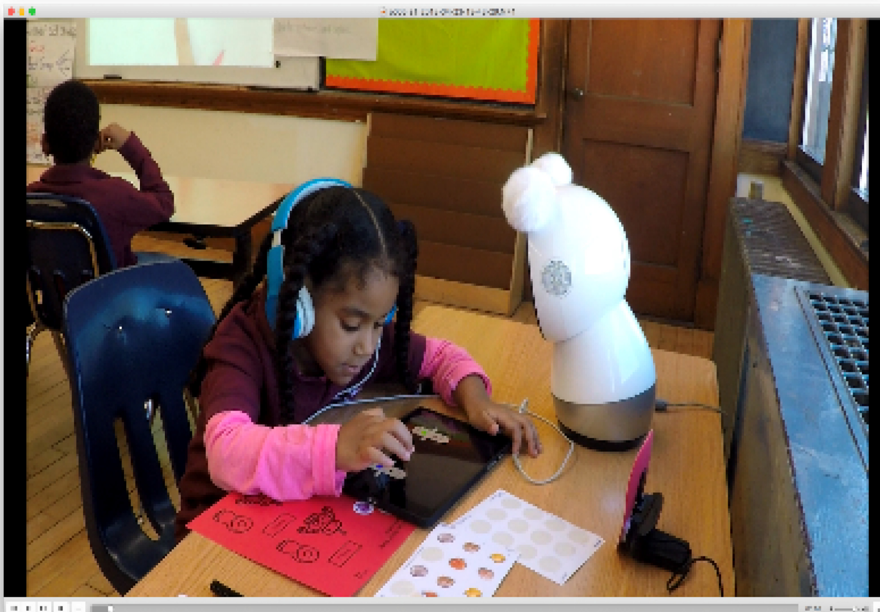


Fig. 5. Our peer-like social robot, Jibo. It engaged with children to solve tangram puzzles with our game app.

examples of how the robot behaved, based on its curiosity condition. Note that our study design intentionally did not include a negative curiosity condition, e.g., "I do not want to learn," for ethical reasons. Instead, the non-curious condition was a neutral condition that is absent of pro-curious behaviors, with some curiosity comments omitted or replaced with factual and task-related statements, together with external reward maximization, i.e., collecting shells through selecting easy puzzles. The curious robot did not succeed in solving the puzzles in the last three rounds due to the time limit and the fact that it chose the more complex puzzles. As the more complex puzzles have more pieces, they are more difficult and the **Task Solver** requires more moves in order to solve them. Thus, the robot was less successful in solving them within the time limit.

5.3 Participants

Participants in kindergarten through 2nd grade were recruited from two local schools. A total of 92 children played with the robot (Age = 6.1 ± 1 yrs, female = 46). Twelve children had technical difficulties, caused by the schools' Wi-Fi connectivity issues, and did not reach the end of the interaction. Out of the remaining 80 who completed the task, 15 had technical issues, which were due to the system coming to a halt and requiring a restart to the last game state. The remaining 65 children completed the interaction with no technical issues (Age = 6.2 ± 1 yrs, female = 33). In the curious robot condition there were 30 participants (Age = 6.2 ± 1 yrs, female = 15) and in the non-curious condition 35 participants (Age = 6.2 ± 1 yrs, female = 18).

The interaction took place in an extracurricular room in each school during their school hours. Each child played with the robot for about 20 minutes. The study was approved by the Institutional Review Board, and all participant's parents signed a consent form.

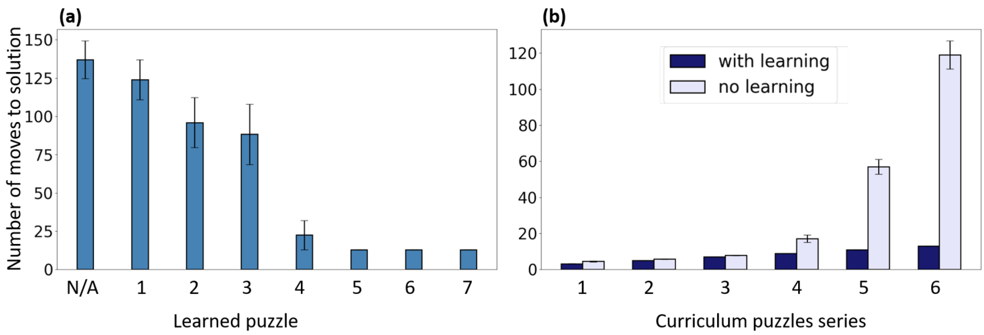


Fig. 6. Validation of Learner. (a) Average number of moves to solution of thefinal puzzle in curriculum, as a function of the example given to the Learner: tangram predictor. (b) Average number of moves to solution as a function of the puzzle in curriculum, given the Learner learned (dark) or not-learned (light) the previous puzzle in the curriculum.

