Toward Effective Robot–Child Tutoring: Internal Motivation, Behavioral Intervention, and Learning Outcomes

2

ADITI RAMACHANDRAN, Yale University
CHIEN-MING HUANG, Johns Hopkins University
BRIAN SCASSELLATI, Yale University

Personalized learning environments have the potential to improve learning outcomes for children in a variety of educational domains, as they can tailor instruction based on the unique learning needs of individuals. Robot tutoring systems can further engage users by leveraging their potential for embodied social interaction and take into account crucial aspects of a learner, such as a student's motivation in learning. In this article, we demonstrate that motivation in young learners corresponds to observable behaviors when interacting with a robot tutoring system, which, in turn, impact learning outcomes. We first detail a user study involving children interacting one on one with a robot tutoring system over multiple sessions. Based on empirical data, we show that academic motivation stemming from one's own values or goals as assessed by the Academic Self-Regulation Questionnaire (SRQ-A) correlates to observed suboptimal help-seeking behavior during the initial tutoring session. We then show how an interactive robot that responds intelligently to these observed behaviors in subsequent tutoring sessions can positively impact both student behavior and learning outcomes over time. These results provide empirical evidence for the link between internal motivation, observable behavior, and learning outcomes in the context of robot–child tutoring. We also identified an additional suboptimal behavioral feature within our tutoring environment and demonstrated its relationship to internal factors of motivation, suggesting further opportunities to design robot intervention to enhance learning. We provide insights on the design of robot tutoring systems aimed to deliver effective behavioral intervention during learning interactions for children and present a discussion on the broader challenges currently faced by robot–child tutoring systems.

CCS Concepts: • **Applied computing** → *Interactive learning environments*; • **Human-centered computing** → *User studies*;

Additional Key Words and Phrases: Child–robot interaction, tutoring, motivation, education

ACM Reference format:
Aditi Ramachandran, Chien-Ming Huang, and Brian Scassellati. 2019. Toward Effective Robot–Child Tutoring: Internal Motivation, Behavioral Intervention, and Learning Outcomes. *ACM Trans. Interact. Intell. Syst.* 9, 1, Article 2 (February 2019), 23 pages.
https://doi.org/10.1145/3213768

The reviewing of this article was managed by associate editor LaViola, Joseph.
Part of this work was published in Proceedings of HRI'16 [57]. This work is supported by the National Science Foundation, grant no. 1139078 (Socially Assistive Robotics).
Authors' addresses: A. Ramachandran and B. Scassellati, Computer Science Department, Yale University, 51 Prospect St, New Haven, CT 06511; emails: {aditi.ramachandran, brian.scassellati}@yale.edu; C. M. Huang, Computer Science Department, Johns Hopkins University, 3400 N. Charles St, Baltimore, MD 21218; email: cmhuang@cs.jhu.edu.

2160-6455/2019/02-ART2 $15.00
https://doi.org/10.1145/3213768

1 INTRODUCTION

Advancing personalized learning, or the shift from traditional *one-size-fits-all* teaching to learning environments tailored to the needs of the individual student, has great potential to bolster conventional models of education and positively impact educational outcomes for children with various abilities and capacities [48, 49]. One aspect of a learner that varies significantly between children is a student's motivation [22, 25]. Differences in motivational goals and tendencies can notably influence students' academic performance and learning outcomes [25, 80], indicating that a child's motivation in learning plays a large role in the type of personalized support the child may require [48]. Motivation is a complex construct that involves various factors driving someone to engage in certain behaviors. For younger learners this can be measured in terms of externally and internally motivated reasons for why they do certain school-related activities [22]. Externally driven motivation involves engaging in activities in order to receive rewards or avoid consequences that are specifically external to the person, such as money or getting in trouble with a teacher. Internally driven motivation refers to engaging in various behaviors to achieve internal satisfaction or act in accordance with one's own goals or values, such as doing work because it is enjoyable or practicing a skill due to a belief that it is important to learn. Internal motivation includes *intrinsic* motivation, which is defined as doing something because it is inherently interesting or enjoyable and reflects a person's highly internal satisfaction in engaging in an academic task [62]. Intrinsic motivation is positively associated with self-regulation and academic success [21, 80]. Students who are intrinsically motivated are often considered to be self-regulated learners, as they typically apply strategies to monitor and evaluate their own learning processes [78, 79]. While intrinsic motivation is important for academic success, there are also other categories of motivation, such as *identified* and *introjected* motivation, that stem from an inner acceptance of the value or utility of a task, making these types of internal motivation positive sources of motivation that are also useful in fostering productive learning [62]. These internal categories of motivation have also been shown to correspond to positive coping strategies with failure experience in an academic setting [61].

Robot tutoring systems, capable of engaging in embodied interactions and maintaining situated awareness of the learning environments and users, hold the promise to deliver effective personalized learning [13]. To fully support personalized learning, it is crucial that these systems consider students' motivation as part of their personalization approaches. However, it is challenging to decipher a person's motivation in learning, which involves hidden factors and processes that cannot be acquired intuitively and directly via computing technologies. While it is difficult to directly access the internal information of a person's motivation, human behavior often reveals information about a person's internal states, including information about attention, emotion, and motivation. Therefore, in practice, we must explore robot tutoring systems that monitor and shape students' observable, manifested behavior without direct access to their internal states. Identifying how such observable behaviors are linked to motivation in learning can help develop a holistic understanding of the learning process and can inform the design of effective robot tutoring systems. We aim to understand whether robot tutoring systems can leverage this understanding in order to provide intelligent intervention behaviors that have the potential to positively impact student learning outcomes.

In this article, we describe the linkage of motivation in learning, observable behavior, and learning outcomes, highlighting how motivation closely relates to behavior within a learning interaction, which, in turn, can be used intelligently by a robot tutoring system to intervene accordingly to strengthen learning outcomes. In a user study involving children interacting with a robot tutoring system over multiple sessions (Figure 1), we first show how motivation in learning as assessed by a self-reported questionnaire directly relates to observed "suboptimal" help-seeking behaviors. We further demonstrate that a robot tutoring system responding intelligently to the occurrence of

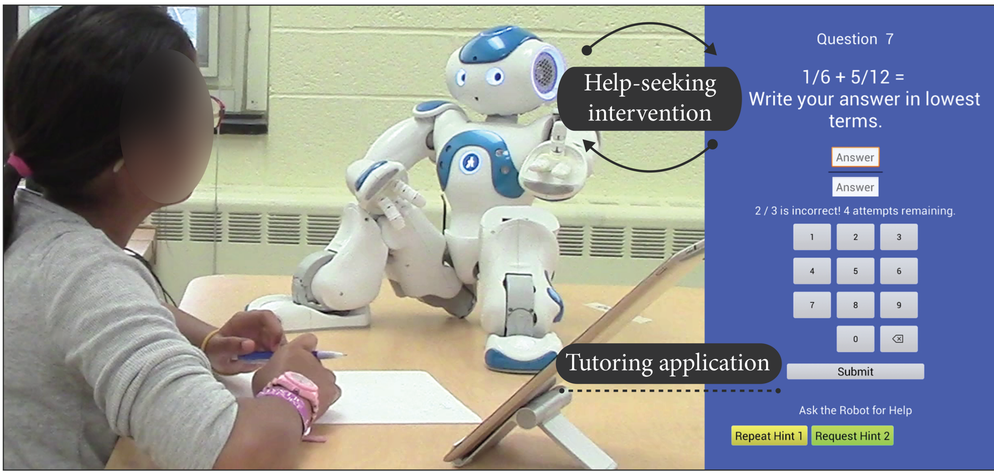

Fig. 1. We designed an interactive robot tutoring system to provide help to children practicing math problems. Children interacted with the robot through a tutoring application on a tablet device and could ask the robot for help through buttons on the tablet screen. The robot system was designed to intervene intelligently based on the help-seeking behavior of the children. The students engaged in autonomous robot tutoring interactions, in which the experimenter did not intervene at any point during the sessions.

these particular suboptimal behaviors can improve behavior and positively affect learning gains [57]. Our results inform future design of robot tutoring systems aiming to deliver effective intervention that supports a broader learning process that includes the role of motivation in learning.

This article provides the following contributions: We describe the unified linkage of motivation in learning, observable behavior, and learning outcomes, which can be used as a tool to design robot intervention behavior. We also empirically evaluate this tool within the context of a user study involving a robot–child tutoring scenario, demonstrating the potential for this tool to be used in building an effective tutoring system that can impact learning. We present empirical findings from our user study showing that a robot that employed shaping strategies to counter suboptimal help-seeking behaviors (requesting too much or too little help) relating to motivation improved the help-seeking behavior and learning gains of students over four tutoring sessions. These findings add to the broader knowledge of how motivation plays a role in robot–child tutoring interactions and informs how robots can leverage the relationship between motivation and behavior to improve behavior and enhance learning as effective tutors for children.

2 BACKGROUND

In this section, we first present related literature that discusses the role of motivation and behavior in a learning context, including examples of suboptimal behaviors in learning environments. We then review existing work on using intelligent tutoring systems (ITSs) to provide personalized learning and summarize the contributions of robots as effective tutoring agents.

2.1 Motivation, Behavior, and Learning Outcomes

Motivation in learning is a particularly complex construct that includes both internal (e.g., to achieve personal satisfaction or to act in accordance with one's values) and external (e.g., to avoid punishment) factors that dictate why children engage in certain academic behaviors. Different motivational profiles of students can affect reactions to successes and failures and can impact a child's cognitive performance [25]. Research in educational psychology has established a positive relationship between motivation and success in learning [21, 77, 80]. In particular, intrinsic motivation contributes crucially to a child's success in an academic environment [21]. Students who were motivated by learning goals, which reflects a more internal orientation toward learning, displayed cognitive achievement, particularly during challenging tasks [26]. Moreover, measures of children's internally oriented motivation have also been positively correlated with traditional achievement measures, such as standardized test scores and grades [29, 63]. Learning outcomes refer generally to measures of performance related to an educational task or program. These can refer to academic measures such as improvements in grades or test scores and are typically used to reliably demonstrate what a learner knows and does not know. To foster motivation in learning, prior work has explored the use of stimulating learning environments, open classrooms, and constructive feedback [21, 32, 44, 68]. These foundational efforts establish the impact of internal motivation on academic performance within the classroom and beyond.

