Expression of Curiosity in Social Robots

Design, Perception, and Effects on Behaviour

Jessy Ceha
jceha@uwaterloo.ca
University of Waterloo
Waterloo, ON, Canada

Nalin Chhibber
nchhibbe@uwaterloo.ca
University of Waterloo
Waterloo, ON, Canada

Joslin Goh
jtcgoh@uwaterloo.ca
University of Waterloo
Waterloo, ON, Canada

Corina McDonald
cmmcdona@uwaterloo.ca
University of Waterloo
Waterloo, ON, Canada

Pierre-Yves Oudeyer
pierre-yves.oudeyer@inria.fr
Inria Bordeaux
Talence, France

Dana Kulić
dana.kulic@uwaterloo.ca
University of Waterloo
Waterloo, ON, Canada

Edith Law
edith.law@uwaterloo.ca
University of Waterloo
Waterloo, ON, Canada

ABSTRACT

Curiosity–the intrinsic desire for new information–can enhance learning, memory, and exploration. Therefore, understanding how to elicit curiosity can inform the design of educational technologies. In this work, we investigate how a social peer robot's verbal expression of curiosity is perceived, whether it can affect the emotional feeling and behavioural expression of curiosity in students, and how it impacts learning. In a between-subjects experiment, 30 participants played the game LinkIt!, a game we designed for teaching rock classification, with a robot verbally expressing: curiosity, curiosity plus rationale, or no curiosity. Results indicate that participants could recognize the robot's curiosity and that curious robots produced both emotional and behavioural curiosity contagion effects in participants.

CCS CONCEPTS

• **Human-centered computing → Empirical studies in HCI**.

KEYWORDS

Curiosity; Social Robot Behaviour; Education

ACM Reference Format:
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 *CHI Conference on Human Factors in Computing Systems Proceedings (CHI 2019), May 4–9, 2019, Glasgow, Scotland, UK*. ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3290605.3300636

ACM ISBN 978-1-4503-5970-2/19/05…$15.00
https://doi.org/10.1145/3290605.3300636

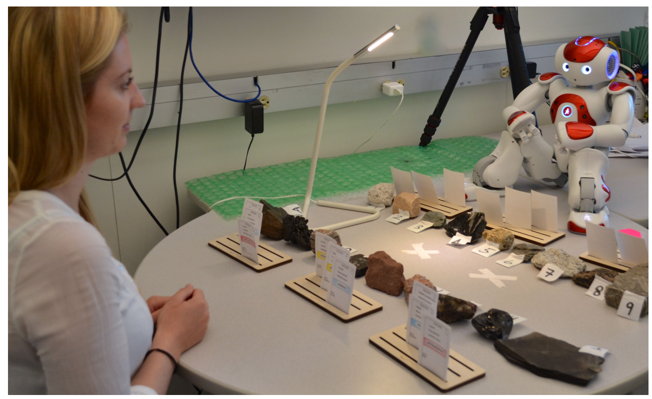

Figure 1: Setup of the game LinkIt!, designed to study human and robot behaviour, and played by participants in the study.

1 INTRODUCTION

Curiosity—the intrinsic desire to explore or obtain unknown information [27]—has been shown to enhance learning [33], promote information-seeking behaviour [39], and improve memory retention (e.g., [18, 20]). Understanding how to elicit and maintain curiosity in students can enable the development of educational technologies that stimulate interest and engagement [2]. Recent research [17, 36] has begun to explore whether educational technologies like robots can foster curiosity in students, and consequently improve learning.

In this work, we investigate how a social peer robot can convey curiosity to students using verbal expressions, and whether those expressions in turn can affect the students' own curiosity, and subsequently enhance learning. A between-subjects experiment was conducted in which participants played LinkIt!, a cooperative game that teaches students how to classify rocks, with a robot designed to: (a) be neutral (expressing no curiosity), (b) express curiosity, or (c) express curiosity and reveal the reason for curiosity. Curiosity was conveyed through on-topic question-asking, i.e., the robot was actively searching for new knowledge on the topic of the specific learning task, by asking questions. We are interested in understanding how robot curiosity is perceived, whether question-asking behaviour can lead to emotional or behavioural curiosity contagion, and what impact it has on learning. Our contributions include:

- A new structured game, *LinkIt!*, that can be used to design and study robot behaviour and human-robot interaction,
- A novel procedure for assessing curiosity through behaviour in a *Free Choice Curiosity Test*, and
- Qualitative and quantitative results showing perception of curiosity in a curious peer robot, and the emotional and behavioural contagion effects on students.

2 RELATED WORK

Robots in Education

Robots and virtual agents have seen increasing deployment in educational settings [5]. Human social interaction has been found to be beneficial for learning [5] and robotic educational systems aim to replicate these benefits by emulating human social roles, such as a teacher or peer.

Teachable peer agents, systems that have the ability to be taught, are of particular interest in education (e.g., [8, 9, 40, 41]), as learning-by-teaching can be a more enriching experience than learning by oneself [15]. Over the past decade, Biswas and colleagues [9], for example, have been developing a teachable virtual agent called Betty's Brain, where students develop concept maps as they 'teach' Betty. The researchers have shown that interaction with Betty's Brain can result in better post-learning gains for students.

In contrast to traditional technologies, students show more interest and higher performance while learning using educational robots [3, 12, 24], and compared to voice-only or virtual agents, the physical presence of robots can produce cognitive learning gains as well as more positive interactions [3, 24]. Social robots must be able to convey social and affective cues in order to interact with people appropriately [10]. Human-like expressions have been successfully implemented in robots through both verbal and non-verbal cues, including expressions of empathy [31], happy/sad states [16], introversion/extroversion [23], positive/negative emotions [45], and curiosity [17].

Curiosity in Social Educational Robots

Promoting curiosity has been shown to play a role in adult [33] and infant [39] learning, and foster early academic achievement, particularly for children with low socio-economic status [35]. The majority of research on curiosity has focused on its elicitation through stimuli that are novel, surprising, conceptually conflicting, or uncertain [7, 22, 37]. In contrast, we are interested in understanding curiosity that is shaped by the social environment.

Shiomi et al. [36] investigated whether a robot acting as a knowledgeable peer could foster curiosity in elementary school children. The robot encouraged the children to ask science questions by prompting them to explain parts of the day's lesson to it, and asking whether they had any questions. The study measured curiosity through a questionnaire administered before and after interaction with the robot, and found that children were more likely to be curious if they asked the robot science-related questions.

Gordon, Breazeal, and Engel [17] compared a curious virtual agent (on a tablet), a curious robot, and a non-curious robot in the context of supporting young children in learning how to read. The agents were portrayed as less-knowledgeable peers learning to read, prompting the children to teach new words. Two curiosity-driven behaviours were implemented in the curious conditions: free exploration and uncertainty seeking. The curious agents were enthusiastic about learning and exploration, challenged the child, and suggested novel moves in their co-player tablet app. Children's curiosity was measured by amount of information-seeking behaviour as a metric of free exploration, question generation, and uncertainty seeking through "The Fish Task" [19], a tablet app recording users' choices of uncertain options. The curious robot resulted in significantly higher free exploration and uncertainty seeking than the non-curious robot, with no differences in question generation. Therefore, only those behaviours modelled by the robot had an impact on children's curiosity. Additionally, even though the children's curiosity was higher, the curious robot did not result in learning gains.