6 ARCHITECTURE VALIDATION

In this section, we present the validation of the cognitive architecture with relation to the Artificial Curiosity components, i.e., the **Learner**, **Learnability Estimator**, and **Task Selection**.

6.1 Learner

The first validation is whether the **Learner** Component of the robot, i.e., the tangram predictor, improves with more information. To test this, we calculated the number of moves, required by the **Task Solver**, to solve the final and most complex tangram involving seven pieces. We note that while in the actual game there is a time limit for each round, in order to better estimate the architecture's performance we have removed this time constraint. The **Learner** provides the initial input to the **Task Solver** by setting the **Task Solver**'s nodes to be the **Learner**'s output nodes (see Figure 4). Thus, an improvement of the **Learner** means it supplies a better starting point for the **Task Solver**. Figure 6(a) shows the number of moves required to solve the final puzzle, after the **Learner** learned a specific puzzle (1–7). The figure clearly shows that more information, i.e., a more complex puzzle as an example, improves the solution, validating that the **Learner** indeed learns.

Figure 6(b) shows the number of moves required to solve a puzzle (2–7), after the **Learner** learned the previous puzzle (1–6). We compared the number of moves with learning and without, i.e., random initial activation of the **Task Solver**. The figure clearly shows that learning the previous puzzle improves the task solution, with an increasing difference the more complex the puzzle.

This validation process has been performed on the simple **Learner** we implemented, with the sparse training set, i.e., previous puzzles. However, the same validation process, i.e., testing whether information supplied during the interaction and learned by the **Learner** improves the **Task Solver**'s ability to solve the task, can be performed on more complex architectures (see Section 8.3).

6.2 Learnability Estimator

To better understand the **Learnability Estimator**, we calculated its output for puzzles $n, n+1$ in the curriculum for $n = 1, \ldots, 6$, given the **Learner** learned all the previous puzzles $m = 1, .., n$. This represents the scenario in which the robot learned the curriculum and now has to estimate the learnability of the next lesson. Figure 7(b) shows that the next puzzle has a higher estimated

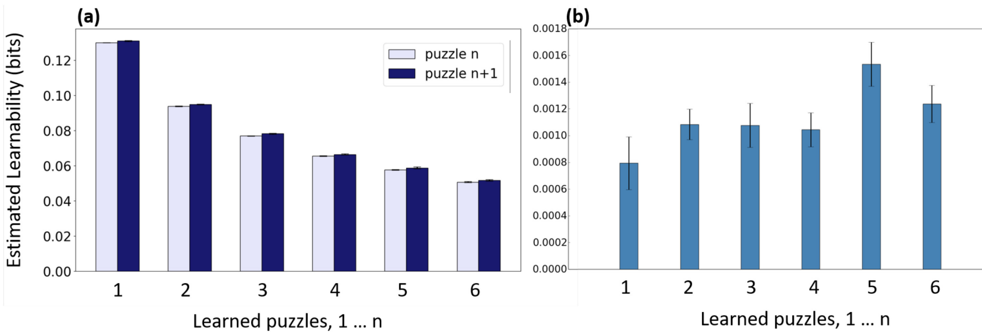


Fig. 7. Validation of Learnability Estimator. (a) Average estimated learnability (average entropy in bits) for puzzles n (light) and $n + 1$ (dark) after learning puzzles $1, \ldots, n$ in the curriculum, for $n = 1, \ldots, 6$. (b) Estimated learnability difference from (b).

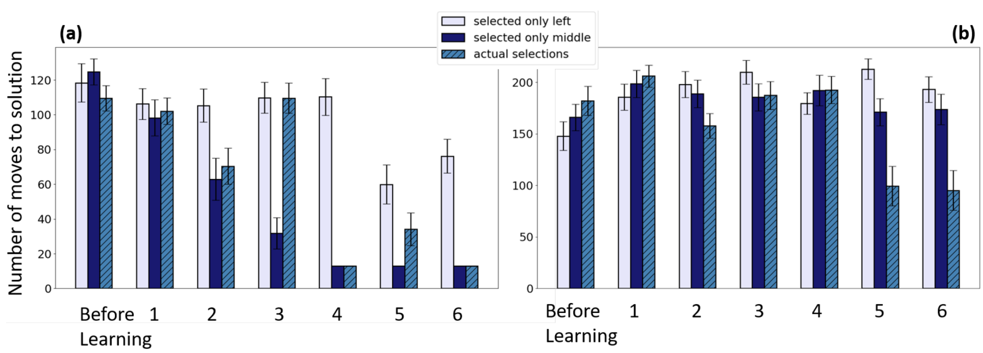


Fig. 8. Validation of Task Selection. Average number of moves to solution of the final puzzle in the (a) known and (b) unknown curricula, as a function of rounds played and based on selection criteria. The actual selections were middle, middle, right, middle, right, middle.

learnability, meaning that a new puzzle will result in more learning than a known one. Furthermore, the learnability decreases as the **Learner** gains more information (see Figure 7(a)), since the previous puzzles in the curriculum have overlapping pieces, thus the **Learner** becomes more confident in most of the puzzle, i.e., lower average entropy.