Students who are intrinsically motivated tend to engage in self-regulated learning (SRL) by applying strategies to monitor their own learning processes and drive them forward [52, 53]. These strategies often involve employing metacognitive skills, such as goal-setting, adaptive help seeking, and persistence through difficulty [53, 78]. Prior research has explored the employment of these skills through the use of questionnaires and interviews with students and teachers and the association of them to academic achievement [53, 60, 80]. In addition to questionnaire and interview probing, more current efforts have made progress on automatically assessing user behavior through user computer traces and observation [5, 31, 79]. One prominent effort to capture SRL was by using a computer-based study environment that provided the learner with opportunities to make notes, search for information, receive help, and chat with fellow students and analyzing when and how frequently students engaged in these behaviors [76]. Other attempts to assess student use of SRL have relied on learning environments that record students thinking aloud or allow them to write down their thoughts during learning and then coding and categorizing these responses based on whether they demonstrate SRL processes [30, 66]. Another method of automatically evaluating student use of a specific SRL strategy—that is, help seeking—involved developing a model of appropriate help-seeking behavior and assessing student actions based on whether they were in line with the model [2]. This body of work has primarily focused on identifying productive behaviors and showing that using them positively impacts learning outcomes.

Identifying students who do not use these SRL strategies is critical, as they may require more tailored support from an interactive learning environment to be able to meet their learning goals. To this end, other work has focused on the complementary problem of identifying behavior productive to learning, which is identifying unproductive strategies that do not lead to more effective learning. This is particularly useful, as the successful identification of these unproductive behaviors provides an interactive tutoring system the opportunity to monitor suboptimal behavior, meaning behaviors that do not make effective use of the tutoring system, and potentially intervene accordingly. For example, ITS systems have explored this specifically in the context of help-seeking behavior, as productive help-seeking is a critical skill positively associated with SRL [27, 47, 59]. The existing ITS literature has shown that the use of exploitative help-seeking behavior ("gaming the system") and help-averse behavior (sometimes called "help avoidance") in students who interact with learning environments can negatively impact learning outcomes [1, 10]. We describe each of these behaviors and provide examples in the context of ITS.

Gaming the System. In the context of tutoring systems, gaming the system has been defined as "attempting to succeed in a learning environment by exploiting properties of the system rather than by learning the material and trying to use that knowledge to answer correctly" [8]. There are many examples of this behavior observed within intelligent tutoring systems. Two notable examples include inputting answers quickly and systematically and rapid hint requests [10]. Specifically related to help-seeking behavior, rapid hint requests typically involve the learner trying to acquire the answer or information about the answer without expending considerable effort thinking through each hint and the problem. In many of the intelligent tutoring systems in which gaming behavior has been detected, each problem will contain a series of hints; the last hint is often called a "bottom-out hint" because it contains very specific information that is necessary to solve the problem [58].

Help-aversion. Another noteworthy suboptimal behavior identified in tutoring environments that impacts learning is help aversion [1]. Help-averse behavior typically involves the lack of use of help features in a learning environment when it is likely to benefit the learner. In most interactive learning environments, there are help features built into the system, for example, in the form of a button on the screen, where the user can request help when needed. Help aversion is typically observed when the student makes many incorrect attempts but ignores the help button altogether, ultimately failing to use the help available from the tutoring system.

Because these suboptimal help-seeking behaviors demonstrate lack of use of SRL behavior, we believe that the use of unproductive help-seeking behavior has a relationship to a child's motivation in learning. In this work, we explore this relationship between motivation in learning and suboptimal help-seeking behavior and evaluate how these behaviors can impact learning outcomes through an empirical user study involving a robot tutoring system.

2.2 Personalization within Intelligent Tutoring Systems

The ITS community has made large contributions to the development of effective one-on-one tutoring environments. Successful work in ITS initially focused on building robust systems designed to capture the knowledge state or estimate skill level of students interacting with the system [7, 18, 23, 50]. Many systems could also successfully use this information to personalize a variety of aspects of the tutoring platform, including the types of exercises, feedback, and hints provided to the student [14, 23, 71]. Many of these learning systems have been shown to be effective in promoting learning and have largely influenced the broader discussion on the promise of personalized learning within educational settings [49, 72].

Work in ITS has also explored student modeling that goes beyond student skill estimation. More complex student models have been built that capture a broad variety of user characteristics, including meta-cognitive behaviors and affective states [23]. Many studies have explored the use of sensors to track affect and attention during learning [4, 17, 24]. There has also been a significant number of systems that have highlighted the importance of learner motivation and engagement during learning [9, 11, 12, 16, 20]. For example, Clement et al. demonstrate the importance of personalizing tutoring activity selection in a way that maintains learner motivation by providing a student with activities of the appropriate difficulty or challenge [16]. Additionally, ITSs have modeled students' usage of meta-cognitive strategies, such as help-seeking behaviors, effective use of the learning system, and problem-solving strategies [15, 40, 58, 73]. The inclusion of these user-centric attributes and internal states allows tutoring systems to model the broader learning process to better promote effective learning across a diverse population of learners. Our work explores the role of robots as physically embodied tutoring agents and seeks to understand whether robot tutors can be used to shape behaviors and promote learning gains over time.

2.3 Robot Tutoring Systems

Building on the promising results of ITS research, the field of human–robot interaction has focused on further advancing personalized learning by highlighting the efficacy of physically present social robots in various educational settings [37, 39, 67]. Social robots are physically embodied to afford effective human–robot interactions, leading to increased compliance and enjoyment as well as greater social and emotional support [6, 41, 51, 54, 55, 64, 74]. Kanda et al. show the establishment of a stronger relationship between the child and a robot that exhibits social behavior while also citing the positive effects of having a physically present entity [38]. Howley et al. describe the potential advantages of a robot tutor over a human tutor in some situations due to distinctions in social role and evaluation apprehension regarding seeking help within a learning interaction [33]. Moreover, Mohseni-Kabir et al. reinforce the idea that a robot dialogue can promote learning gains within the context of a bidirectional coaching model between an adult and a robot [46]. Ramachandran et al. also demonstrate that children show increased engagement and compliance with meta-cognitive strategy support during learning when the assistance is delivered by a physically present robot [55]. These studies show the promise of the use of physically present robots as effective tutoring agents specifically designed to interact with children and shape behavior over time.

There has also been substantial work establishing the benefits of personalized robot behavior or intervention on learning performance [28, 56]. Leite et al. modeled the affective reactions of children and then used a robot to adapt its social support based on the observed affective states [41]. Szafir and Mutlu monitored the attentional state of the user interacting with a robot and adaptively used this information to reengage the user when needed, which improved the information recall ability of the users [70]. Leyzberg et al. showed that providing personalized lesson support for a cognitive puzzle task significantly impacts learning gains [43]. More recent work in robot tutoring systems has directly investigated using robot tutors to foster SRL behaviors in children during learning owing to their positive impact on academic performance. Jones et al. have demonstrated the value of using open-learner models and scaffolding to specifically encourage students to build SRL skills while completing an educational task [35, 36]. After investigating the effectiveness of this robot system on student behavior over several weeks, students who interacted with the adaptive scaffolding robot tutor built more SRL skills than those who interacted with a robot tutor that did not use scaffolding [36]. These studies demonstrate how personalized robot systems that are being used for behavioral intervention can positively impact the learning interaction. While behavioral interventions presented by tutoring systems have shown success in shaping learning outcomes, how they are linked to measures of motivation in the context of a tutoring interaction is unclear. In this work, we seek to empirically explore the linkage of motivation in learning, observable behavior, and learning outcomes. Our results provide a holistic understanding of learning process and shed light on designing tutoring systems for effective behavioral intervention that leverages the influence of motivation in learning.

3 USER STUDY: SHAPING PRODUCTIVE HELP-SEEKING BEHAVIOR

To illustrate the linkage of motivation in learning, observable behavior within a learning environment, and learning gains (Figure 2), we here present empirical evidence from a user study involving children interacting with a robot tutoring system over multiple sessions [57]. We first introduce the context for and the robot tutoring system used in this user study. We then describe our experimental design, conditions, and procedure, as well as measures for evaluation and participants.

3.1 Study Context

We contextualized our study in a robot–child tutoring interaction involving four one-on-one sessions spanning approximately two weeks (Figure 3). The study was conducted in local elementary schools in Connecticut. Given the recent body of work discussed in Section 2.3 indicating that physical robots make promising tutoring agents owing to the increased engagement and compliance that they foster [6, 51, 54, 74], we chose a physically embodied robot to act as the tutor in

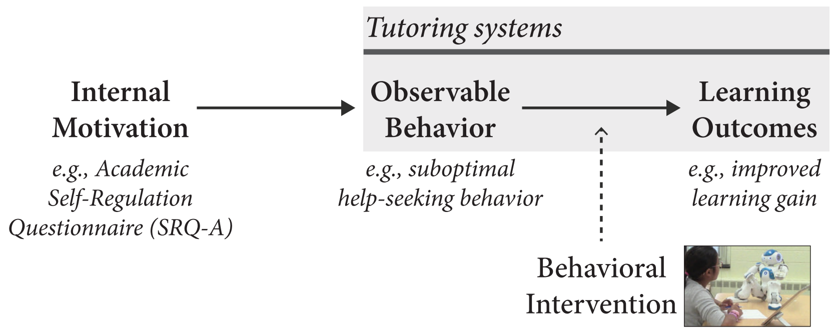

Fig. 2. Linkage between internal motivation, observable behavior, and learning outcomes. We use empirical evidence from a user study to establish a relationship between measures of internal motivation (as measured by a self-report questionnaire) and observable behaviors during a tutoring interaction with a robot. We further show that a robot tutoring system that provides behavioral intervention based on this observable behavior leads to learning gains.

our user study. Throughout the tutoring sessions, a robot acted as a tutoring agent that helped participating children solve math problems (Figure 1). In particular, the robot tutoring system was designed to provide hints as requested by the student on how to solve the math problems. It was not our goal in this work to isolate the benefits of using a physical robot. In this study, we sought to (1) understand how children's motivation in learning was related to their help-seeking behaviors and (2) evaluate how the robot's help-seeking intervention strategies may shape children's suboptimal help-seeking behaviors and subsequently influence their learning outcomes.