In Gordon et al.'s study, the curious agents were enthusiastic about learning and exploration, and frequently challenged the child, whereas the non-curious robot did not. The curious agents therefore differed from the neutral one along multiple dimensions, making it difficult to identify what influenced the observed changes in curiosity. In our work, we focused on using question-asking behaviour to convey the robot's curiosity about the learning task at hand. Furthermore, in addition to measuring behavioural curiosity contagion, we also measure emotional curiosity contagion to provide a more detailed account of participants' curiosity. Emotional curiosity contagion refers to the transfer of mood and affect (i.e., the emotional state of feeling curious), as opposed to behavioural curiosity contagion which is the transfer of behaviour indicative of curiosity (e.g., free exploration, uncertainty seeking). Lastly, we are also interested in understanding how robot curiosity is perceived by participants.

Revealing an Agent's Internal State

Effective teaching requires an understanding of the learning progress of the student, and not being able to figure out the internal state of an agent can lead to a loss of trust, satisfaction, and acceptance [28]. One way of providing information about the internal model of a system is by providing explanations as to why or why not a system did something [25]. Therefore, to investigate whether improved understanding of the robot's internal state has additional emotional or behavioural contagion effects on participants, we included a condition in which the robot expressed its curiosity *and* revealed the reason behind it.

The remainder of this paper describes the study design, beginning with a description of the game LinkIt!, an explanation of the experimental conditions and methodology, followed by our results and a discussion of the implications.

3 STUDY DESIGN

LinkIt!

LinkIt!, a novel educational game played between a robot and student, is designed to teach classification of rocks into rock types (sedimentary, metamorphic, or igneous) based on visual features, such as fossils. Players (robot and student) sit opposite each other, with a row of 9 rocks in front of each of them (3 sedimentary, 3 metamorphic, and 3 igneous, in random order; setup shown in Figure 1). Every rock in a player's row has a card associated with it, containing the 'ground truth' (Figure 2), i.e., three visual features, the type, and the name, which are not visible to the other player. The features were determined in consultation with geologists and are therefore standard descriptors of the rocks.

The aim of the game is for players to *link* each other's rocks by finding those with a common feature or type. Each round starts with the student taking the top card from an upside down deck placed in between the players. This card contains either a feature or rock type. During the round, both the robot and the student try to identify one of the other player's rocks which they think has this property.

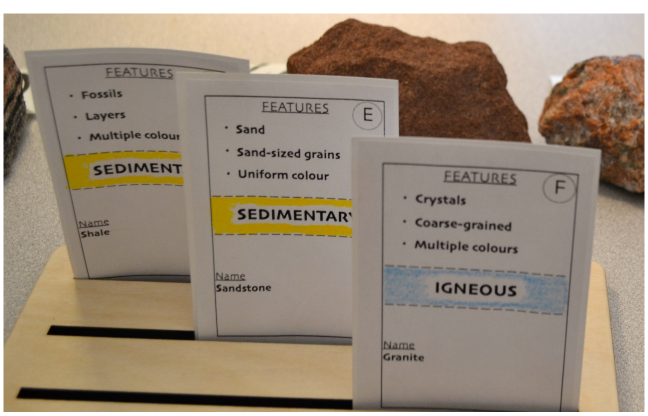

Figure 2: The ground truth cards used in the game.

Once a player makes a guess, the other player silently reads the associated ground truth card to check whether the guess is correct. If both players are correct in identifying the property in each other's rocks, they have successfully linked them. Once linked, players are able to talk about the other features, rock types, or names that are listed on the ground truth cards for the linked rocks. The rocks can then be placed to the side of the game area, and a new card is pulled from the upside down deck. If a player guesses incorrectly, the other player can give them a hint, before trying again. If they guess incorrectly a second time, the round is over, and they will have to pick a new card from the upside down deck. The game ends once all cards in the upside down deck have been played, or all rocks have been linked.

Experimental Conditions

There were three conditions in the study: 1) *Curious+Reveal* – the robot is curious and reveals the reason for its curiosity, 2) *Curious* – the robot is curious but does not reveal the reason, and 3) *Neutral* – the robot is not curious. In contrast to prior work, our robot expressed its curiosity through on-topic question-asking, and in all three conditions, the robot exhibited general enthusiasm for the game. In all conditions, the robot made both correct and incorrect guesses in the game with the same level of accuracy, to convey to the student its role as a peer, rather than a knowledgeable teacher.

Curious+Reveal In the Curious+Reveal condition the robot was curious about rocks and the participant's curiosity about rocks, *and* it made statements that revealed why it was curious. The reason behind the robot's curiosity varied; it was either because of novelty (e.g., "*I have never seen shiny rocks before! What could make them so shiny? Do you not wonder that too?*") or expectation violation (e.g., "*Huh. I would have expected the rock to look different. I am curious though, do the holes form when gas bubbles get trapped when the lava cools? Do you have any idea?*"), both of which are known causes of curiosity [6].

Curious The Curious condition also had the robot exhibit curiosity about rocks and the participant's curiosity, but it did not reveal its rationale. For example, the robot said, "*I am curious. Do the holes form when gas bubbles get trapped when the lava cools? Do you have any idea?*".

Neutral In the Neutral condition, the robot did not express any curiosity for the rocks or the participant's curiosity. For example, the robot said, "*I believe holes can form when gas bubbles cool*".

In both curious conditions, in addition to asking the participant general rock formation questions (as described above), the robot asked more specific questions about the participant's rocks, e.g., "*What type of rock is your rock?*", or "*What other features does your rock have?*". In the Neutral condition, rather than asking these specific questions, the robot would share information about its own rock, i.e., "*My rock is metamorphic*". In this way, we aimed to keep the amount of words said by the robot relatively consistent across conditions.

Table 1: Illustration of LinkIt! dialogue, with examples of differences between conditions (P = Participant, R = Robot).

Neutral	Curious	Curious+Reveal
	P: I'll pick a new card. It's a feature card, glassy surface.	
	R: Who goes first?	
	P: I think number 4 has a glassy surface. Does it?	
	R: Alright, let me check my card. Yeah! That's correct! Hmmm, let's have a look at your rocks. Can you show me rock H?	
R: Some rocks have a really glassy surface. They probably cooled too quickly to form crystals.	**R**: I am curious why some rocks have a glassy surface. Could they have cooled too quickly to form any crystals? What do you think?	**R**: A glassy surface is new to me! I am curious why some rocks have that. Could they have cooled too quickly to form any crystals? What do you think?
	P: Um. I think you're right. So, if they cool down really fast they become glassy. Do you choose rock H?	
	R: Alright. I pick that rock.	
	P: Indeed, it has a glassy surface. You're right!	
	R: Yay! We linked the rocks!	
	P: Woohoo!	
R: My rock is igneous.	**R**: What rock type is your rock?	**R**: What rock type is your rock?
	P: It's igneous.	
	R: Cool! Mine too!	
R: I think all igneous rocks come from volcanoes.	**R**: I am curious whether all igneous rocks come from volcanoes. Do you think so?	**R**: I am curious whether all igneous rocks come from volcanoes. Do you think so?
	P: I think all igneous rocks come from lava or magma. So, I guess we could say they all come from volcanoes.	
R: My rock is shiny.	**R**: What other features does your rock have?	**R**: What other features does your rock have?
	P: It has multiple colours.	
	R: I see! They do look kind of similar.	
R: My rock is called Obsidian.	**R**: What's the name of your rock?	**R**: What's the name of your rock?
	P: Obsidian.	