6.3 Task Selection

To test whether the learnability-based selection improves performance, we computed the number of moves it takes the **Task Solver** to solve the final puzzle in the curriculum after each round. We compared three scenarios (Figure 8): (i) "Selected only left" represents selecting the left puzzle in each round, and corresponds to the non-curious condition. (ii) "Selected only middle" represents selecting the middle puzzle in each round, and corresponds to the selection based solely on the **Learnability Estimator**. (iii) "Actual selections" represents the actual selections the curious robot selected in the game according to the treasure selection algorithm (see Table 3). The actual selections were middle, middle, right, middle, right, middle.

Figure 8(a) shows the number of moves to solve the *presented* curriculum's final puzzle, whereas Figure 8(b) shows the number of moves to solve the *unknown* (right, "?") curriculum's final puzzle.

As can be seen, overall the non-curious selection is the most inefficient, since it learns the least, having chosen the lowest estimated learnability. While the solely **Learnability**

Estimator–based selection is always the most efficient in solving the curriculum's final puzzle (Figure 8(a)), it does not perform well on the unknown final puzzle (Figure 8(b)). In contrast, the curious robot's behavior, which also periodically selects the unknown puzzle, performs well on both curricula, showing that explorations into the unknown, while hindering specific task-learning, can promote generalization.

7 SOCIAL CURIOUS ROBOT EFFECTS ON CHILDREN

7.1 Children's Selections

We next studied the influence of the conditions, i.e., curious vs. non-curious robot, on the child's behavior. We used all 92 children's data, since we were interested in moment-by-moment selections, thus even partial data from the interaction could be used. We analyzed whether the robot's selections predict the child's selections. We found that the robot's selection is a significant predictor of the child's selection of the unknown puzzle, compared to the left puzzle ($\chi^2(4) = 10, p = 0.04$ multinomial regression, with robot's selection of left $\beta = -1.0, p = 0.002$ and middle puzzles $\beta = -0.8, p = 0.03$ being significant predictors of child's selection of unknown puzzle compared to left puzzle). In other words, if the robot did not select the unknown puzzle, the child chose significantly less the unknown puzzle.

Furthermore, since only the curious robot selected the unknown puzzle, we found that, unsurprisingly, the condition is a significant predictor of the child's selection of the unknown puzzle compared to the left one ($\chi^2(2) = 6.5, p = 0.039$ multinomial regression, with robot's condition, $\beta = 0.59, p = 0.012$ being a significant predictor of child's selection of unknown puzzle compared to left puzzle). These results show that the robot's behavior, dictated by whether it is curious or not, influences the child's selections.

7.2 Children's Simulated Curiosity

Based on the Artificial Curiosity Component of the robot, we have formulated a curiosity estimation for each child, wherein we extracted their simulated curiosity variable, C_{child}. We used only the 65 children who completed the interaction without any technical issues, since we were interested in the entire interaction as a whole. Furthermore, we defined the initial simulated curiosity as the average of the first two rounds' simulated curiosity, and the final simulated curiosity as the average of the last two rounds.

We first tested whether there was any dynamical condition-agnostic effect on the simulated curiosity. A repeated measures analysis shows there was a significant decrease from the initial to final simulated curiosity (initial: $M = 0.185, SD = 0.658$; final: $M = -0.177, SD = 0.558$, $F(1, 64) = 10.634, p = 0.002$; within-subjects repeated measures). In other words, children's simulated curiosity, as calculated by their selection and Equation (1), *decreased* during the game.

We next performed an in-depth analysis to understand the effect of robot conditions in the change of children's simulated curiosity between the first two and the final two rounds. The analysis showed that the decrease in curiosity trend was mainly driven by the non-curious condition. A two-tailed t-test showed a non-significant drop in the curious condition, $\Delta = -0.115 \pm 0.188$, $t(29) = -0.611, p = 0.546$, and a significant drop in the non-curious condition, $\Delta = -0.573 \pm 0.119$, $t(34) = -4.79, p < 0.001$. The difference in curiosity change (Δ) between the conditions was found to be significant ($F(1, 63) = 4.47, p = 0.038$; one-way ANOVA).

These results show that the robot's curiosity, expressed as a puzzle selection behavior, influenced children's selection behavior (as discussed in Section 7.1), and that the same trend exists in children's simulated curiosity measured using a similar algorithm that drove the robot's puzzle selection.

8 DISCUSSION AND FUTURE WORK

We introduced a novel expressive cognitive architecture that links intrinsic motivation variables to social expressivity. We validated the artificial curiosity module and presented initial results from a user study that showed that a robot's externalized curiosity-driven behaviors can influence children's behaviors.