3.2 Robot Tutoring System

We built an interactive robot tutoring system consisting of a NAO robot to act as the tutoring agent and a tablet to display the tutoring application containing math problems (Figure 1). The math problems were displayed and completed on the tablet positioned in front of the students. Each of the four sessions contained eight math problems on fractions concepts. All problems followed state curriculum standards and were designed for students in fifth or sixth grade.

The robot provided verbal feedback on a child's answer to a math problem. For example, it used phrases such as "great job!" for a correct answer and "give it another shot!" for an incorrect answer. Additionally, the robot verbally provided hints at the child's request via the tutoring application on the tablet. Each problem had exactly three hints associated with it and the hints had to be requested in order although they could be repeated. Each successive hint provided more information; the third hint contained the most information relevant to the given problem.

The robot operated autonomously in real time with each child, requiring no input or intervention from the experimenter throughout the duration of each interaction. It reacted exclusively to tablet input from the child throughout the course of the tutoring sessions. The robot's behaviors were designed to be consistent across all participants over all sessions.

3.3 Experimental Design and Conditions

As described in Section 2.1, suboptimal help-seeking behaviors have been identified to impede effective learning with ITSs [1]. In this work, we aimed to understand how these behaviors are related to a student's motivation in learning and how a robot tutoring system can effectively intervene in response to these behaviors to improve learning outcomes. To this end, we employed the Academic Self-Regulation questionnaire (SRQ-A), detailed in Section 3.4.1, to measure a student's motivation in learning prior to the beginning of the first tutoring session. We designed

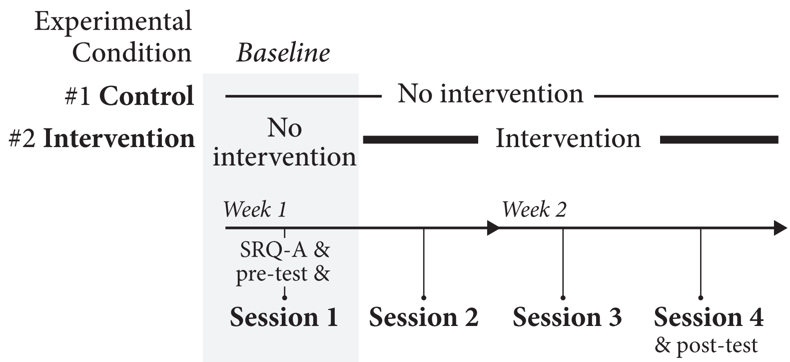

Fig. 3. Experimental design for the four-session robot–child tutoring interaction study that we conducted. Baseline help-seeking behavior for both groups was assessed during Session 1, in which neither group received intervention behaviors from the robot. The intervention group received intervention strategies from the robot during Sessions 2 through 4. Children in both groups completed the four tutoring sessions over the course of two weeks.

a between-subjects experiment in which participating students were randomized into one of two conditions: *control* (participants use on-demand help with the robot) and *intervention* (participants received intervention strategies from the robot). We chose to use a repeated measures design to examine behavior change and learning gains of the participants over time. We established each child's baseline behavior during the first session without introducing the confound of the robot's shaping strategies. Figure 3 illustrates the study design.

One established way of providing help in a learning environment is to allow the student to use on-demand help. On-demand help refers to help provided by the learning environment that must be actively solicited by the learner [3]. Participants in the control condition relied on the buttons on the tablet interface (pictured in Figure 1) to make up to three help requests per question to the robot whenever they wanted to during each session, thereby using on-demand help features of the application.

Participants in the intervention condition followed this same method of requesting help; however, the robot also employed two strategies informed by the literature aimed at countering suboptimal help-seeking behavior: *help aversion* and *help overuse* [1, 10]. Below, we describe the implementation of these two strategies for our robot tutoring system.

- S_1: If the participant makes two consecutive incorrect attempts on a problem without asking for any hints, the robot will automatically provide the participant with the next hint that the participant has not yet requested.
- S_2: If the participant makes three consecutive hint requests on a problem without an attempt in between, the robot will deny the participant the third hint, and request that the participant attempt the problem before asking for more help.

While triggering S_1 may not indicate that the participant was completely help averse, the behavior involved was still considered suboptimal in our tutoring context. This trigger indicated that the participant was not using the help features of the tutoring system in the most productive way. S_2 is a strategy used to counter the suboptimal behavior of making successive hint requests to receive the most information before attempting the problem. This trigger indicated that the participant was not trying to use the information presented in previous hints. Asking the student to make an attempt before requesting more help can encourage the participant to make a bigger effort to understand and use the presented hints. These are not the only suboptimal help-seeking behaviors that can be defined in this tutoring context and the absence of these two suboptimal behaviors does not represent optimal help-seeking behavior across all students. Other suboptimal help-seeking behaviors can be measured by how help requests are distributed through time within and across sessions. However, we did not attempt to quantify these or design intervention strategies to counter these types of behaviors in this study. We chose two suboptimal help-seeking behaviors that were based on frequencies of behaviors representing help aversion and help overuse within our tutoring context that could be easily detected during each question. This led to the design of the simple shaping strategies that can directly counter these behaviors to shape more productive help use over time.

Because these two suboptimal behaviors specifically related to help use during a tutoring session have been shown to impact learning, we hypothesized that building simple robot-shaping strategies to directly counter each of these behaviors would lead to improved behavior and learning regardless of "why" students engaged in these behaviors. These particular strategies were derived from countering using "too much" or "too little" help in the context of the robot tutoring interaction that we designed. Though we planned to analyze whether suboptimal help-seeking behaviors relate to a child's self-reported academic motivation, we did not measure this relationship before conducting the multi-session user study. Rather than trying to classify a child's motivational state based on the child's use of suboptimal help-seeking behavior, we investigate whether a robot can effectively shape suboptimal use of the help features in our tutoring system. We then further attempt to elucidate whether a child's motivation in learning is related to the child's use of suboptimal help-seeking behaviors in order to provide a broader understanding for why shaping these behaviors may lead to more effective learning within tutoring.

3.4 Measures

In this section, we describe the SRQ-A and the measures that we used to assess the relationship between students' motivation in learning and their suboptimal help-seeking behavior observed in Session 1 (baseline). We then report the metrics that we use to evaluate the effectiveness of the robot employing help-seeking intervention strategies from Session 2 to Session 4.

3.4.1 The Academic Self-Regulation Questionnaire (SRQ-A). The SRQ-A explores why children complete their schoolwork and is designed for children in late elementary school and middle school [61]. The questionnaire asks children to explicitly consider reasons for completing academic tasks that they are familiar with, such as homework. There are four main questions on the SRQ-A, each requiring responses to eight individual items, making 32 items in total. Each response required is on a four-point scale: "very true" is scored as four, "sort of true" is scored as three, "not very true" is scored as two, and "not true at all" is scored as one. Each of the 32 responses is associated with one of four subscales, which are categories of reasons for academic achievement: external, introjected, identified, and intrinsic. Ryan and Connell define the four categories of reasons as follows: "*External* reasons were those where behavior is explained by reference to external authority, fear of punishment, or rule compliance. *Introjected* reasons were framed in terms of internal, esteem-based pressures to act, such as avoidance of guilt and shame or concerns about self- and other-approval. *Identifications* were captured by reasons involving acting from one's own values or goals, and typically took the form of 'I want.' Finally, and where applicable, we included *intrinsic* reasons for action where the behavior is done simply for its inherent enjoyment or for fun." [61]. An example of one of the four main questions on the questionnaire is: Why do I do my homework? This question is followed by eight reasons, each for which the student would circle one of the four options on the scale indicating the extent of their agreement with that particular reason. The external category is associated with reasons like "because I'll get in trouble if I don't." A reason representing the introjected category is "because I want the teacher to think I'm a good student." An example of a reason that fits identification is "because it's important to me to do my homework." Last, "because it's fun" is an example that corresponds with intrinsic reasons for action. The four subscales are organized in order from external to internal orientation toward why they complete school-related activities. Of these four subscales, the *external* category measures externally oriented motivation, while the other three categories—*introjection*, *identification*, and *intrinsic* motivation—can be grouped under the broader category of internal motivation. The ordering of the subscales indicates that identification is more internal than introjection, and the intrinsic category is considered to be the most internal category of motivation on this scale. The *introjected* category of motivation includes reasons for behavior involving the approval of others, which appears to be an example of external motivation. However, it also includes reasons such as "because I will feel really proud of myself if I do well" and focuses on behaving according to internal pressures, allowing the introjected category to be considered somewhat internally oriented. *Identified* reasons for academic behavior reflect a strong internal value on learning, which is considered to be a positive source of motivation and also allows for identification to be considered internal. Scores for each subscale are calculated by averaging the values of the responses that correspond with each category of reasons.

While there are multiple ways of using the SRQ-A (e.g., combining the four subscales into one score), we chose to treat the subscales separately, because the subscales are not mutually exclusive. For example, a student can score highly on both external and intrinsic subscales. The high internal consistency for the external (9 items, Cronbach's α = .813), introjected (9 items, Cronbach's α = .885), identified (7 items, Cronbach's α = .792), intrinsic (7 items, Cronbach's α = .874) subscales confirmed their reliability. The average scores across all participants for each of the four subscales are as follows: mean (M) = 3.07, standard deviation (SD) = .57 (external); M = 3.08, SD = .59 (introjected); M = 3.51, SD = .45 (identified); M = 2.68, SD = .62 (intrinsic).