Robot speech was carefully designed to convey the same informational content about rocks to participants in all three conditions. The content used in the statements or questions made by the robot were taken directly from an introductory video about rock formation and classification that participants and robot were shown prior to playing the game.

Research Questions and Hypotheses

Our main research questions are:

(1) Are the curious robots perceived to be more curious than the neutral one? If so, why?
(2) Does curious question-asking behaviour in a robot produce emotional and/or behavioural curiosity contagion?
(3) What impact does providing rationale for curiosity have?
(4) Does curious robot behaviour affect learning?

Our hypothesis was that the robot's curious behaviour would elicit both behavioural and emotional curiosity contagion effects in participants, that could be about rocks and/or the robot. Specifically, participants would exhibit more of the curiosity-driven behaviour modelled by the robot, i.e., question-asking. Motivated by the elicited curiosity-driven behaviour, participants may pay more attention during the game and think more in depth about the content, thus learning more in the process, or actively seek information to gain more knowledge on rocks. Additionally, we hypothesized that the robot revealing the rationale behind its curiosity would provide participants with a better understanding of the robot's internal model, resulting in more pronounced effects on learning, and curiosity contagion.

Methodology

Participants. 30 students [20 female, 8 male, 2 other; age range: 18-49, mean 23, median 22] were recruited from a research-based university and randomly assigned to a condition. Participants volunteered for the study by responding to posters, and varied in their degree programs (i.e., Computer Science to Psychology) and level of education (7% PhD, 10% Masters, and 83% Bachelors students). Both native (63%) and non-native (37%) English speakers participated in the study.

Materials. The materials used in the game are shown in Figure 1. In addition to the robot NAO, a small humanoid robot developed by SoftBank Robotics, there were 18 rocks, playing cards, and stands for the cards. There were 14 additional rocks which were used for the quizzes and free period (both described in more detail in Procedure).

Wizard of Oz Interface. The SoftBank Robotics Python SDK (version 2.1.4.13) was used to teleoperate the robot. An interface (programmed in Javascript, and deployed as a new application inside the robot, accessible at NAO's IP address) was implemented to allow for quick and easy selection from a set of predetermined statements and questions—supporting a more 'natural' interaction, as the possibility for long pauses was reduced, and consistency of robot responses between trials was ensured. Additionally, hand and arm gestures were evoked through the *ALAnimatedSpeech* module of the Python SDK, and the built-in "joyful" style was applied to NAO's voice to convey a positive attitude towards the game. The wizard handled participants speaking out of context by replying that it does not know about anything other than rocks.

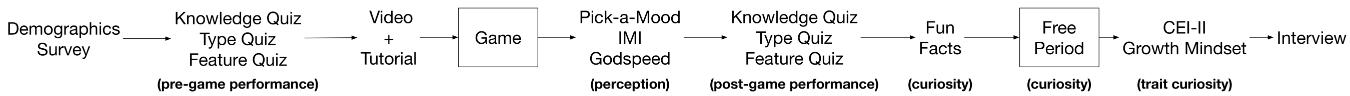

Figure 3: Study Procedure for Session 1 - with the variables being measured shown in brackets

Procedure

The procedure consisted of Session 1 (90 minutes) and Session 2 (a 30-minute session one week later). Participants interacted with the robot in Session 1; the procedure is shown in Figure 3. Both sessions were audio and video recorded.

Session 1—Pre-Game Session 1 began with the information letter, consent form, a demographics questionnaire, and three short quizzes (taking 10 minutes total) on rock classification, to provide a baseline from which learning gains could be measured. The *Knowledge quiz* involved multiple choice questions on rock formation. The *Type* and *Feature quizzes* had participants inspect the 18 rocks that would be used in the game (game rocks), plus 6 rocks not in the game (non-game rocks), and identify their types and features.

After the quizzes, participant and robot together watched a 3-minute video describing rock formation and classification. They watched together so that the participant was aware of how the robot knew about rocks. Next, LinkIt! was explained to both the participant and robot, during which the robot asked clarification questions, introduced itself to the participant, and asked for their name, in order to calibrate the participant's expectations on the capabilities of the robot and how they could converse with it. After the explanation, participant and robot were told they would be left alone in the room for 30 minutes while they played the game.

Session 1—Post-Game Following the game, participants were given a set of questionnaires about their experience and their perception of the robot. First, was the "pick-a-mood" self-report scales [14]; one for themselves and one for the robot, asking them to select one (or more) out of 9 characters expressing emotion. Second, was the standardized Intrinsic Motivation Inventory (IMI), measuring their self-reported feeling of enjoyment, competence, effort, and relatedness with the robot [13, 43, 44]. Third, participants were given the Godspeed questionnaire [4], which consists of semantic differential scales on anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of the robot. After the questionnaires, participants proceeded to complete the post-game Knowledge, Type, and Feature quizzes, which were identical to the pre-game quizzes. In addition, to understand what rocks participants may be curious about, participants were given the opportunity to ask the researcher for a *fun fact* about any of the 24 rocks in the quizzes, with the constraint that the researcher would only tell them about a maximum of three of the chosen rocks.

Session 1—Free Period Next, participants were given 5-10 minutes to freely choose how to interact with the robot. In pilot studies, we found that immediately after participants were told the study was over, they exhibited curiosity-driven behaviours, e.g., asking questions about the remaining rocks in the game, or about the robot. The free period allowed us to apply a novel procedure for assessing curiosity through behaviour in a *Free Choice Curiosity Test*, which we designed to enable systematic observation and quantifying of a participant's curiosity about rocks and the robot.

Participants were given eight rocks not used in the game or quizzes, and four 'ground truth' cards (with three features, the type, and the name listed on them) for four of the eight rocks. Participants were told they were free to do what they wanted with the rocks and the robot, and that the researcher would step out for 10 minutes. By presenting participants with eight new rocks, four with the ground truth, and four without, we constructed a measure of curiosity: If participants decided to ask the robot about any of the rocks for which they had the ground truth, it suggested that the participant was more curious about testing the robot's knowledge than knowing about the rocks. However, if they chose to ask the robot about the rocks for which they did not have the ground truth, it indicated curiosity for the rock.

If the robot was asked about any of the rocks, it would correctly identify the features and type, because, if the robot was incorrect for any of the four rocks for which participants had the ground truth, they may notice that the robot was incorrect and therefore, even though they are curious about the other rocks, not ask the robot about them. The robot's personality: Neutral, Curious, or Curious+Reveal, carried over from the game to the free period.