8.1 Cognitive Architecture for HRI

The proposed cognitive architecture relates specifically to HRI. Refs [2, 6, 68] also presented HRI-related architectures, but have not directly connected internal variables related to curiosity to the expressiveness of the robot. This unique link, between computational variables to social interaction, are an important step toward fully autonomous social robots, that can not only interact with people, but also explain their inner workings. We have shown, in the experimental study, that such an approach can also influence and promote children's expressions of curiosity.

8.2 Curious Social Robots

Previous studies have shown that robots that behave like curious agents affect their human counterparts [11, 26]. Nevertheless, the novel results presented here have important implications: first, the field of HRI strives for such repeatable results as they give more confidence that indeed children can "catch curisoity" from a social robot [26]. Second, the robot was fully autonomous, as opposed to Ref. [11]. Third, our results show that children's moment-by-moment selections are affected by the curiosity of the robot. Finally, the robot in the curious condition was, by our definition, truly curious, as its behavior was driven by the goal to maximize learnability. Hence, this contribution is the first, to the best of our knowledge, that implements a truly curious social expressive robot. This is an important step toward open-ended learning robots in social contexts.

However, the link between the verbal and non-verbal expressions and action selections was predefined in our setup. In future work, we will explore how this can be generalized. Curiosity is defined as joy of learning, and seeking to learn, so non-verbal expressions of happiness either in learning or anticipation of learning can be generalizable. Verbal expressions are more context dependent, yet general expressions can be used such as "I know how to solve the other ones so I will try the X," where X comes from the context.

8.3 Artificial Curiosity

We implemented a simplified artificial curiosity module, whereas state-of-the-art algorithms involve **deep reinforcement learning** (**DRL**) modules [29, 55]. However, the generalized architecture and validation scheme can still be applied to more complex computational models. Thus, for example, the intrinsic reward signal that comes from the intrinsic motivation module in DRL algorithms can be linked to social expressivity. It is important to keep in mind, though, that our architecture is used in real, short social interactions, whereas DRL algorithms require vast amounts of data and training time, and rarely use the intrinsic reward after training [55]. Hence, the integration of these modules is left for future work.

Furthermore, most artificial curiosity algorithm implementations [25] apply to non-social robots, where their actions are related to their embodiment, objects in their surroundings, and other robots. In this contribution, we did not apply the curiosity algorithm on the social interaction itself [55], but have connected the robot's inner variables to social interaction, via verbal and non-verbal expressions.

8.4 Effects on Children's Behaviors

In our implementation, the robot's curiosity determined both its task-related action, e.g., tangram selection, and its social expressivity. Previous studies have shown that interacting with social robots may cause entrainment, i.e., [31], people tending to adopt the behaviors of their robotic partners. Thus, children may be entrained to the robot's choices, and not necessarily be influenced by its social expressivity. However, studies have shown that subtle changes in verbal and non-verbal expressivity can result in changes in curiosity [26] and growth mindset [52] expressions. Further research is needed in order to control the individual effects of the task- and social-related actions on children's behaviors.

We have shown that children express curious behaviors, dependent on the curiosity of the robot. However, the cause of this change is unclear. It can stem from social mimicry [60], but whether the children perceive the robot as their peer (or in-group) [60] or as a teacher [61] is unknown. It can also originate from an increased sense of safety [3, 21], that enables more expressions of curiosity, or result from a real change in intrinsic motivation [4, 16]. One way to resolve this uncertainty is to use novel neuronal recording, e.g., fMRI, to detect "curiosity states" in the brain [27, 35] during and after the interaction. In future work, we will also investigate whether there is evidence of transfer of such curious behaviors to different tasks.

8.5 Curiosity Assessment

We have used similar algorithms for both robot's behavior (Table 3) and assessing the curiosity of the children (Equation (1)). However, curiosity is a multi-faceted characteristic and different model-based tools have been recently developed to assess uncertainty seeking [33], free exploration [26, 64], and question asking [67]. A more comprehensive study is required to validate our suggested model-based tool for simulated curiosity and to add it to the growing toolbox of curiosity assessment tools.

Furthermore, the aforementioned curiosity assessment tools cannot be used during the intervention, but only as pre-post tests. Our suggested simulated curiosity approach can be generalized to other curiosity-based interventions as a dynamical moment-by-moment assessment tool.

9 CONCLUSIONS

To conclude, the novel architecture and evaluation studies presented here suggest that curious social robots, which are capable of choosing their actions based on intrinsic learning motivations, as well as expressing these motivations to their social environment, may positively influence the behavior of their compatriots humans.

ACKNOWLEDGMENTS

Any opinions, findings, and conclusions, or recommendations expressed in this article are those of the authors and do not represent the views of the grant agency. G. G. is a Jacobs Foundation Fellow.