3.4.2.2 Help-Seeking Behavior Change. To evaluate the effectiveness of the robot intervention of suboptimal help-seeking behavior, we counted the number of times strategies $\mathbf{S_1}$ and $\mathbf{S_2}$ would have been triggered for each participant i for a given session S:

$$num\_triggers(i, S) = num\_auto\_hints(i, S) + num\_denied\_hints(i, S).$$

This count involves num_auto_hints, representing the number of times a hint would be automatically given, and num_denied_hints, representing the number of hints that would be denied. We calculate $num\_triggers(i, S)$ for each participant. where $S = 1$ and $S = 4$. This metric represents the number of suboptimal help-seeking behaviors observed for a participant in Session 1 (baseline help-seeking behavior) and in Session 4. We further define $\Delta_{triggers}$ as a metric that captures the difference between number of triggers from Session 1 to Session 4 for participant i:

$$\Delta_{triggers}(i) = num\_triggers(i, 4) - num\_triggers(i, 1).$$

We employ this metric to understand the change in a participant's use of help-seeking behavior.

3.4.3 Learning Gains. In this user study, we investigated a specific learning outcome, learning gains, to measure each student's improvement in knowledge from before to after the tutoring interactions. To assess how the robot intervention may influence a student's learning outcomes, particularly learning gains, we asked participants to complete a pretest before Session 1 and a posttest after Session 4. Participants completed the pretest on the same day as Session 1 and the posttest on the same day as Session 4. Both the pretest and posttest consisted of eight questions, containing the same types of problems that were presented during the four tutoring sessions. The questions on both pretest and posttest were almost identical, with differing numbers within each

Table 1. Participant Demographic Information for the Two Experimental Conditions in Our User Study

Condition	N	Age	Gender	Pretest accuracy	Ethnicity				
					Asian	Caucasian	Hispanic	More than one	Did not report
Control	15	M=10.9 SD=.80	8 males 7 females	M=.51 SD=.27	13.3%	60.0%	13.3%	6.7%	6.7%
Intervention	14	M=10.68 SD=.54	8 males 6 females	M=.31 SD=.29	7.1%	78.6%	0.0%	7.1%	7.1%

problem. We scored both tests by awarding one point for each correct answer and dividing the number of correct answers by the total number of questions, resulting in scores ranging from zero to one (accuracy). The difference in test scores between the pretest and the posttest, Δ_{score}, is a within-subjects measure of learning gains over the course of the entire experiment. All measures of learning gains are calculated relative to each participant's individual pretest score. In our assessment, we employ *normalized learning gain* as defined below for each participant i, allowing us to control for individuals starting at different levels of expertise:

$$\Delta_{score}(i) = \frac{score_{\text{post}}(i) - score_{\text{pre}}(i)}{1 - score_{\text{pre}}(i)}.$$

3.5 Procedure

Both parent and child consent forms were obtained for each student prior to the student's participation in this study. Participants were escorted from their classrooms by the experimenter. Before Session 1, each child was introduced to the robot. The robot greeted the participant, saying "Hello! My name is Nao, your personal robot tutor. I'm really excited to meet you and work on some problems together." After this introduction, the participants completed a pretest and the SRQ-A addressing why they do certain school-related activities. They then completed four distinct tutoring sessions with the robot, spaced over approximately two weeks. Upon finishing Session 4 of the tutoring interaction, participants completed a posttest. Each child was then given stickers and pencils for participating in the study and was escorted back to the classroom by the experimenter.

3.6 Participants

The participants in this study were fifth and sixth grade students from local public schools in Connecticut. A total of 33 students were recruited; however, four participants were excluded from our data analysis (three for not completing the study owing to school absences, and one for non-compliance). Table 1 displays the demographic information for both the control and intervention groups. Of the 29 students included in our data analysis, 15 and 14 participants were in the control and intervention conditions, respectively. In the control group, there were eight males and seven females with a mean age of 10.9 years (SD = .80). The ethnicity of each participant was reported by parents: 13.3% Asian, 60.0% Caucasian, 13.3% Hispanic, 6.7% reported more than one ethnicity, and 6.7% did not report. The average pretest score of the control group was .51 (SD = .27). In the intervention group, there were eight males and six females with a mean age of 10.68 years (SD = .54). The ethnicities of the participants as reported by parents were as follows: 7.1% Asian, 78.6% Caucasian, 7.1% reported more than one ethnicity, and 7.1% did not report. The average pretest score of the intervention group was .31 (SD = .29). As we did not screen children based on their pretest scores, the randomized distribution of students in each group resulted in a lower average pretest score for the intervention group as compared to the control group. Pretest scores were not statistically significantly different between the two experimental groups: $t(27) = 1.904$, $p = .068$. Furthermore, all measures of learning improvement for participants were calculated relative to each individual participant's pretest score, accounting for differing levels of incoming knowledge. Our two experimental groups were gender balanced and there were no significant gender differences in pretest score between males ($M = .39$, $SD = .32$) and females ($M = .44$, $SD = .25$); $t(27) = .471$, $p = .641$.

Almost all students used the system's help features to some extent in the initial tutoring session. Of the participants, 93.1% requested at least one hint during the first session. Only two students (6.9%) did not request help during this session. During Session 1, students requested 9.07 hints ($SD = 6.82$), on average. The average number of hints requested during Session 1 did not significantly differ between the control ($M = 7.20$, $SD = 5.51$) and intervention ($M = 11.07$, $SD = 7.70$) groups; $t(27) = -1.566$, $p = .129$. In addition, we did not find any significant gender differences in the average number of hints requested during Session 1 between males ($M = 10.44$, $SD = 6.99$) and females ($M = 7.38$, $SD = 6.49$); $t(27) = 1.208$, $p = .238$.

4 RESULTS

In this section, we provide empirical findings that demonstrate how motivation closely relates to behavior during learning, which, in turn, can be effectively shaped by a robot tutoring system to improve learning outcomes. We first present results on the relationship between a student's motivation in learning as assessed by the SRQ-A and the occurrence of the student's suboptimal help-seeking behaviors during the tutoring interaction with the robot. We then show how the robot's intervention strategies impact changes in the number of suboptimal help-seeking behaviors ($\Delta_{triggers}$) and test scores (Δ_{score}) for each participant. Additionally, we identified another suboptimal behavior that relates to a student's motivation in learning, further demonstrating the opportunity for robot tutoring systems to leverage observable behavior to drive tutoring intervention behavior. All statistical tests reported below employed an α of .05 for significance. We used non-parametric statistical tests when needed based on the distribution of our empirical data.

4.1 Relationship Between Motivation and Suboptimal Help-Seeking Behaviors

In exploring the relationship between students' motivation in learning and their suboptimal help-seeking behavior, we focused on correlations between the SRQ-A subscales and the baseline help-seeking behavior assessed in Session 1 of the user study (see Table 2). The baseline help-seeking behavior was represented as *num_triggers* defined in Section 3.4.2. We consider these two suboptimal behaviors together as they both relate specifically to suboptimal use of the hints in the tutoring system and were the two behaviors that we designed strategies for our robot tutor to counter in our user study prior to this analysis. The average number of triggers of suboptimal help-seeking behavior (*num_triggers*) for students in Session 1 was 2.90 ($SD = 2.51$). There were no significant gender differences in the average number of triggers between males ($M = 3.44$, $SD = 2.45$) and females ($M = 2.23$, $SD = 2.52$) in Session 1; $t(27) = 1.302$, $p = .204$. There were also no significant differences in the baseline number of triggers from Session 1 between the control ($M = 2.07$, $SD = 2.66$) and intervention ($M = 3.78$, $SD = 2.08$) groups; $t(27) = -1.929$, $p = .064$.

Across 29 participants, we found that the frequency of suboptimal help-seeking behavior in Session 1 is negatively correlated with the *identification* subscale of the SRQ-A,; $r(27) = -.37$, $p = .048$ (Pearson correlation). This negative correlation indicates that higher scores on the identification subscale of the SRQ-A correlate with a lower number of suboptimal help-seeking behaviors. The identification subscale captures motivation driven by one's own values and goals and corresponds to reasons on the questionnaire such as "because it's important to me" and "because I want to understand the subject." This finding suggests that children who feel motivated to complete work due to internally viewing it as important are less likely to engage in suboptimal help-seeking

Table 2. Results of Correlation Analyses Linking Suboptimal Behaviors Occurring in Session 1 to Subscales of the SRQ-A

	Motivation Subscales	Cronbach's α	Average Scores	Pearson Correlation to Suboptimal Help-Seeking Behavior (*num_triggers*) (S1)	Pearson Correlation to Number of fast attempts (S1)
External Orientation	**External** *e.g., I'll get in trouble if I don't do my homework.*	.813	M=3.07 SD=.57	r(27) = -.090 *p = .643*	r(27) = -.064 *p = .740*
Internal Orientation	**Introjected** *e.g., I want the teacher to think I'm a good student.*	.885	M=3.08 SD=.59	r(27) = -.321 *p = .090*	r(27) = -.162 *p = .402*
	Identified *e.g., It's important to me to do my homework.*	.792	M=3.51 SD=.45	**r(27) = -.370** ***p = .048****	r(27) = -.326 *p = .084*
	Intrinsic *e.g., I do my homework because it's fun.*	.874	M=2.68 SD=.62	r(27) = -.24 *p = .211*	**r(27) = -.376** ***p = .045****

Cronbach's α values for each subscale indicate the reliability of the questionnaire subscales. Participants' number of triggers of suboptimal help-seeking behavior from Session 1 had a negative relationship with the identied category of motivation (Section 4.1). Additional results showed that number of fast attempts from Session 1 was negatively correlated to the intrinsic category of motivation (Section 4.4).

behaviors, such as asking for all available hints consecutively or not requesting available help. One potential explanation for this finding is that those who score highly on the identification subscale of the SRQ-A are motivated to expend effort throughout the learning interaction and tend not to avoid help or rely too heavily on help features in the tutoring environment.

This finding shows how motivational factors can be linked to observable behaviors in a learning environment. It provides evidence showing that internal values regarding a learning task can substantially influence a person's suboptimal behaviors during a learning interaction. Next, we demonstrate that creating shaping strategies that directly counter these suboptimal behaviors over multiple tutoring interactions led to fewer occurrences of these undesirable behaviors and greater learning gains.