Session 1—Post-Free Period After the free period, participants filled out the PERTS Growth Mindset Assessment [32] and the Curiosity and Exploration Inventory-II (CEI-II) [21], both measuring *trait* curiosity–the characteristic of always having an interest in learning or obtaining new information. In contrast, *state* curiosity describes curiosity elicited by external situations—measured in our study by fun facts, verbal behaviour during the free period and game, and Likert scale questions on curiosity (described below).

Finally, we conducted a semi-structured interview with participants to understand their perception of the robot, the rationale behind their behaviour during the free period, as well as their overall experience. To measure participants'

Table 2: Data Table Summary

Background	gender, age, degree, native_english_speaker
Perception	anthropomorphism, animacy, likeability, perceived_intelligence, perceived_safety, robot_mood, participant_mood
Trait	stretching, exploration, growth_mindset
Experience	enjoyment, competence, effort, relatedness
Curiosity	[pre|post|1wk]_curiosity_rocks, [pre|post|1wk]_curiosity_robot, [post|1wk]_funfact_rock[x]
Behaviour	avg lines, avg words, qa, ontask_qa, social_qa, uncertainty, argument, justificaiton, suggestion, agreement, disagreement, idea_verbalization, information_sharing, hypothesis_generation
Learning	[pre|post|1wk]_quiz_accuracy, [pre|post|1wk]_type_accuracy, [pre|post|1wk]_feature_[precision|recall|accuracy|specificity]

perception of the robot's curiosity, we first asked participants whether they thought the robot had any particular personality. This was followed by separate Likert scale rating questions on how enthusiastic, engaged, and curious participants felt the robot was during the game. The questions were asked in this order so as not to make curiosity a focal point of the interview and minimize priming effects and bias. In general, we explicitly avoided the use of the word *curiosity* in any questionnaires or interview questions until the end of the session. Participants were also asked whether *they* felt enthusiastic, engaged, and curious during the game.

Session 2 One week later, participants returned for Session 2. They were given the one-week Knowledge, Type, and Feature quizzes (which were the same as the pre- and post-game quizzes), and they could again ask the researcher for a fun fact about any of the 24 rocks in the quizzes. They were subsequently interviewed about whether they had thought about, or had any questions come to mind about rocks or the robot during the preceding week, and whether they had looked up any information related to either. We also wanted to know what they remembered from their time with the robot. Lastly, participants were debriefed on the study.

4 ANALYSIS

Data Preparation

There were several sources of ambiguity in the quizzes as a result of being filled out on paper. In the Feature quizzes, participants on a few rare occasions marked both a feature and 'I don't know'. In this case, we took the marked feature as the final answer. We discarded those answers when participants marked both a feature and 'None of the Above'. We discarded the data of one participant (C8) in the Curiosity condition entirely, since it was unclear whether the participant understood the procedure of the experiment. In total, we retained the data of 29 participants for analysis.

The audio of the game was transcribed and using the theoretical framework of curiosity [38], ten verbal curiosity-driven behaviours were coded: (1) *Uncertainty:* Lack of sureness about something or someone, (2) *Argument:* Reasons, statements, or facts supporting or establishing a view point, (3) *Justification:* Providing information to make something clearer, (4) *Suggestion:* An idea or plan suggested by one group member, usually to get the other to do something, (5) *Agreement:* Group members' opinions or feelings are in harmony, (6) *Disagreement:* The opposite of agreement, (7) *Question Asking:* (on task vs. social) Any questions that are asked, not related to logistics of the game or task, (8) *Idea Verbalization:* Communication of an idea, (9) *Sharing Information:* Verbally communicating results, findings, or discoveries to the other player, and (10) *Hypothesis Generation:* Providing theories to explain something. Each statement/question could be associated with multiple codes.

The Free Choice Curiosity Test was designed to measure whether elicited curiosity was directed towards the robot or the rocks. The audio of the free period was transcribed, and both the audio and video were coded with a rank assigned to each rock to indicate the order in which participants chose to talk about them. The premise was that participants would choose the rocks that they were most interested in first.

Measures

We collected a variety of data about participants' demographics, curiosity, perception of and experience with the robot, behaviour in the game and free period, and learning. In this paper, we focus on those measures that provide, primarily, insights into participants' mental state of curiosity and curiosity-related behaviour, and secondarily, into learning.

The dependent variables include demographics information, perception of the robot (e.g., Godspeed measures, pick-a-mood scale), trait curiosity (e.g., curiosity and exploration inventory, growth mindset survey) as well as intrinsic motivation. Table 2 briefly summarizes all the variables we have considered in our analysis along these dimensions. Only a few of the variables are found to be significant through the step-wise selection process; these variables and their associated results are reported in the findings.

5 RESULTS

Perception of Robot Curiosity

During the Session 1 interview, participants were asked to rate on a 7-point Likert scale how curious they thought the robot was from 1 (not at all curious) to 7 (extremely curious). Table 3 shows the average rating for each condition. Participants in the Curious and Curious+Reveal conditions perceived the robot to be more curious than participants in the Neutral condition. The proportional odds model [1] showed that the robot's curiosity was ranked significantly differently by the participants in the different conditions ($\chi^2(2, N = 28) = 7.46, p = 0.02$). This demonstrates that our design created a robot that behaves in a manner considered curious by participants.

Table 3: Perception of Robot Curiosity about Rocks

Condition	Average Rating (SD)
Neutral	3.78 (2.33)
Curious	5.89 (0.78)
Curious+Reveal	5.90 (0.97)

Approximately half of the participants in the curious conditions, and one participant in the Neutral condition, described the robot as 'curious', 'inquisitive', or having a desire to learn—before being asked specifically by the interviewer about the robot's curiosity. The majority of participants in the Curious (C) and Curious+Reveal (CR) conditions stated that it was the question-asking behaviour which indicated to them that the robot was curious: "The robot actually actively asked me for more knowledge" (C1), and, "He's very curious because he kept asking a lot of questions" (CR6). The curious robots were perceived to be curious not only about rocks, but also about the participants' state of mind. Participant C2 said, "The robot seemed curious about my opinions on things. So he would make a connection and say 'What do you think?'". Participants also noticed a link between asking questions and a desire for new knowledge: "[The robot] wants to know more" (C7); "He kind of always wanted to learn about, and go a step further to learn a little bit more about the rock. He was kind of trying to go deeper" (CR5).

Rather than asking questions, participant C5 found the robot to be curious because "when we were playing the game, I was moving to the next card, and he [was] still talking about the rock". Curiosity was also perceived from the fact that "he always notices little details about [the rocks], and connects it to what he knows and tries to share with me" (CR2).

In contrast, the participant who rated the robot's curiosity the lowest in the Curious+Reveal condition, felt that "if NAO were to be curious, then it would ask me different kinds of questions every time" (CR4). The participant was referring to the robot asking about the participant's rocks features, type, or name, after every time the rocks linked: "I felt that if you ask the same thing again and again...it's more robotic".