REFERENCES

[1] Joshua Achiam and Shankar Sastry. 2017. Surprise-based intrinsic motivation for deep reinforcement learning. arXiv: 1703.01732. Retrieved March 2017 from http://arxiv.org/abs/1703.01732.

[2] Carole Adam, Wafa Johal, Damien Pellier, Humbert Fiorino, and Sylvie Pesty. 2016. Social human-robot interaction: A new cognitive and affective interaction-oriented architecture. In *International Conference on Social Robotics*. Springer, 253–263.

[3] Mary D. Ainsworth and Silvia M. Bell. 1970. Attachment, exploration, and separation: Illustrated by the behavior of one-year-olds in a strange situation. *Child Development* 41, 1 (1970), 49–67. DOI: https://doi.org/10.2307/1127388

[4] Gianluca Baldassarre, Tom Stafford, Marco Mirolli, Peter Redgrave, Richard M. Ryan, and Andrew Barto. 2014. Intrinsic motivations and open-ended development in animals, humans, and robots: An overview. *Frontiers in Psychology* 5 (2014). DOI: https://doi.org/10.3389/fpsyg.2014.00985

[5] A. G. Barto, S. Singh, and N. Chentanez. 2004. Intrinsically motivated learning of hierarchical collections of skills. In *International Conference on Developmental Learning (ICDL'04).* DOI: https://doi.org/citeulike-article-id:1693122

[6] Paul E. Baxter, Joachim de Greeff, and Tony Belpaeme. 2013. Cognitive architecture for human-robot interaction: Towards behavioural alignment. *Biologically Inspired Cognitive Architectures* 6 (Oct. 2013), 30–39. DOI: https://doi.org/10.1016/j.bica.2013.07.002

[7] Martin Bichsel and Peter Seitz. 1989. Minimum class entropy: A maximum information approach to layered networks. *Neural Networks* 2, 2 (Jan. 1989), 133–141. DOI: https://doi.org/10.1016/0893-6080(89)90030-0

[8] Cynthia Breazeal. 2009. Role of expressive behaviour for robots that learn from people. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 364, 1535 (2009), 3527–3538. http://rstb.royalsocietypublishing.org/content/364/1535/3527.short

[9] C. Breazeal and B. Scassellati. 1999. How to build robots that make friends and influence people. In *Proceedings of the 1999 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'99)*, Vol. 2. 858–863. DOI: https://doi.org/IROS.1999.812787

[10] Cynthia L. Breazeal. 2004. *Designing Sociable Robots.* MIT Press.

[11] 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 (CHI'19).* ACM, New York, NY, 406:1–406:12. DOI: https://doi.org/10.1145/3290605.3300636event-place: Glasgow, Scotland Uk.

[12] Chaona Chen, Oliver G. B. Garrod, Jiayu Zhan, Jonas Beskow, Philippe G. Schyns, and Rachael E. Jack. 2018. Reverse engineering psychologically valid facial expressions of emotion into social robots. In *2018 13th IEEE International Conference on Automatic Face Gesture Recognition (FG'18).* 448–452. DOI: https://doi.org/10.1109/FG.2018.00072

[13] Chaona Chen, Laura B. Hensel, Yaocong Duan, Robin A. A. Ince, Oliver G. B. Garrod, Jonas Beskow, Rachael E. Jack, and Philippe G. Schyns. 2019. Equipping social robots with culturally-sensitive facial expressions of emotion using data-driven methods. In *2019 14th IEEE International Conference on Automatic Face Gesture Recognition (FG'19).* 1–8. DOI: https://doi.org/10.1109/FG.2019.8756570

[14] Cédric Colas, Pierre Fournier, Mohamed Chetouani, Olivier Sigaud, and Pierre-Yves Oudeyer. 2019. CURIOUS: Intrinsically motivated modular multi-goal reinforcement learning. In *International Conference on Machine Learning*. 1331–1340. http://proceedings.mlr.press/v97/colas19a.html.

[15] Edoardo Conti, Vashisht Madhavan, Felipe Petroski Such, Joel Lehman, Kenneth Stanley, and Jeff Clune. 2018. Improving exploration in evolution strategies for deep reinforcement learning via a population of novelty-seeking agents. In *Advances in Neural Information Processing Systems 31*, S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett (Eds.). Curran Associates, Inc., 5027–5038. http://papers.nips.cc/paper/7750-improving-exploration-in-evolution-strategies-for-deep-reinforcement-learning-via-a-population-of-novelty-seeking-agents.pdf.

[16] Diana I. Cordova and Mark R. Lepper. 1996. Intrinsic motivation and the process of learning: Beneficial effects of contextualization, personalization, and choice. *Journal of Educational Psychology* 88, 4 (1996), 715. http://psycnet.apa.org/journals/edu/88/4/715/.