4.2 Help-Seeking Behavior Change

Figure 4 summarizes the results of behavior change from Session 1 to Session 4 for the participants in the two experimental conditions. A Wilcoxon signed-rank test showed that participants in the intervention group exhibited significantly fewer suboptimal behaviors in Session 4 (median [Mdn] = 2.0, interquartile range [IQR] = 1) than they did in Session 1 (Mdn = 4.0, IQR = 2); $Z = -2.605$, $p = .009$. Conversely, we did not see such improvement in the control group; there was no significant difference in the number of suboptimal behaviors that participants in the control group (using on-demand help) exhibited in Session 1 (Mdn = 1.0, IQR = 4) and Session 4 (Mdn = 1.0, IQR = 3); $Z = -.213$, $p = .832$ (Wilcoxon signed-rank test). Moreover, the decrease in number of triggers, $\Delta_{triggers}$, was significantly greater for the intervention group (Mdn = −1.5, IQR = 2.0) than for the control group (Mdn = 0.0, IQR = 0), indicated by a Mann-Whitney test; $U = 45.000$, $p = .008$.

Together, these results demonstrate that the robot's intervention strategies aimed at shaping productive help-seeking behavior were successful in mitigating the occurrences of suboptimal help-seeking behaviors over time. Participants in the intervention condition significantly decreased their number of suboptimal help-seeking behaviors over time while participants in the control condition did not. Though the behavior change that we observed may be a short-term effect, we did observe more productive help-seeking behavior over the course of the two weeks, indicating the effectiveness of the intervention strategies in the duration of time that students interacted with the robot.

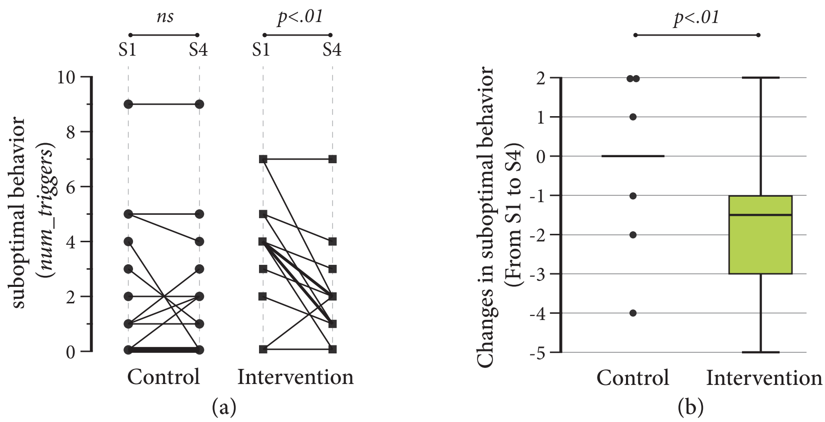

Fig. 4. Results for behavior change indicate that participants in the intervention group significantly decreased their suboptimal help-seeking behaviors from Session 1 (S1) to Session 4 (S4) and decreased these behaviors significantly more than the control group. (a) Number of suboptimal behaviors, as captured by *num_triggers*, from Session 1 and Session 4 for each participant in the two experimental conditions. Thicker lines represent multiple participants with the same values. (b) Change in *num_triggers* for participants in each group. In the boxplot, the darker line inside the box represents the median, and the extents of the box represent the first and third quartiles. Because of the large number of participants with no difference in suboptimal behavior from S1 to S4 in the control group, the first and third quartiles and the median for the control group are all zero, resulting in the flat box. Remaining data points are shown as points (outliers) owing to control-group data having IQR = 0.

4.3 Learning Gains

Figure 5 summarizes the results for learning gains for the participants in the two experimental conditions. For participants who received the help-seeking intervention strategies, posttest scores (Mdn = .62, IQR = .63) were significantly higher than pretest scores (Mdn = .25, IQR = .63); $Z = 3.089$, $p = .002$ (Wilcoxon signed-rank test). On the other hand, for participants in the control group that did not receive intervention strategies, there was no significant change in score from pretest (Mdn = .50, IQR = .38) to posttest (Mdn = .75, IQR = .38); $Z = 1.615$, $p = .106$ (Wilcoxon signed-rank test). We further sought to understand differences in normalized learning gains between groups. An independent sample t-test revealed that participants in the intervention group ($M = .45$, $SD = .34$) improved their score from pretest to posttest significantly more than those in the control group ($M = .06$, $SD = .59$); $t(27) = -2.169$, $p = .039$.

These results show that participants receiving intervention strategies from the robot were able to improve their test scores effectively, while the group that relied on using on-demand help were not. These results, together with the results of behavior change from the previous section, indicate that the shaping strategies employed by the robot improved help-seeking behaviors, which thereby positively impacted learning outcomes for participants.

4.4 Additional Findings: Relationship Between Motivation and Fast Attempts

Thus far, we have provided empirical findings illustrating the linkage of motivation, behaviors, and outcomes in a learning scenario in which children practiced math problems with a tutoring robot. In particular, we have shown that internal motivations, especially *identification*, were negatively correlated to suboptimal help-seeking behaviors during learning and that the robot tutoring system effectively intervened in response to the undesirable help-seeking behaviors to improve behavioral and learning outcomes. Our findings of the linkage of motivation, behaviors,

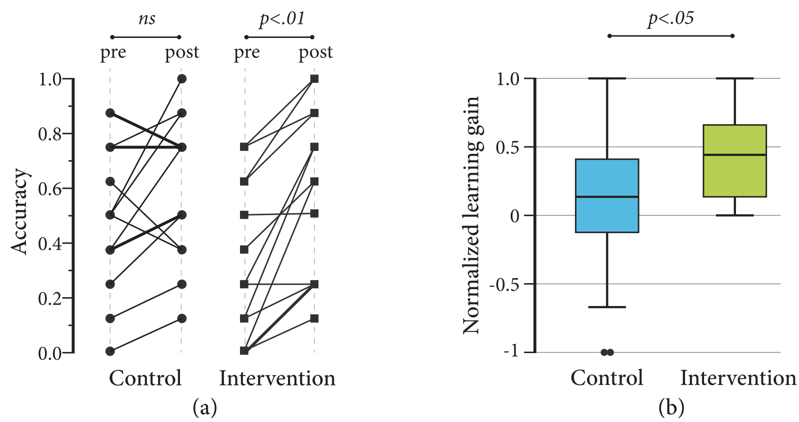

Fig. 5. Results for learning gains demonstrate that participants in the intervention group significantly improved their scores from pretest to posttest and improved their score significantly more than the control group, as measured by normalized learning gain. (a) Pretest and posttest scores for each participant in the two experimental conditions. Thicker lines represent multiple participants with the same scores. (b) Normalized learning gain for participants in each group. In the boxplot, the darker line inside the box represents the median, and the extents of the box represent the first and third quartiles.

and outcomes in learning encouraged us to explore different dimensions of suboptimal learning behaviors that robot tutoring systems hold potential to address through effective intervention strategies. Below, we report on a further analysis of our present data revealing the relationship between another internal factor of motivation, *instrinsic* motivation, and the suboptimal behavior of *fast attempts* (consecutive incorrect attempts within a small time frame). These results are also displayed in Table 2.

Another undesirable behavior that we observed in the tutoring interactions was children making successive attempts in a small window of time as if they did not make enough effort to work out the problems presented, indicating the likelihood of guessing. We considered such behavior as fast attempts and defined $num\_fast\_attempts$ as the number of attempts made within a small threshold of seconds (empirically determined to be 20 seconds) from the previous attempt. This behavior is considered to be suboptimal in this learning interaction, as the questions given to the students were on fractions and contained multiple steps to complete, meaning that a successive attempt within a few seconds of an incorrect attempt does not give the student adequate time to rework the problem. The average number of fast attempts ($num\_fast\_attempts$) for the 29 students in Session 1 was 2.79 (SD = 4.77). We did not consider fast attempts in advance when designing our robot shaping strategies for the user study that we conducted, in which we designed shaping strategies for suboptimal behavior specifically relating to hint use in the system. Because the robot did not actively track and counter this behavior, we consider it separately from our measure of suboptimal help-seeking behavior.

In exploring the relationship between motivation and the behavior of fast attempts, we found that $num\_fast\_attempts$ is negatively correlated with the *intrinsic* motivation subscale on the SRQ-A; $r(27) = -.38$, $p = .045$ (Pearson correlation). The intrinsic motivation subscale on the SRQ-A corresponds to reasons for doing various types of schoolwork for implicit enjoyment in learning and completing work. Scoring highly on this subscale may indicate a tendency to put effort into each attempt, even after an incorrect attempt, thereby making less "guessing" attempts. We further found that $num\_fast\_attempts$ is negatively correlated with posttest score; $r(27) = -.47$, $p = .011$ (Pearson correlation), although we did not include shaping behaviors aimed at countering fast attempts in our user study. This negative correlation dictates that higher posttest scores correspond with a lower incidence of fast attempts in Session 1 (baseline). This finding suggests an opportunity for designing robot intervention aiming to shape the suboptimal behavior of fast attempts to improve learning outcomes.

4.5 Discussion

The results of our user study demonstrate the linkage of motivation, behaviors, and learning outcomes that underlies learning processes (Figure 2). We employed the SRQ-A, involving four motivational factors, including *external*, *introjected*, *identified*, and *intrinsic*, to assess children's motivation in engaging in learning tasks. We found a negative correlation between the *identified* factor of motivation and the suboptimal behavior of help-seeking (e.g., help-averse and help-overuse behaviors) as well as a negative correlation between the *intrinsic* factor of motivation and the suboptimal behavior of fast attempts. Moreover, we demonstrated that a robot tutoring system can employ help-seeking intervention strategies effectively to reduce the suboptimal behavior of help-seeking and to improve learning gains. We also note that the correlations between the *introjected* factor of motivation and suboptimal help-seeking behaviors were marginally significant (Table 2). The *identified* category of motivation showed a marginally significant correlation to the suboptimal behavior of fast attempts as well. This aligns with our results, as each of these three categories of motivation are grouped together as internally oriented, indicating that these more positive sources of motivation all correspond to a lower incidence of suboptimal behaviors. However, not all categories of motivation with an internal orientation were correlated to the behaviors measured during the tutoring sessions. For example, we did not see a correlation between intrinsic motivation and the suboptimal help-seeking behaviors despite intrinsic motivation being the most internally oriented subscale of motivation. While we cannot be sure about why this is, this may be a limitation of our smaller sample size. Alternatively, it may be the case that certain behaviors, such as help-seeking behaviors, correlate more strongly with specific measures of internal motivation. Additional research is required to understand which subcategories of internal motivation may prove most informative in the context of our motivation-behavior-learning outcomes pipeline. Similarly, the external category of motivation, which is considered to be a less positive source of motivation in learning, did not significantly correlate to our measures of suboptimal help-seeking behaviors or fast attempts. Though a student's external motivation may still be useful to understand within a learning context, our empirical results indicate the value of understanding suboptimal behavior in relation to categories of internally oriented motivation.