Unexpectedly, when asked: "Do you think [the robot] was curious or not curious about rocks during the game?", 6 out of the 10 participants in the Neutral (N) condition felt that it was curious. However, on average, they gave a significantly lower Likert scale rating of the robot's curiosity (Table 3), implying that participants in this condition did not think it was as curious as participants who interacted with the curious robots. We discovered that, in the absence of the robot asking questions to know more about rocks, participants picked up on cues that we had not intended to convey curiosity through: "He was curious because he *knew* a lot about [the rocks]" (N1); "NAO is really *interested* in playing the game... really pushing the game to get it going" (N2); and "He would tell me random facts" (N3). Non-verbal cues of curiosity were also noticed by some participants, i.e., "He tilts his head" (N3), and, "I think he's actively looking for certain traits when you give him a rock" (N7).

Additionally, two participants mentioned that the robot asking questions related to game play indicated curiosity to them, e.g.: "I think he was curious about what I would do, like [he would say] 'Do you want to go first?'" (N3). The remaining participants in the Neutral condition found that the robot was not curious specifically because "he didn't ask me questions" (N4). Participant N9 explained: "To me curiosity comes from wanting to know something....So the way he sounds is curious, but I think that's from the enthusiasm but in terms of actual curiosity maybe not". The observation about question-asking by some participants in the Neutral condition, further supports that our design of verbal question-asking behaviour by the robot can be used to convey curiosity.

Participant's Curiosity

About Rocks. In the demographics questionnaire (pre-study) we asked participants to rate their level of interest in rocks on a 7-point Likert scale from 1 (not at all) to 7 (very). 17% of participants were not at all interested in rocks, 73% were interested, and 10% were very interested. During the interviews in Session 1 (post-interaction) and Session 2 (one-week), in addition to rating the robot's curiosity, participants were asked to rate their *own* curiosity about rocks, also on a 7-point Likert Scale, from 1 (not at all curious) to 7 (extremely curious). Table 4 shows the average rating in each condition at pre-study, post-interaction, and after one-week. The differences in pre-study curiosity on rocks between the conditions were significant. We controlled for this bias by comparing the magnitude of change in curiosity across the three time points, between conditions. Compared to the other conditions, participants in the Curious condition experienced a greater change (from pre-study) in curiosity about rocks after the game ($F(2, 26) = 4.12, p = 0.03$) and after one week ($F(2, 26) = 5.47, p = 0.01$).

Table 4: Average Participant Curiosity about Rocks (SD)

Condition	Pre-Study	Post-Interaction	One-Week
Neutral	4.00 (1.05)	4.90 (1.35)	4.90 (1.45)
Curious	3.11 (1.05)	4.61 (1.32)	3.67 (1.00)
Curious+Reveal	4.50 (1.27)	4.50 (1.43)	3.75 (1.55)

For the requested fun facts, we ran a Poisson regression model with Neutral as the baseline category, to which Curious and Curious+Reveal were compared. Participants in the Curious ($\hat{\beta} = 2.09, t(364) = 2.51, p = 0.01$) and Curious+Reveal ($\hat{\beta} = 1.21, t(364) = 1.96, p = 0.05$) conditions tended to ask to learn more fun facts about the rocks the robot had commented on during the game. After one week, the conditions did not appear to be significantly different anymore, $\hat{\beta} = 1.45, t(364) = 1.55, p = 0.12$ (Curious) and $\hat{\beta} = -0.46, t(364) = -0.67, p = 0.50$ (Curious+Reveal). However, there is evidence that the more questions asked by the robot during the game about the participant's rocks type, the lower the probability that a participant would inquire for a fun fact about that particular rock a week later, $\hat{\beta} = -0.22, t(364) = -2.06, p = 0.04$.

Analysis of the Free Choice Curiosity Test, under the Kruskal-Wallis test, showed that the rank order for all rocks was similar in all conditions, suggesting that participants did not prefer the rocks with the ground truth over those without, and vice versa. However, in addition to coding rank order, participants were asked to explain in the interview what they did during the free period and why. Only participants in the curious conditions described their behaviour as being solely about rocks (five participants in the Curious condition and five in the Curious+Reveal condition). Common explanations for behaviour were: using the robot to clarify or get advice on distinguishing rock types, to get more knowledge on rocks, or using the robot to test whether their own knowledge was correct. All participants in the Neutral condition stated their behaviour was either directed solely towards the robot, or both the rocks and the robot.

In Session 2, participants were asked in the interview whether they had had any questions or had thought about rocks at all since Session 1, and whether they had looked up any information on the topic. Seven participants in the Curious condition and six in the Curious+Reveal condition, compared to only three in the Neutral condition, stated that they had questions about rocks, and three participants in each condition explained they had actively searched for information about rocks. Participants in the Neutral and Curious+Reveal conditions who said they looked up more information, were mostly interested in clarifying their understanding between sedimentary and metamorphic rocks, because they were unsure of how to categorize those types. In contrast, participants in the Curious condition stated that they searched for information because, "it's interesting" (C4), and, "I was just thinking about it and just wonder, I'm like more intrigued basically, because of the process" (C6).

About the Robot. Similar to ranking their curiosity about rocks, participants were asked in the demographics questionnaire (pre-study) to rate their interest in robots on a 7-point Likert scale from 1 (not at all) to 7 (very). 3% of participants were not at all interested in robots, 67% were interested, and 30% rated themselves very interested. In the interview during Session 1 (post-interaction) and Session 2 (one-week), participants were asked to rate their level of curiosity about the robot on a 7-point Likert scale from 1 (not at all curious) to 7 (extremely curious). The average ratings per condition at each of the three time points is shown in Table 5. In contrast to curiosity about rocks, participants' curiosity about the robot, in all conditions, started relatively high, increased immediately after the game, and decreased one week later. However, these changes are not statistically different between conditions after the game ($F(2, 24) = 0.84, p = 0.44$) and one week later ($F(2, 21) = 0.23, p = 0.79$).

Table 5: Average Participant Curiosity about Robot (SD)

Condition	Pre-Study	Post-Interaction	One-Week
Neutral	4.90 (1.60)	6.11 (0.78)	5.60 (1.26)
Curious	4.89 (0.60)	6.50 (0.53)	5.36 (1.03)
Curious+Reveal	5.30 (0.95)	5.80 (1.81)	5.50 (1.32)

All but one of the participants rated their curiosity about the robot equal to, or higher than, their curiosity about rocks post-interaction. The participant who was not curious about the robot stated, "Well it's just a computerized robot, right?" (CR7). The other participants were mainly curious about how the robot worked, what else it could do, what else the robot knew, and whether it was learning. For example: "I'm pretty curious about NAO... [If I] phrased things differently or asked a question in a non-direct way...I couldn't stump her. That was cool" (CR8), and, "I was very curious about how he is identifying [the rocks]" (N10), but also, "I was curious about whether he'll build on, you know, different experiences. So after my trial, would it differ from someone else's?" (CR6). Participant C9 made an interesting observation, stating: "I think I'm a lot more interested in NAO. Like he has a *contagious* influence".