[17] Jan de Wit, Thorsten Schodde, Bram Willemsen, Kirsten Bergmann, Mirjam de Haas, Stefan Kopp, Emiel Krahmer, and Paul Vogt. 2018. The effect of a robot's gestures and adaptive tutoring on children's acquisition of second language vocabularies. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-robot Interaction.* ACM, 50–58.

[18] David DeSteno, Cynthia Breazeal, Robert H. Frank, David Pizarro, Jolie Baumann, Leah Dickens, and Jin Joo Lee. 2012. Detecting the trustworthiness of novel partners in economic exchange. *Psychological Science* (2012), 0956797612448793.

[19] Carol S. Dweck. 2012. Mindsets and human nature: Promoting change in the Middle East, the schoolyard, the racial divide, and willpower. *American Psychologist* 67, 8 (2012), 614. http://psycnet.apa.org/journals/amp/67/8/614/.

[20] Richard C. Endsley and Susan A. Clarey. 1975. Answering young children's questions as a determinant of their subsequent question-asking behavior. *Developmental Psychology* 11, 6 (1975), 863.

[21] Susan Engel. 2015. *The Hungry Mind: The Origins of Curiosity in Childhood*. Harvard University Press.

[22] Matan Epstein and Goren Gordon. 2018. Social robots as physical curiosity assessment tools. In *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN).* 503–508. DOI: https://doi.org/10.1109/ROMAN.2018.8525790 ISSN: 1944-9437.

[23] Gordon (Ed.). 2018. *The New Science of Curiosity.* Nova Science Publishers, Inc, Hauppauge, N.Y.

[24] 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.

[25] Goren Gordon. 2020. Infant-inspired intrinsically motivated curious robots. *Current Opinion in Behavioral Sciences* 35 (Oct. 2020), 28–34. DOI: https://doi.org/10.1016/j.cobeha.2020.05.010

[26] Goren Gordon, Cynthia Breazeal, and Susan Engel. 2015. Can children catch curiosity from a social robot? In *Proceedings of the 10th Annual ACM/IEEE International Conference on Human-Robot Interaction (HRI'15)*. ACM, New York, NY, 91–98. DOI: https://doi.org/10.1145/2696454.2696469

[27] Matthias J. Gruber, Bernard D. Gelman, and Charan Ranganath. 2014. States of curiosity modulate hippocampus-dependent learning via the dopaminergic circuit. *Neuron* 84, 2 (Oct. 2014), 486–496. DOI: https://doi.org/10.1016/j.neuron.2014.08.060

[28] H. W. Park, M. Gelsomini, J. J. Lee, and C. Breazeal. 2017. Telling stories to robots: The effect of backchanneling on a child's storytelling. In *12th ACM/IEEE International Conference on Human Robot Interaction. ACM.* Vienna, Austria. http://robotic.media.mit.edu/wp-content/uploads/sites/14/2017/01/park2017hri-backchannel.pdf

[29] Muhammad Burhan Hafez, Cornelius Weber, Matthias Kerzel, and Stefan Wermter. 2019. Deep intrinsically motivated continuous actor-critic for efficient robotic visuomotor skill learning. *Paladyn, Journal of Behavioral Robotics* 10, 1 (2019), 14–29. DOI: https://doi.org/10.1515/pjbr-2019-0005

[30] Sandy H. Huang, Martina Zambelli, Jackie Kay, Murilo F. Martins, Yuval Tassa, Patrick M. Pilarski, and Raia Hadsell. 2019. Learning gentle object manipulation with curiosity-driven deep reinforcement learning. arXiv: 1903.08542. Retrieved March 2019 from https://arxiv.org/abs/1903.08542.

[31] Takamasa Iio, Masahiro Shiomi, Kazuhiko Shinozawa, Takahiro Miyashita, Takaaki Akimoto, and Norihiro Hagita. 2009. Lexical entrainment in human-robot interaction: Can robots entrain human vocabulary? In *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 3727–3734.

[32] Jamie Jirout. 2011. *Curiosity and the Development of Question Generation Skills*. In *2011 AAAI Fall Symposium Series*, Technical Report FS-11-04. AAAI, Arlington, VA.

[33] 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. http://www.sciencedirect.com/science/article/pii/S0273229712000123.

[34] N. Justesen, P. Bontrager, J. Togelius, and S. Risi. 2019. Deep learning for video game playing. *IEEE Transactions on Games* (2019), 1–1. DOI: https://doi.org/10.1109/TG.2019.2896986

[35] Min Jeong Kang. 2010. *Three Experimental Studies of Reward and Decision Making*. PhD. California Institute of Technology. http://resolver.caltech.edu/CaltechETD:etd-09082009-143650.

[36] James Kennedy, Paul Baxter, and Tony Belpaeme. 2017. The impact of robot tutor nonverbal social behavior on child learning. *Frontiers in ICT* 4 (2017). DOI: https://doi.org/10.3389/fict.2017.00006Publisher: Frontiers.