Participants who reported a lower level of identified regulation were the ones who had higher levels of baseline suboptimal help usage. As higher identified regulation indicates agreement with reasons involving an internal sense of importance for completing academic work, this suggests that motivation in learning can influence how a student uses resources in a learning environment, such as the help features of the system. Our results on behavior change and learning gains lead us to believe that shaping these unproductive help-seeking behaviors during a learning interaction can cause students to engage more effectively with a robot tutoring system, thereby impacting their learning outcomes over time. Together, these findings add a broader lens on why a robot shaping suboptimal help-seeking behavior was effective in promoting learning.

Through this process, we introduced a basic approach for designing robot intervention, which involves finding observable behaviors that correspond to motivational factors and employing a robot to respond to these behaviors to create more effective tutoring interactions. In further exploring this approach, we identified an additional observable behavior that had a relationship to a measure of intrinsic motivation. Similarly to suboptimal help-seeking behavior, students who had lower intrinsic motivation as measured by the SRQ-A made more fast attempts, indicating that lower intrinsic motivation manifests itself during learning through suboptimal behaviors that can be monitored by a robot tutoring system. Because students' use of these fast attempts correlates negatively with posttest score, this suggests that a robot tutoring system specifically designed to intervene based on this behavior that signals suboptimal use of the system might lead to improved learning.

Prior work in tutoring systems has attempted to take into account a student's motivation in several different ways. Some systems automatically provide exercises or lessons that fit a student's ability, assuming that a student will stay motivated if experiencing the right level of challenge [16, 43]. Other work has sought to increase motivation in learning through game-based environments or has focused on building tutoring environments that enhance student motivation through various strategies, such as providing choices or using social support [19, 34, 64]. Our results highlight a simple way to incorporate types of motivation into a learning environment by demonstrating the usefulness of linking measures of motivation with observable behavior within a learning environment, which further lends itself to designing robot intervention behavior that has the potential to foster learning gains. We evaluated our approach in the context of an autonomous robot tutoring system for children; however, we believe this process can also be useful in designing behaviors for a variety of agent-based tutoring systems as well.

The influence of motivation on behavior is certainly not limited to learning outcomes within a tutoring setting. Extensive research has demonstrated how motivation impacts a variety of important health-related outcomes, such as physical fitness, smoking cessation, and weight loss [69]. Using an intelligent system to understand how motivation relates to observable behaviors can be leveraged to provide support that will be useful in promoting these types of impactful health outcomes. The design approach that we describe in this article is not limited to the domain of education but rather has potential to be applied to improving health-related outcomes as well. Providing empirical support for this design tool demonstrates the feasibility of leveraging the relationship between internal motivation and behaviors to inform the design of intervention behaviors for a supportive robot. Applying a similar process outside a tutoring setting in other application domains in which motivation plays a crucial role should be further explored.

5 IMPLICATIONS AND DIRECTIONS FOR FURTHER RESEARCH

5.1 The Promise and Challenge of Autonomous Robot–Child Tutoring

Robot tutoring systems have demonstrated great potential to provide effective, personalized learning for children, as they can leverage their embodied nature, social presence, and situated awareness to engage users and thereby enhance learning. These systems have been successfully used to promote learning in a variety of educational domains including math concepts, reading skills, and language learning tasks [28, 45, 56]. However, challenges still exist in building successful robot tutoring systems that can support the diverse needs of children. Though we explore these challenges in the context of robot tutoring systems, many of these challenges apply to a wider range of autonomous tutoring systems as well. One of the main challenges is understanding the internal state of the student during a learning interaction. Within a learning interaction, an effective tutoring agent should be able to provide support for the learner. Ideally, a robot tutor would be able to take into account a learner's motivation and tailor the interaction based on this user state. Though there are psychological theories that support taking into account student motivation in human tutoring, this is a much greater challenge for autonomous robot–child tutoring. The vast number of latent features and traits of a student during learning, including motivation, cannot be directly observed, making it computationally challenging for a robot to respond intelligently to these user states. Robot tutoring systems, therefore, must rely on observable approximations of behavior that convey information about a user's internal state. Research has shown the value of using sensor-based systems to better understand the internal state of a student involving gaze cues, valence, and attention level [41, 42, 70]. However, for high-level aspects of the user, such as motivation, it is still unclear what the relationship is between certain observable cues and high-level user states.

As we begin to build robot tutoring systems that respond to and support users based on their internal states, robots still face additional challenges owing to the particularly complex nature of these internal states. A child's motivation in learning is a manipulable construct that can be affected by a variety of external factors, such as a child's family, culture, and school. These dynamic influences make it challenging for a robot tutoring system to address a user's motivation in learning directly. Understanding how autonomous robot tutoring systems will fit into learning contexts in schools, afterschool care, and homes will help to inform what types of internal states and external influences these systems must account for when interacting with children.

Another challenge in autonomous robot–child tutoring deals with differences between children in a learning domain. Even if latent user internal states can be sensed through observable behavior, these behaviors can vary greatly between individual children. In our study, many behaviors within the tutoring application differed between children, including features such as average time to complete problems and number of hints requested. Even observable behavior within a robot tutoring system must be interpreted in relation to an individual user's baseline performance. The diverse learning abilities within children indicate that not all students express their internal states through the same behaviors, making it even more challenging for robot tutoring systems to autonomously intervene. Additionally, students may even have varied preferences in what types of interventions should be employed during learning interactions, including teaching styles and level of autonomy desired within the system. Children also have widely differing personality traits, such as extraversion level, which may influence their acceptance of or willingness to engage with a robot tutoring system over time. Robot tutoring systems have a rich space to explore in order to provide truly personalized tutoring interactions for children with a wide variety of behaviors and preferences.

5.2 Toward Effective Robot–Child Tutoring

Prior work in robot tutoring has often focused on providing content-based personalization for students interacting with the system by providing personalized behavior directly involving the learning task at hand [43, 65, 75]. The learning process is complex and also involves a variety of different dimensions that are complementary to the cognitive task at hand, such as attitudes toward learning, affective and emotional support, and help-seeking behavior. The user study that we conducted goes beyond providing content-based personalization to students and explores an aspect of learning that has not been widely studied within robot tutoring systems, namely, help-seeking behavior. We showed the effectiveness of a robot tutoring system that targeted suboptimal help-seeking behaviors that relate to self-regulated learning. Demonstrating that shaping these types of behaviors had an impact on the learning outcomes of the students receiving the robot intervention strategies shows the promise of robot tutoring systems to handle these crucial aspects of a user during learning. Given that our results showed that the intervention group improved their test scores significantly more than the control group, it is critical for these systems to monitor the help-seeking behaviors of children. Furthermore, our ability to demonstrate the relationship between motivation in learning and these suboptimal help-seeking behaviors helps to elucidate possible reasons for why these behaviors may occur. This new perspective may broaden the types of robot intervention that should be considered in the future, as strategies that seek to bolster the use of self-regulated learning may also be beneficial for children.

This work provides empirical evidence demonstrating how the link between motivation in learning and observable behavior can be leveraged to design robot intervention behavior to impact learning. Though it is not possible to directly understand a child's motivation in learning through computing technology, we have outlined a method to identify behaviors that occur within a given learning context that relate to a child's motivation in learning. We then demonstrated that a robot tutoring system that intelligently intervenes when these specific suboptimal behaviors occur can effectively reduce the occurrence of the undesired behavior over time. We outlined our method of identifying specific behaviors that relate to a child's self-reported measure of the child's motivation in learning and then employing a robot tutoring system to respond to these behaviors to enhance learning as a process that can be used to design effective robot intervention strategies. We also showed how the process of designing robot intervention strategies can be applied by identifying an additional suboptimal behavioral feature collected from our empirical data and demonstrating its relation to motivation in learning. This is one plausible method for designing robot intervention that can be used to explore the different dimensions of learning and their relationship to learning outcomes. Our results demonstrate the usefulness of this process and add to the existing work demonstrating the potential for robot tutoring systems to provide effective personalized instruction to children.

5.3 Limitations and Future Work

Though this work demonstrates the promise of linking motivation in learning, behavioral intervention, and learning outcomes to inform the design of effective robot tutoring systems for children, there are limitations that should be acknowledged and addressed by future work in robot–child tutoring. Our user study was conducted with a limited number of participants. Furthermore, the correlations that we observed between certain subscales of motivation and suboptimal behaviors during tutoring were not strong, which may have been a consequence of trying to relate measures capturing intricate constructs such as motivation types to more simple actions within a learning environment, such as requesting three hints in a row. We did not use these correlations to classify a child's internal motivational state but rather presented empirical evidence supporting the presence of a relationship between measures of complex categories of motivation to reliable, observable behaviors within an automated learning environment. Nonetheless, future work should investigate linking measures of motivation to observable behavior and designing robot behavior accordingly with a much larger population to more thoroughly evaluate the validity of this approach.

We focused on identifying observable behaviors that could be detected reliably during the learning interaction and relied on features that could be collected from the child's input using the tablet device. However, given other related work investigating the use of sensor-based approaches to understand information about the user (such as attention through EEG [70], or engagement through facial-feature detection [28, 41]), future work should leverage these channels of information to understand how these richer features relate to a student's motivation in learning as well. Our work also relied on self-reported questionnaire data (SRQ-A) to measure a child's motivation for completing academic work. Though this was a validated questionnaire designed to be used by children, not all children excel at accurately assessing themselves, especially at a meta-cognitive level. It would also be beneficial to consider other measures of a child's motivation in learning from adults who have interacted with them for a longer period of time, such as parents or teachers.