When asked to explain what they chose to do during the free period, only one participant in the Curious condition stated they were only focused on the robot. They explained: "I kind of tested him...how this robot can...like what's his intelligence [and] I specifically used the rocks with the cards to kind of see if he can correctly, precisely identify features" (C3). Across all conditions, when participants showed some curiosity towards the robot, most explained how they tested it to figure out what more it could do and what more it knew. Participant CR3 taught the robot the names of the rocks and then tested to see whether it could remember them. Participant N1 wanted to see whether the robot could distinguish between numbers and letters that looked similar. Another stated, "I wanted to know if [the robot] could do anything wrong. I wanted to challenge him... I wanted to see if he was gonna get mad" (N3).

In Session 2, the average rating for curiosity about the robot was approximately the same in each condition. Six people in each of the three conditions said they still had some questions pertaining to the robot, however very few searched for more information.

Participant's Verbal Behaviour

Participants' verbal behaviour was coded using the theoretical framework of curiosity [38]. As a validity test, two independent coders each coded half of the game transcript dialogues. On average, the inter-rater agreement (cohen kappa score) was around 0.74, indicating substantial agreement. The lowest agreement rate was 0.50 (i.e., moderate agreement) for Justification, and the highest was 0.85 (i.e., almost perfect agreement) for Question Asking.

Table 6: Average Number of Participant Responses by Type in Game (SD)

Condition	Neutral	Curious	Curious+ Reveal	p
Question Asking (QA)	1.30 (1.34)	9.67 (7.26)	3.40 (2.37)	***
On-Task QA	1.00 (1.25)	8.89 (7.79)	3.20 (2.49)	**
Social QA	0.30 (0.48)	0.78 (1.99)	0.20 (0.63)	
Uncertainty	0.40 (0.70)	8.44 (3.50)	6.00 (3.83)	***
Argument	0.00 (0.00)	0.00 (0.00)	0.10 (0.32)	
Justification	0.40 (0.52)	2.33 (2.24)	2.00 (2.16)	.
Suggestion	0.00 (0.00)	0.11 (0.33)	0.00 (0.00)	
Agreement	9.60 (4.93)	9.67 (2.12)	10.5 (4.53)	
Disagreement	0.00 (0.00)	0.67 (0.71)	1.00 (0.82)	**
Idea Verbalization	0.00 (0.00)	0.11 (0.33)	0.00 (0.00)	
Information Sharing	5.00 (4.35)	4.00 (3.84)	1.50 (1.65)	.
Hypothesis Generation	0.00 (0.00)	3.78 (1.30)	1.90 (2.38)	***

Table 6 shows how participants' responses differed by condition, with the statistically significant response types in bold. Our findings, using ANOVA, show significant differences in the frequency of certain types of participant responses between conditions. Participants in the curious conditions produced a greater number of responses in the game that can be categorized as expressing curiosity, compared to the Neutral condition. In particular, there was more question-asking, specifically on-task question-asking ($F(2, 26) = 7.27, p = 0.003$), in the Curious and Curious+Reveal conditions compared to the Neutral condition. There are also statistically significant differences in responses expressing uncertainty ($F(2, 26) = 18.15, p < 0.001$), hypothesis generation ($F(2, 26) = 13.65, p < 0.001$), and disagreement ($F(2, 26) = 6.72, p = 0.004$). Across all conditions, there was very high verbal agreement with the robot. This is most likely a consequence of the participants not having enough knowledge on rocks to disagree Disagreement with the robot, although low, was significantly higher in the Curious+Reveal condition—possibly a result of the robot providing rationale, and giving participants more explanation with which they could disagree.

Other Measures

Learning. For the Knowledge and Type quizzes, the score is 1 if the answer was correct for a particular question/rock, zero otherwise. For the Feature quiz, as the answer could include a number of features, we computed 5 different measures—precision, recall, F1, specificity, and accuracy, to capture both how many features participants correctly identified and how many features they missed. Condition had no effect on the Knowledge, Type, and Feature quiz scores, i.e., learning; nor did it have an effect one week later, i.e., retention. However, through step-wise linear regression, we did find that participants who reported being more curious about rocks after the game had higher Knowledge quiz scores a week later, $\hat{\beta} = 0.32, t(26) = 2.34, p = 0.03$.

Results show that participants performed much better on determining the type of rocks they saw in the game than the rocks not in the game, as reflected in the Type quiz scores immediately after the game ($F(2, 83) = 22.64, p < 0.001$) and one week later ($F(2, 83) = 20.69, p < 0.001$). Similarly, there was significantly different performance for game rocks versus non-game rocks for both the post and one-week Feature identification quiz. These results show that the LinkIt! game has some success in teaching participants how to identify features in rocks and determining the rock type.

To investigate the measures that impact the number of questions participants answered correctly, different Binomial regression models were used for the post and one-week Knowledge quiz, respectively. Through step-wise selection, the resulting models show that participants' initial number of correct answers significantly influence the number of correct answers in the post Knowledge quiz ($\hat{\beta} = -0.37, t(27) = 2.24, p = 0.03$) but not the one-week Knowledge quiz ($\hat{\beta} = -0.13, t(25) = -0.27, p = 0.79$).

Perception of the Robot - General. There were no significant differences between condition in the four Godspeed measures: anthropomorphism, animacy, likeability and perceived intelligence.

Perception of the Robot - Gender. The robot's voice was child-like, without any explicit manipulation of pitch to indicate gender, and the researcher conducting the sessions referred to the robot with the pronoun 'he'. However, other cues appeared to influence perception of the robot's gender, as all participants in the Neutral condition used 'he', but two participants in the Curious condition and three in the Curious+Reveal condition, referred to the robot as 'she', while the rest used 'he' or 'it/the robot'.

Intrinsic Motivation Inventory. We found no statistical significance between different conditions in terms of participants' self reported measures of enjoyment, competence, effort and relatedness on the Intrinsic Motivation Inventory.

6 DISCUSSION

In this work, we present and use a novel structured game, LinkIt!, for the design and study of robot behaviour and human-robot interaction, and introduce a procedure for assessing curiosity through behaviour, called the Free Choice Curiosity Test. Our findings show that verbally-expressed curiosity through on-task, topic-directed question-asking, can reliably be recognized as curiosity in a social peer robot. Additionally, we found evidence for curiosity contagion effects; i.e., the robot's verbal expression of curiosity about rocks increased the participants' emotional *and* behavioural curiosity about rocks. Participants in the Curious condition reported a greater increase in curiosity about rocks after the game, and participants in the Curious and Curious+Reveal conditions tended to ask to learn more fun facts about rocks that the robot had commented on. In Session 2, more participants in the curious conditions than in the Neutral condition had questions about rocks and were interested in gathering new information about rocks. Furthermore, results from the Free Choice Curiosity Test indicate that the curious robots were more effective at directing participants' curiosity towards the rocks than the Neutral robot.

As with Gordon et al.'s findings [17], our results show behavioural curiosity contagion effects as a result of interacting with a curious robot. Different from Gordon et al., our study focused on the verbal behaviour of participants interacting with curious robots, rather than curiosity-driven exploration and uncertainty seeking. Using the theoretical framework of curiosity [38] to code dialogue, we found a significant increase in on-task question-asking by participants in the Curious condition. This result could indicate the *chameleon effect* [11], the unconscious mimicry of behaviour–a well-documented phenomenon in human social interactions. Our study shows that within a 30 minute interaction, students appear to mimic the most pronounced verbal behaviour of a robot dialogue partner. In addition, beyond on-task question-asking, participants in the Curious condition also more frequently expressed other curiosity-driven behaviours such as uncertainty and hypothesis generation in their conversation with the robot, indicating emotional curiosity contagion effects as well as behavioural.