[37] Varun Raj Kompella, Marijn Stollenga, Matthew Luciw, and Juergen Schmidhuber. 2017. Continual curiosity-driven skill acquisition from high-dimensional video inputs for humanoid robots. *Artificial Intelligence* 247 (June 2017), 313–335. DOI: https://doi.org/10.1016/j.artint.2015.02.001

[38] Varun Raj Kompella and Laurenz Wiskott. 2017. Intrinsically motivated acquisition of modular slow features for humanoids in continuous and non-stationary environments. arXiv: 1701.04663. Retrieved Jan. 2017 from http://arxiv.org/abs/1701.04663.

[39] Yulia Kotseruba, Oscar Avella, and John Tsotsos. 2016. A review of 40 years of cognitive architecture research: Focus on perception, attention, learning and applications. arXiv preprint arXiv:1610.08602 (2016). 1–74.

[40] Tejas D. Kulkarni, Karthik Narasimhan, Ardavan Saeedi, and Josh Tenenbaum. 2016. Hierarchical deep reinforcement learning: Integrating temporal abstraction and intrinsic motivation. In *Advances in Neural Information Processing Systems.* 3675–3683.

[41] Kate Lawrence, Ruth Campbell, and David Skuse. 2015. Age, gender, and puberty influence the development of facial emotion recognition. *Frontiers in Psychology* 6 (2015), 761.

[42] Jordan A. Litman and Paul J. Silvia. 2006. The latent structure of trait curiosity: Evidence for interest and deprivation curiosity dimensions. *Journal of Personality Assessment* 86, 3 (2006), 318–328. http://www.tandfonline.com/doi/abs/10.1207/s15327752jpa8603_07.

[43] George Loewenstein. 1994. The psychology of curiosity: A review and reinterpretation.*Psychological Bulletin* 116, 1 (1994), 75.

[44] Diana Löffler, Nina Schmidt, and Robert Tscharn. 2018. Multimodal expression of artificial emotion in social robots using color, motion and sound. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction (HRI'18)*. Association for Computing Machinery, New York, NY, 334–343. DOI: https://doi.org/10.1145/3171221.3171261

[45] Christina Moro, Shayne Lin, Goldie Nejat, and Alex Mihailidis. 2019. Social robots and seniors: A comparative study on the influence of dynamic social features on human–robot interaction. *International Journal of Social Robotics* 11, 1 (Jan. 2019), 5–24. DOI: https://doi.org/10.1007/s12369-018-0488-1

[46] F. D. Naylor. 1981. A state-trait curiosity inventory. *Australian Psychologist* 16, 2 (March 1981), 172–183. DOI: https://doi.org/10.1080/00050068108255893

[47] Kemal Oflazer. 1993. Solving tangram puzzles: A connectionist approach. *International Journal of Intelligent Systems* 8, 5 (Jan. 1993), 603–616. DOI: https://doi.org/10.1002/int.4550080502

[48] P. Y. Oudeyer, J. Gottlieb, and M. Lopes. 2016. Intrinsic motivation, curiosity, and learning. *Progress in Brain Research* 229 (Jan. 2016), 257–284. DOI: https://doi.org/10.1016/bs.pbr.2016.05.005

[49] Pierre-Yves Oudeyer and Frederic Kaplan. 2007. What is intrinsic motivation? A typology of computational approaches. *Frontiers in Neurorobotics* 1 (Nov. 2007). DOI: https://doi.org/10.3389/neuro.12.006.2007

[50] P. Y. Oudeyer, F. Kaplan, and V. V. Hafner. 2007. Intrinsic motivation systems for autonomous mental development. *IEEE Transactions on Evolutionary Computation* 11, 2 (April 2007), 265–286. DOI: https://doi.org/10.1109/TEVC.2006.890271

[51] Raul Benites Paradeda, Mojgan Hashemian, Rafael Afonso Rodrigues, and Ana Paiva. 2016. How facial expressions and small talk may influence trust in a robot. In *Social Robotics (Lecture Notes in Computer Science)*, Arvin Agah, John-John Cabibihan, Ayanna M. Howard, Miguel A. Salichs, and Hongsheng He (Eds.). Springer International Publishing, Cham, 169–178. DOI: https://doi.org/10.1007/978-3-319-47437-3_17

[52] Hae Won Park, Rinat Rosenberg-Kima, Maor Rosenberg, Goren Gordon, and Cynthia Breazeal. 2017. Growing growth mindset with a social robot peer. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction (HRI'17)*. ACM, New York, NY, 137–145. DOI: https://doi.org/10.1145/2909824.3020213

[53] Deepak Pathak, Pulkit Agrawal, Alexei A. Efros, and Trevor Darrell. 2017. Curiosity-driven exploration by self-supervised prediction. 16–17. http://openaccess.thecvf.com/content_cvpr_2017_workshops/w5/html/Pathak_Curiosity-Driven_Exploration_by_CVPR_2017_paper.html.