The user study that we conducted monitored observable suboptimal behavior within a learning environment and then leveraged a robot tutor to intervene specifically when these behaviors occurred to shape more productive help-seeking behavior. While designing robot intervention to counter these behaviors directly was effective in this context, robot tutoring intervention could be expanded to more broadly address the underlying ties to internal motivation. Robot intervention strategies designed based on an elaborated user model of a student's motivation in learning should be explored. For example, a robot tutor might employ different teaching tactics depending on whether a student's subscale scores on the SRQ-A demonstrate that they are more motivated by external factors than internal factors. Furthermore, robot intervention strategies that specifically try to foster an increase in internal motivation itself should also be explored.

Additional work must also explore longer-term studies in the domain of robot–child tutoring. Robot tutoring systems must be built to interact with children in real-world settings, such as classrooms and homes, for long periods of time and should be designed robustly to successfully engage their users during these extended deployments. Furthermore, it is of particular importance in the application of tutoring to understand whether learning gains and behavior change as a result of robot intervention during tutoring lasts over time and yields robust learning. From the perspectives of both large-scale system deployments and student learning outcomes, it is crucial that we understand whether robot tutoring systems can be designed to handle long-term interactions that can have lasting impact.

6 CONCLUSION

In this article, we presented empirical findings to illustrate the linkage between motivation in learning, behavioral intervention, and learning outcomes within the domain of robot–child tutoring. We conducted a user study involving 29 fifth and sixth grade children doing math problems with a robot over four tutoring sessions. Our empirical analysis revealed the relationship between self-reported measures of children's motivation in learning and their use of certain suboptimal help-seeking behaviors. We also demonstrated the effectiveness of a robot tutor that employed intervention strategies to shape participants' suboptimal behaviors over a control group that did not receive the intervention strategies. We found that participants in the intervention group improved their suboptimal help-seeking behavior and learning significantly more than the control group over the four sessions, indicating that productive help-seeking behavior can be shaped by a robot tutor and does impact learning outcomes. Together, these results demonstrate a process for how to design robot intervention during robot–child tutoring that supports broader aspects of the learner, including the student's motivation in learning, and should be used in service of the goal of providing effective robot–child tutoring for all learners.

REFERENCES

[1] Vincent Aleven, Bruce McLaren, Ido Roll, and Kenneth Koedinger. 2004. Toward tutoring help seeking. In *Intelligent Tutoring Systems*. Springer, 227–239.

[2] Vincent Aleven, Ido Roll, Bruce M. McLaren, and Kenneth R. Koedinger. 2010. Automated, unobtrusive, action-by-action assessment of self-regulation during learning with an intelligent tutoring system. *Educational Psychologist* 45, 4 (2010), 224–233.

[3] Vincent Aleven, Elmar Stahl, Silke Schworm, Frank Fischer, and Raven Wallace. 2003. Help seeking and help design in interactive learning environments. *Review of Educational Research* 73, 3 (2003), 277–320.

[4] Ivon Arroyo, David G. Cooper, Winslow Burleson, Beverly Park Woolf, Kasia Muldner, and Robert Christopherson. 2009. Emotion sensors go to school. In *Artificial Intelligence in Education*, Vol. 200. IOS Press, 17–24.

[5] Roger Azevedo. 2007. Understanding the complex nature of self-regulatory processes in learning with computer-based learning environments: An introduction. *Metacognition and Learning* 2, 2–3 (2007), 57–65.

[6] Wilma A. Bainbridge, Justin Hart, Elizabeth S. Kim, and Brian Scassellati. 2008. The effect of presence on human-robot interaction. In *17th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN'08)*. IEEE, 701–706.

[7] Ryan S. J. Baker, Albert T. Corbett, and Vincent Aleven. 2008. More accurate student modeling through contextual estimation of slip and guess probabilities in Bayesian knowledge tracing. In *International Conference on Intelligent Tutoring Systems*. Springer, 406–415.

[8] Ryan S. J. Baker, Albert T. Corbett, Kenneth R. Koedinger, Shelley Evenson, Ido Roll, Angela Z. Wagner, Meghan Naim, Jay Raspat, Daniel J. Baker, and Joseph E. Beck. 2006. Adapting to when students game an intelligent tutoring system. In *International Conference on Intelligent Tutoring Systems*. Springer, 392–401.

[9] Ryan Shaun Baker, Albert T. Corbett, Kenneth R. Koedinger, and Angela Z. Wagner. 2004. Off-task behavior in the cognitive tutor classroom: When students game the system. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 383–390.

[10] Ryan S. J. Baker, Albert T. Corbett, Ido Roll, and Kenneth R. Koedinger. 2008. Developing a generalizable detector of when students game the system. *User Modeling and User-Adapted Interaction* 18, 3 (2008), 287–314.

[11] Carole R. Beal, Lei Qu, and Hyokyeong Lee. 2006. Classifying learner engagement through integration of multiple data sources. In *Proceedings of the National Conference on Artificial Intelligence*, Vol. 21. Menlo Park, CA; Cambridge, MA, MIT Press; London, AAAI Press; 1999, 151.

[12] Joseph E. Beck. 2005. Engagement tracing: Using response times to model student disengagement. In *Proceedings of the 2005 Conference on Artificial Intelligence in Education: Supporting Learning Through Intelligent and Socially Informed Technology*. IOS Press, 88–95.

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

[14] Peter Brusilovsky and Christoph Peylo. 2003. Adaptive and intelligent web-based educational systems. *International Journal of Artificial Intelligence in Education* 13, 2–4 (2003), 159–172.

[15] Min Chi and Kurt VanLehn. 2010. Meta-cognitive strategy instruction in intelligent tutoring systems: How, when, and why. *Educational Technology & Society* 13, 1 (2010), 25–39.

[16] Benjamin Clement, Didier Roy, Pierre-Yves Oudeyer, and Manuel Lopes. 2015. Multi-armed bandits for intelligent tutoring systems. *Journal of Educational Data Mining* 7, 2 (2015), 20–48.

[17] Cristina Conati and Heather Maclaren. 2009. Empirically building and evaluating a probabilistic model of user affect. *User Modeling and User-Adapted Interaction* 19, 3 (2009), 267–303.

[18] Albert T. Corbett and John R. Anderson. 1994. Knowledge tracing: Modeling the acquisition of procedural knowledge. *User Modeling and User-adapted Interaction* 4, 4 (1994), 253–278.

[19] 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.

[20] Angel De Vicente and Helen Pain. 2002. Informing the detection of the students' motivational state: An empirical study. In *International Conference on Intelligent Tutoring Systems*. Springer, 933–943.

[21] E. E. Deci and R. M. Ryan. 1985. *Intrinsic motivation and self-determination in human behavior*. Plenum Press (1985).

[22] Edward L. Deci, Robert J. Vallerand, Luc G. Pelletier, and Richard M. Ryan. 1991. Motivation and education: The self-determination perspective. *Educational Psychologist* 26, 3–4 (1991), 325–346.

[23] Michel C. Desmarais and Ryan S. Baker. 2012. A review of recent advances in learner and skill modeling in intelligent learning environments. *User Modeling and User-Adapted Interaction* 22, 1–2 (2012), 9–38.

[24] Sidney K. D'Mello, Scotty D. Craig, Amy Witherspoon, Bethany Mcdaniel, and Arthur Graesser. 2008. Automatic detection of learner's affect from conversational cues. *User Modeling and User-Adapted Interaction* 18, 1–2 (2008), 45–80.

[25] Carol S. Dweck. 1986. Motivational processes affecting learning. *American Psychologist* 41, 10 (1986), 1040.

[26] Elaine S. Elliott and Carol S. Dweck. 1988. Goals: An approach to motivation and achievement. *Journal of Personality and Social Psychology* 54, 1 (1988), 5.

[27] Sharon Nelson-Le Gall. 1985. Help-seeking behavior in learning. *Review of Research in Education* 12, 1 (1985), 55–90.

[28] Goren Gordon, Samuel Spaulding, Jacqueline Kory Westlund, Jin Joo Lee, Luke Plummer, Marayna Martinez, Madhurima Das, and Cynthia Breazeal. 2016. Affective personalization of a social robot tutor for children's second language skills. In *Proceedings of the 30th AAAI Conference on Artificial Intelligence*. AAAI Press, 3951–3957.

[29] Adele Eskeles Gottfried. 1985. Academic intrinsic motivation in elementary and junior high school students. *Journal of Educational Psychology* 77, 6 (1985), 631–45.

[30] Jeffrey Alan Greene and Roger Azevedo. 2007. Adolescents' use of self-regulatory processes and their relation to qualitative mental model shifts while using hypermedia. *Journal of Educational Computing Research* 36, 2 (2007), 125–148.

[31] Allyson F. Hadwin, John C. Nesbit, Dianne Jamieson-Noel, Jillianne Code, and Philip H. Winne. 2007. Examining trace data to explore self-regulated learning. *Metacognition and Learning* 2, 2–3 (2007), 107–124.

[32] Susan Harter. 1981. A new self-report scale of intrinsic versus extrinsic orientation in the classroom: Motivational and informational components. *Developmental Psychology* 17, 3 (1981), 300.

[33] Iris Howley, Takayuki Kanda, Kotaro Hayashi, and Carolyn Rosé. 2014. Effects of social presence and social role on help-seeking and learning. In *Proceedings of the 2014 ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 415–422.

[34] G. Tanner Jackson and Danielle S. McNamara. 2013. Motivation and performance in a game-based intelligent tutoring system. *Journal of Educational Psychology* 105, 4 (2013), 1036.

[35] Aidan Jones, Susan Bull, and Ginevra Castellano. 2018. "I know that now, I'm going to learn this next" promoting self-regulated learning with a robotic tutor. *International Journal of Social Robotics* 10, 4 (2018), 439–454.