Our findings show significant differences in the frequency of certain types of participant responses between conditions, which can be the result of the dynamics of the conversation itself. For example, question-asking by the curious robots may have provided more opportunities for participants to generate hypotheses, than the neutral robot that only shared information. We leave detailed conversational analysis to future work, however note that verbal mimicry of question-asking, did not occur significantly more in the Curious+Reveal condition, even though this robot also asked questions. Furthermore, participants in this condition did not rate their curiosity about rocks higher after playing the game. Tsai et al. [42] propose that emotional contagion effects may be hindered by cognitive workload. The addition of rationale in the Curious+Reveal condition may have been too informationally overwhelming for participants, leading to neither behavioural *nor* emotional curiosity contagion—indicating that the mimicking of question-asking behaviour by participants in the Curious condition may have aided in emotional curiosity contagion.

Similar to prior work [17], we found that the manipulations had no effect on learning; neither immediately after the interaction nor one week later. This remains an interesting finding as our study considered a different age group, interaction protocol, and expression of curiosity. The result may be due to the novelty effect—an increase in performance when interacting with a new technology, as a result of *interest* in the new technology rather than an actual improvement in learning. For all participants, interacting with a social humanoid robot was a novel experience, and many expressed surprise and curiosity about the robot. In the future, this effect could be mitigated by having participants interact with the robot over multiple sessions and measuring their overall learning performance. We did however, find that participants who reported being more curious about rocks after the interaction performed better on the Knowledge quiz one week later, indicating that curiosity elicited in Session 1 may have helped retain knowledge for Session 2.

A few participants noticed themselves become curious about rocks as a consequence of learning more about them. This supports the information gap theory of curiosity [27], which postulates that curiosity originates from a gap between what one knows and what one wants to know. Once people become aware of this gap, their curiosity is piqued. It appears that the question-asking by the curious robots acted to bring attention to the information gap and make participants aware of its existence, and in this way elicited curiosity. It also supports the Learning Progress hypothesis [30], which suggests that the experience of learning causally influences curiosity and an intrinsic motivation for new knowledge.

The robot in this study was naive; it only knew about what it saw in the introductory video and what it learned from the game. As a result, even when participants became aware of their own knowledge gaps and were asking the robot for answers, the robot was not always able to provide them. Previous research has shown that when people do not expect to close their knowledge-gap quickly, not knowing can affect their subjective experience of curiosity [29]. Similarly, Shiomi et al. [36] found that students were motivated to ask more questions when they *knew* the robot could answer them. In future studies, giving the robot the ability to answer participants' questions, or providing participants with material they could search to resolve their information gaps, may lead to further increases in curiosity.

The novel Free Choice Curiosity Test did not show significant differences in how participants chose rocks to present to the robot. Instead, we gained insights into participants' curiosity from their explanations—indicating a need for development of how the test is conducted, as well as analyzed, in order for it to be an effective objective measure of curiosity.

Limitations

A limitation of our study is the relatively small sample size in each condition and the fact that all participants are university students. This makes it difficult to generalize the results to larger or different populations. To handle small sample size, we avoided including more covariates in the ANOVA and selected the simplest possible model to explain our data in order to prevent overfitting, while providing some general directions for future research involving a larger sample.

Both small sample size and large individual differences may have contributed to the lack of significant differences between conditions in certain measures. In our analysis, we considered individual differences by using curiosity trait measures as independent variables, however, did not find these to be significant. In future work, we aim to run the study with younger students and include more participants. A within-subject experimental design may be adopted to combat large individual differences, and other traits could be measured that are known to influence learning and social interactions, such as executive, emotional, and social functioning, as well as learning goals (e.g., [26, 34]).

Furthermore, we did not code non-verbal indicators of curiosity during the game. In the future, video could be coded for such indicators to provide a more complete picture of emotional and behavioural curiosity contagion. For example, one possibility is to use facial landmark detection to analyze facial action units to infer affective states that often occur with curiosity as Sinha, Bai, and Cassell [38] have done with human-human social interaction.

Design Implications

Our findings indicate that care must be taken in the design of robot dialogue to prevent high cognitive workload in the conversational partner. Furthermore, curiosity can be inadvertently communicated through cues other than question-asking (e.g., body language, task-relevant behaviours, etc.). Lastly, more research can be done on behavioural measures of curiosity, such as the Free Choice Curiosity Test.

7 CONCLUSION

In this work, we contribute to the understanding of social curiosity contagion with a robot. We show that emotional and behavioural curiosity contagion can occur in university students, and that question-asking is a curiosity-driven behaviour which students mimic from the robot. Our work contributes a new structured game, LinkIt!, for the design and study of human-robot behaviour and its impact on curiosity, as well as a novel procedure for assessing curiosity through behaviour.

ACKNOWLEDGMENTS

This work was made possible by funding from NSERC Discovery Grant RGPIN-2015-0454.

REFERENCES

[1] Alan Agresti. 1990. *Categorical data analysis.* Wiley, New York [u.a.]. XV, 558 S. pages.

[2] Marilyn P. Arnone, Ruth V. Small, Sarah A. Chauncey, and H. Patricia McKenna. 2011. Curiosity, interest and engagement in technology-pervasive learning environments: a new research agenda. *Educational Technology Research and Development* 59, 2 (2011), 181–198.

[3] 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.

[4] Christoph Bartneck, Dana Kulić, Elizabeth Croft, and Susana Zoghbi. 2009. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International Journal of Social Robotics* 1, 1 (2009), 71–81.

[5] Tony Belpaeme, James Kennedy, Aditi Ramachandran, Brian Scassellati, and Fumihide Tanaka. 2018. Social robots for education: A review. *Science Robotics* 3, 21 (2018), eaat5954.

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

[7] Timothy Bickmore, Daniel Schulman, and Langxuan Yin. 2010. Maintaining engagement in long-term interventions with relational agents. *Applied Artificial Intelligence* 24, 6 (2010), 648–666.

[8] Gautam Biswas, Krittaya Leelawong, Daniel Schwartz, Nancy Vye, and The Teachable Agents Group at Vanderbilt. 2005. Learning by teaching: A new agent paradigm for educational software. *Applied Artificial Intelligence* 19, 3-4 (2005), 363–392.

[9] Gautam Biswas, James R. Segedy, and Kritya Bunchongchit. 2016. From design to implementation to practice a learning by teaching system: Betty's brain. *International Journal of Artificial Intelligence in Education* 26, 1 (2016), 350–364.

[10] Cynthia Breazeal and Rodney Brooks. 2005. Robot emotions: A functional perspective. In *Who Needs Emotions?: The Brain Meets the Robot*, Jean-Marc Fellous and Michael A. Arbib (Eds.). Oxford University Press, 271–310.