[54] Eyal Peer, Joachim Vosgerau, and Alessandro Acquisti. 2014. Reputation as a sufficient condition for data quality on Amazon Mechanical Turk. *Behavior Research Methods* 46, 4 (2014), 1023–1031. Springer.

[55] Ahmed Hussain Qureshi, Yutaka Nakamura, Yuichiro Yoshikawa, and Hiroshi Ishiguro. 2018. Intrinsically motivated reinforcement learning for human-robot interaction in the real-world. *Neural Networks* 107 (Nov. 2018), 23–33. DOI: https://doi.org/10.1016/j.neunet.2018.03.014

[56] Astrid M. Rosenthal-von der Pütten, Nicole C. Krämer, and Jonathan Herrmann. 2018. The effects of humanlike and robot-specific affective nonverbal behavior on perception, emotion, and behavior. *International Journal of Social Robotics* 10, 5 (Nov. 2018), 569–582. DOI: https://doi.org/10.1007/s12369-018-0466-7

[57] Maha Salem, Friederike Eyssel, Katharina Rohlfing, Stefan Kopp, and Frank Joublin. 2013. To err is human(-like): Effects of robot gesture on perceived anthropomorphism and likability. *International Journal of Social Robotics* 5, 3 (Aug. 2013), 313–323. DOI: https://doi.org/10.1007/s12369-013-0196-9

[58] Alexei V. Samsonovich. 2010. *Biologically Inspired Cognitive Architectures 2010: Proceedings of the First Annual Meeting of the BICA Society*. IOS Press. Google-Books-ID: viYYT8hAnqIC.

[59] S. Schaat, K. Doblhammer, A. Wendt, F. Gelbard, L. Herret, and D. Bruckner. 2013. A psychoanalytically-inspired motivational and emotional system for autonomous agents. In *IECON 2013-39th Annual Conference of the IEEE Industrial Electronics Society*. 6648–6653. DOI: https://doi.org/10.1109/IECON.2013.6700232

[60] Johanna E. van Schaik and Sabine Hunnius. 2016. Little chameleons: The development of social mimicry during early childhood. *Journal of Experimental Child Psychology* 147 (July 2016), 71–81. DOI: https://doi.org/10.1016/j.jecp.2016.03.003

[61] Johanna E. van Schaik and Sabine Hunnius. 2018. Modulating mimicry: Exploring the roles of inhibitory control and social understanding in 5-year-olds' behavioral mimicry. *PLoS ONE* 13, 3 (March 2018), e0194102. DOI: https://doi.org/10.1371/journal.pone.0194102.Public Library of Science.

[62] Jessica L. Schleider and John R. Weisz. 2015. Using Mechanical Turk to study family processes and youth mental health: A test of feasibility. *Journal of Child and Family Studies* 24, 11 (Nov. 2015), 3235–3246. DOI: https://doi.org/10.1007/s10826-015-0126-6

[63] J. Schmidhuber. 1990. *A Possibility for Implementing Curiosity and Boredom in Model-Building Neural Controllers*. MIT Press.

[64] Keren Ben-Tov Sher, Michal Levi-Keren, and Goren Gordon. 2019. Priming, enabling and assessment of curiosity. *Educational Technology Research and Development* 67 (April 2019), 931–952. DOI: https://doi.org/10.1007/s11423-019-09665-4

[65] Christopher Stanton and Jeff Clune. 2018. Deep curiosity search: Intra-life exploration can improve performance on challenging deep reinforcement learning problems. arXiv: 1806.00553. Retrieved June 2018 from http://arxiv.org/abs/1806.00553.

[66] Alexandr Ten, Pramod Kaushik, Pierre-Yves Oudeyer, and Jacqueline Gottlieb. 2020. *Humans Monitor Learning Progress in Curiosity-Driven Exploration*. Technical Report. PsyArXiv. DOI: https://doi.org/10.31234/osf.io/7dbr6

[67] Noam Tor and Goren Gordon. 2020. Digital interactive quantitative curiosity assessment tool: Questions worlds. *International Journal of Information and Education Technology* 10, 8 (2020).

[68] J. Gregory Trafton, Laura M. Hiatt, Anthony M. Harrison, Franklin P. Tamborello II, Sangeet S. Khemlani, and Alan C. Schultz. 2013. ACT-R/E: An embodied cognitive architecture for human-robot interaction. *Journal of Human-Robot Interaction* 2, 1 (2013), 30–55.

[69] Michelle Tran, Laura Cabral, Ronak Patel, and Rhodri Cusack. 2017. Online recruitment and testing of infants with Mechanical Turk. *Journal of Experimental Child Psychology* 156 (2017), 168–178. Elsevier.

Received April 2020; revised January 2021; accepted February 2021