[36] Aidan Jones and Ginevra Castellano. 2018. Adaptive robotic tutors that support self-regulated learning: A longer-term investigation with primary school children. *International Journal of Social Robotics* 10, 3 (2018), 357–370.
[37] Takayuki Kanda, Takayuki Hirano, Daniel Eaton, and Hiroshi Ishiguro. 2004. Interactive robots as social partners and peer tutors for children: A field trial. *Human-computer Interaction* 19, 1 (2004), 61–84.
[38] Takayuki Kanda, Michihiro Shimada, and Satoshi Koizumi. 2012. Children learning with a social robot. In *Proceedings of the 7th Annual ACM/IEEE International Conference on Human-Robot Interaction.* ACM, 351–358.
[39] James Kennedy, Paul Baxter, Emmanuel Senft, and Tony Belpaeme. 2015. Higher nonverbal immediacy leads to greater learning gains in child-robot tutoring interactions. In *International Conference on Social Robotics.* Springer, 327–336.
[40] Krittaya Leelawong and Gautam Biswas. 2008. Designing learning by teaching agents: The Betty's brain system. *International Journal of Artificial Intelligence in Education* 18, 3 (2008), 181–208.
[41] Iolanda Leite, Ginevra Castellano, André Pereira, Carlos Martinho, and Ana Paiva. 2012. Modelling empathic behaviour in a robotic game companion for children: An ethnographic study in real-world settings. In *Proceedings of the 7th Annual ACM/IEEE International Conference on Human-Robot Interaction.* ACM, 367–374.
[42] Séverin Lemaignan, Fernando Garcia, Alexis Jacq, and Pierre Dillenbourg. 2016. From real-time attention assessment to with-me-ness in human-robot interaction. In *11th ACM/IEEE International Conference on Human Robot Interaction.* IEEE Press, 157–164.
[43] Daniel Leyzberg, Samuel Spaulding, and Brian Scassellati. 2014. Personalizing robot tutors to individuals' learning differences. In *Proceedings of the 2014 ACM/IEEE International Conference on Human-Robot Interaction.* ACM, 423–430.
[44] Thomas W. Malone. 1981. Toward a theory of intrinsically motivating instruction. *Cognitive Science* 5, 4 (1981), 333–369.
[45] Joseph E. Michaelis and Bilge Mutlu. 2017. Someone to read with: Design of and experiences with an in-home learning companion robot for reading. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems.* ACM, 301–312.
[46] Anahita Mohseni-Kabir, Charles Rich, Sonia Chernova, Candace L. Sidner, and Daniel Miller. 2015. Interactive hierarchical task learning from a single demonstration. In *Proceedings of the 10th Annual ACM/IEEE International Conference on Human-Robot Interaction.* ACM, 205–212.
[47] Richard S. Newman. 1990. Children's help-seeking in the classroom: The role of motivational factors and attitudes. *Journal of Educational Psychology* 82, 1 (1990), 71.
[48] National Academy of Engineering. 2008. *Grand Challenges for Engineering*. Technical Report. National Academy of Engineering, Washington, DC. Retrieved March 28, 2017 from http://www.engineeringchallenges.org/File.aspx?id=11574&v=ba24e2ed.
[49] John F. Pane, Elizabeth D. Steiner, Matthew D. Baird, and Laura S. Hamilton. 2015. *Continued Progress: Promising Evidence on Personalized Learning*. Technical Report. RAND Corporation, Santa Monica, CA. Retrieved March 28, 2017 from http://www.rand.org/content/dam/rand/pubs/research_reports/RR1300/RR1365/RAND_RR1365.pdf.
[50] Zachary A. Pardos and Neil T. Heffernan. 2010. Modeling individualization in a Bayesian networks implementation of knowledge tracing. In *International Conference on User Modeling, Adaptation, and Personalization.* Springer, 255–266.
[51] André Pereira, Carlos Martinho, Iolanda Leite, and Ana Paiva. 2008. iCat, the chess player: The influence of embodiment in the enjoyment of a game. In *Proceedings of the 7th International Joint Conference on Autonomous Agents and Multiagent Systems-Volume 3.* International Foundation for Autonomous Agents and Multiagent Systems, 1253–1256.
[52] Paul R. Pintrich. 2000. The role of goal orientation in self-regulated learning. In *Handbook of Self-Regulation.* Academic Press, San Diego, 451–502.
[53] Paul R. Pintrich and Elisabeth V. De Groot. 1990. Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology* 82, 1 (1990), 33.
[54] Aaron Powers, Sara Kiesler, Susan Fussell, and Cristen Torrey. 2007. Comparing a computer agent with a humanoid robot. In *2nd ACM/IEEE International Conference on Human-Robot Interaction (HRI'07).* IEEE, 145–152.
[55] Aditi Ramachandran, Chien-Ming Huang, Edward Gartland, and Brian Scassellati. 2018. Thinking aloud with a tutoring robot to enhance learning. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction.* ACM, 59–68.
[56] Aditi Ramachandran, Chien-Ming Huang, and Brian Scassellati. 2017. Give me a break!: Personalized timing strategies to promote learning in robot-child tutoring. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction.* ACM, 146–155.
[57] Aditi Ramachandran, Alexandru Litoiu, and Brian Scassellati. 2016. Shaping productive help-seeking behavior during robot-child tutoring interactions. In *11th ACM/IEEE International Conference on Human-Robot Interaction.* IEEE Press, 247–254.
[58] Ido Roll, Vincent Aleven, Bruce M. McLaren, and Kenneth R. Koedinger. 2011. Improving students' help-seeking skills using metacognitive feedback in an intelligent tutoring system. *Learning and Instruction* 21, 2 (2011), 267–280.

[59] Allison M. Ryan and Paul R. Pintrich. 1997. "Should I ask for help?" The role of motivation and attitudes in adolescents' help seeking in math class. *Journal of Educational Psychology* 89, 2 (1997), 329.
[60] Allison M. Ryan and Huiyoung Shin. 2011. Help-seeking tendencies during early adolescence: An examination of motivational correlates and consequences for achievement. *Learning and Instruction* 21, 2 (2011), 247–256.
[61] Richard M. Ryan and James P. Connell. 1989. Perceived locus of causality and internalization: Examining reasons for acting in two domains. *Journal of Personality and Social Psychology* 57, 5 (1989), 749.
[62] Richard M. Ryan and Edward L. Deci. 2000. Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology* 25, 1 (2000), 54–67.
[63] Cyril J. Sadowski and Helen R. Woodward. 1981. Relationship between origin climate, perceived responsibility, and grades. *Perceptual and Motor Skills* 53, 1 (1981), 259–261.
[64] Martin Saerbeck, Tom Schut, Christoph Bartneck, and Maddy D. Janse. 2010. Expressive robots in education: Varying the degree of social supportive behavior of a robotic tutor. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 1613–1622.
[65] B. R. Schadenberg, M. A. Neerincx, F. Cnossen, and R. Looije. 2017. Personalising game difficulty to keep children motivated to play with a social robot: A Bayesian approach. *Cognitive Systems Research* 43 (2017), 222–231.
[66] Bernhard Schmitz and Bettina S. Wiese. 2006. New perspectives for the evaluation of training sessions in self-regulated learning: Time-series analyses of diary data. *Contemporary Educational Psychology* 31, 1 (2006), 64–96.
[67] Elaine Short, Katelyn Swift-Spong, Jillian Greczek, Aditi Ramachandran, Alexandru Litoiu, Elena Corina Grigore, David Feil-Seifer, Samuel Shuster, Jin Joo Lee, Shaobo Huang, et al. 2014. How to train your dragonbot: Socially assistive robots for teaching children about nutrition through play. In *23rd IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN'14)*. IEEE, 924–929.
[68] Daniel Solomon and Arthur J. Kendall. 1976. Individual characteristics and children's performance in open and traditional classroom settings. *Journal of Educational Psychology* 68, 5 (1976), 613.
[69] Victor J. Strecher, Brenda McEvoy DeVellis, Marshall H. Becker, and Irwin M. Rosenstock. 1986. The role of self-efficacy in achieving health behavior change. *Health Education Quarterly* 13, 1 (1986), 73–92.
[70] Daniel Szafir and Bilge Mutlu. 2012. Pay attention!: Designing adaptive agents that monitor and improve user engagement. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 11–20.
[71] Kurt VanLehn. 2006. The behavior of tutoring systems. *International Journal of Artificial Intelligence in Education* 16, 3 (2006), 227–265.
[72] Kurt VanLehn. 2011. The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist* 46, 4 (2011), 197–221.
[73] Kurt VanLehn, Winslow Burleson, Maria-Elena Chavez Echeagaray, Robert Christopherson, Javier Gonzalez Sanchez, Jenny Hastings, Yoalli Hidalgo Pontet, and Lishan Zhang. 2011. The affective meta-tutoring project: How to motivate students to use effective meta-cognitive strategies. In *19th International Conference on Computers in Education, Chiang Mai, Thailand*. Asia-Pacific Society for Computers in Education (APSCE), 128–130.
[74] Joshua Wainer, David J. Feil-Seifer, Dylan A. Shell, and Maja J. Mataric. 2007. Embodiment and human-robot interaction: A task-based perspective. In *16th IEEE International Symposium on Robot and Human interactive Communication (RO-MAN'07)*. IEEE, 872–877.
[75] Jacqueline Kory Westlund and Cynthia Breazeal. 2015. The interplay of robot language level with children's language learning during storytelling. In *Proceedings of the 10th Annual ACM/IEEE International Conference on Human-Robot Interaction Extended Abstracts*. ACM, 65–66.
[76] Philip H. Winne, John C. Nesbit, Vive Kumar, Allyson F. Hadwin, Susanne P. Lajoie, Roger Azevedo, and Nancy E. Perry. 2006. Supporting self-regulated learning with gStudy software: The learning kit project. *Technology Instruction Cognition and Learning* 3, 1/2 (2006), 105.
[77] Christopher A. Wolters, L. Yu Shirley, and Paul R. Pintrich. 1996. The relation between goal orientation and students' motivational beliefs and self-regulated learning. *Learning and Individual Differences* 8, 3 (1996), 211–238.
[78] Barry J. Zimmerman. 1990. Self-regulated learning and academic achievement: An overview. *Educational Psychologist* 25, 1 (1990), 3–17.
[79] Barry J. Zimmerman. 2008. Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal* 45, 1 (2008), 166–183.
[80] Barry J. Zimmerman and Manuel Martinez-Pons. 1988. Construct validation of a strategy model of student self-regulated learning. *Journal of Educational Psychology* 80, 3 (1988), 284.

Received June 2017; revised May 2018; accepted October 2018