[11] Tanya L. Chartrand and John A. Bargh. 1999. The chameleon effect: the perception–behavior link and social interaction. *Journal of Personality and Social Psychology* 76, 6 (1999), 893.

[12] Kai-Yi Chin, Zeng-Wei Hong, and Yen-Lin Chen. 2014. Impact of using an educational robot-based learning system on studentsâĂŹ motivation in elementary education. *IEEE Transactions on Learning Technologies* 7, 4 (2014), 333–345.

[13] 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.

[14] Pieter M. A. Desmet, Martijn H. Vastenburg, and Natalia Romero. 2016. Mood measurement with Pick-A-Mood: review of current methods and design of a pictorial self-report scale. *Journal of Design Research* 14, 3 (2016), 241–279.

[15] David Duran. 2017. Learning-by-teaching. Evidence and implications as a pedagogical mechanism. *Innovations in Education and Teaching International* 54, 5 (2017), 476–484.

[16] Rachel Gockley, Jodi Forlizzi, and Reid Simmons. 2006. Interactions with a moody robot. In *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-Robot Interaction*. ACM, 186–193.

[17] 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*. ACM, 91–98.

[18] Matthias J. Gruber, Bernard D. Gelman, and Charan Ranganath. 2014. States of curiosity modulate hippocampus-dependent learning via the dopaminergic circuit. *Neuron* 84, 2 (2014), 486–496.

[19] 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.

[20] Min Jeong Kang, Ming Hsu, Ian M. Krajbich, George Loewenstein, Samuel M. McClure, Joseph Tao-yi Wang, and Colin F. Camerer. 2009. The wick in the candle of learning: Epistemic curiosity activates reward circuitry and enhances memory. *Psychological Science* 20, 8 (2009), 963–973.

[21] Todd B. Kashdan, Matthew W. Gallagher, Paul J. Silvia, Beate P. Winterstein, William E. Breen, Daniel Terhar, and Michael F. Steger. 2009. The curiosity and exploration inventory-II: Development, factor structure, and psychometrics. *Journal of Research in Personality* 43, 6 (2009), 987–998.

[22] 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.

[23] Kwan Min Lee, Wei Peng, Seung-A Jin, and Chang Yan. 2006. Can robots manifest personality?: An empirical test of personality recognition, social responses, and social presence in human-robot interaction. *Journal of Communication* 56, 4 (2006), 754–772.

[24] Daniel Leyzberg, Samuel Spaulding, Mariya Toneva, and Brian Scassellati. 2012. The physical presence of a robot tutor increases cognitive learning gains. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, Vol. 34.

[25] Brian Y. Lim, Anind K. Dey, and Daniel Avrahami. 2009. Why and why not explanations improve the intelligibility of context-aware intelligent systems. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2119–2128.

[26] Jordan A. Litman. 2008. Interest and deprivation factors of epistemic curiosity. *Personality and individual differences* 44, 7 (2008), 1585–1595.

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

[28] Bonnie M. Muir. 1994. Trust in automation: Part I. Theoretical issues in the study of trust and human intervention in automated systems. *Ergonomics* 37, 11 (1994), 1905–1922.

[29] Marret K. Noordewier and Eric van Dijk. 2017. Curiosity and time: from not knowing to almost knowing. *Cognition and Emotion* 31, 3 (2017), 411–421.

[30] Pierre-Yves Oudeyer, Jacqueline Gottlieb, and Manuel Lopes. 2016. Intrinsic motivation, curiosity, and learning: Theory and applications in educational technologies. In *Progress in Brain Research*. Vol. 229. Elsevier, 257–284.

[31] André Pereira, Iolanda Leite, Samuel Mascarenhas, Carlos Martinho, and Ana Paiva. 2010. Using empathy to improve human-robot relationships. In *International Conference on Human-Robot Personal Relationship*. Springer, 130–138.

[32] Project for Education Research That Scales (PERTS). 2015. Growth Mindset Assessment. https://survey.perts.net/share/toi

[33] Thomas G. Reio Jr. and Albert Wiswell. 2000. Field investigation of the relationship among adult curiosity, workplace learning, and job performance. *Human Resource Development Quarterly* 11, 1 (2000), 5–30.

[34] Dale H. Schunk, Paul R. Pintrich, and Judith L. Meece. 2008. Motivation in education: Theory, research, and applications. (2008).

[35] Prachi E. Shah, Heidi M. Weeks, Blair Richards, and Niko Kaciroti. 2018. Early childhood curiosity and kindergarten reading and math academic achievement. *Pediatric Research* (April 2018). https://doi.org/10.1038/s41390-018-0039-3

[36] Masahiro Shiomi, Takayuki Kanda, Iris Howley, Kotaro Hayashi, and Norihiro Hagita. 2015. Can a social robot stimulate science curiosity in classrooms? *International Journal of Social Robotics* 7, 5 (2015), 641–652.

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

[38] Tanmay Sinha, Zhen Bai, and Justine Cassell. 2017. A new theoretical framework for curiosity for learning in social contexts. In *European Conference on Technology Enhanced Learning*. Springer, 254–269.

[39] Aimee E. Stahl and Lisa Feigenson. 2015. Observing the unexpected enhances infants' learning and exploration. *Science* 348, 6230 (2015), 91–94.

[40] Fumihide Tanaka, Kyosuke Isshiki, Fumiki Takahashi, Manabu Uekusa, Rumiko Sei, and Kaname Hayashi. 2015. Pepper Learns Together with Children: Development of an Educational Application. In *Humanoid Robots (Humanoids)*. IEEE, 270–275.

[41] Fumihide Tanaka and Shizuko Matsuzoe. 2012. Children teach a care-receiving robot to promote their learning: Field experiments in a classroom for vocabulary learning. *Journal of Human-Robot Interaction* 1, 1 (2012), 78–95.

[42] Jason Tsai, Emma Bowring, Stacy Marsella, Wendy Wood, and Milind Tambe. 2012. A study of emotional contagion with virtual characters. In *Intelligent Virtual Agents. IVA 2012. Lecture Notes in Computer Science*, Yukiko Nakano, Michael Neff, Ana Paiva, and Marilyn Walker (Eds.). Vol. 7502. Springer, Berlin, Heidelberg, 81–88.

[43] Robert J. Vallerand, Marc R. Blais, Nathalie M. Brière, and Luc G. Pelletier. 1989. Construction et validation de l'échelle de motivation en éducation (EME). *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement* 21, 3 (1989), 323.

[44] Robert J. Vallerand, Luc G. Pelletier, Marc R. Blais, Nathalie M. Briere, Caroline Senecal, and Evelyne F. Vallieres. 1992. The Academic Motivation Scale: A measure of intrinsic, extrinsic, and amotivation in education. *Educational and Psychological Measurement* 52, 4 (1992), 1003–1017.

[45] Junchao Xu, Joost Broekens, Koen Hindriks, and Mark A. Neerincx. 2014. Robot mood is contagious: effects of robot body language in the imitation game. In *Proceedings of the 13th International Conference on Autonomous Agents and Multiagent Systems*, Alessio Lomuscio, Paul Scerri, Ana Bazzan, and Michael Huhns (Eds.). 973–980.