DEVELOPING AN ENGAGEMENT AND SOCIAL INTERACTION MODEL FOR A ROBOTIC EDUCATIONAL AGENT

A Thesis
Presented to
The Academic Faculty

by

LaVonda N. Brown

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Electrical and Computer Engineering

Georgia Institute of Technology
December 2015

DEVELOPING AN ENGAGEMENT AND SOCIAL INTERACTION MODEL FOR A ROBOTIC EDUCATIONAL AGENT

Approved by:

Professor Patricio Vela,
Committee Chair
School of Electrical and Computer Engineering
Georgia Institute of Technology

Professor Ayanna Howard, Advisor
School of Electrical and Computer Engineering
Georgia Institute of Technology

Professor Andrea Thomaz
School of Interactive Computing
Georgia Institute of Technology

Professor Jeffrey Davis
School of Electrical and Computer Engineering
Georgia Institute of Technology

Professor Elizabeth DiSalvo
School of Interactive Computing
Georgia Institute of Technology

Date Approved: 10 November 2015

To my family – Mom, Dad, Brittany, Maxine, and Blossom.

ACKNOWLEDGEMENTS

First and above all, I praise God for providing all the resources, love and support needed to obtain this degree. This thesis appears in its current form due to the assistance and guidance of several people. I would therefore like to offer my sincere thanks to all of them. In particular, I would like to thank my advisor, Dr. Ayanna Howard, whose support has made this thesis possible. I would also like to thank the reading committee Dr. Patricio Vela, Dr. Andrea Thomaz, Dr. Jeffrey Davis and Dr. Elizabeth DiSalvo for their incredibly valuable time and input. I truly appreciate the HumAnS Lab research group for their support and feedback throughout this project, my weekly accountability group who ensures that I am always making steady strides towards my goals, and VOW for playing an active role in my spiritual growth. My family has also played a major role in my tenure here at Georgia Tech, and I want to thank them for all that they have done to minimize stress and, in effect, enable me to focus on my education. Last but not least, I would like to thank my dog Blossom who has been there with me through the trenches and never left my side.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

Effective educational agents should accomplish four essential goals during a student's learning process: 1) monitor engagement, 2) re-engage when appropriate, 3) teach novel tasks, and 4) improve retention. In this dissertation, we focus on all of these objectives through use of a teaching device (computer, tablet, or virtual reality game) and a robotic educational agent. We begin by developing and validating an engagement model based on the interactions between the student and the teaching device. This model uses time, performance, and/or eye gaze to determine the student's level of engagement. We then create a framework for implementing verbal and nonverbal, or gestural, behaviors on a humanoid robot and evaluate its perception and effectiveness for social interaction. These verbal and nonverbal behaviors are applied throughout the learning scenario to re-engage the students when the engagement model deems it necessary. Finally, we describe and validate the entire educational system that uses the engagement model to activate the behavioral strategies embedded on the robot when learning a new task. We then follow-up this study to evaluate student retention when using this system. The outcome of this research is the development of an educational system that effectively monitors student engagement, applies behavioral strategies, teaches novel tasks, and improves student retention to achieve individualized learning.

CHAPTER I

INTRODUCTION

Effective instructors are able to improve learning by observing and maintaining student engagement in real-time [42]. Through their observation, instructors are able to interject behavioral cues essential for optimal learning in the form of instruction, motivation, and correction. This paradigm is further expressed in Figure 1.

Figure 1: An effective instructor performing three behavioral cues essential for optimal learning: instruction, motivation, and correction.

Instructors are not only limited to academic teachers. They are defined as any individuals who teach new tasks. For example, personal tutors, physical therapists, team-sport coaches, and fitness trainers are all considered instructors for the scope of this research. Furthermore, an academic teacher may instruct subjects such as math or english, whereas a physical therapist may instruct proper movements necessary for motor skill rehabilitation.

The primary issue is that there is a shortage of effective instructors across multiple domains. Namely, there is a great need for more effective academic teachers and physical therapists. According to the U.S. Department of Education, the shortage of teachers in the traditional classroom setting is as high as 94% (Figure 2). According to the U.S. Bureau of Labor, the demand for physical therapists is predicted to grow as much as 27% by 2016 (Figure 3). The need for both academic teachers and physical therapists are increasing at a rate that is difficult for human instructors to adequately fill. Nevertheless, this shortage of professional instructors has led to many innovative solutions such as the use of more learning technologies.

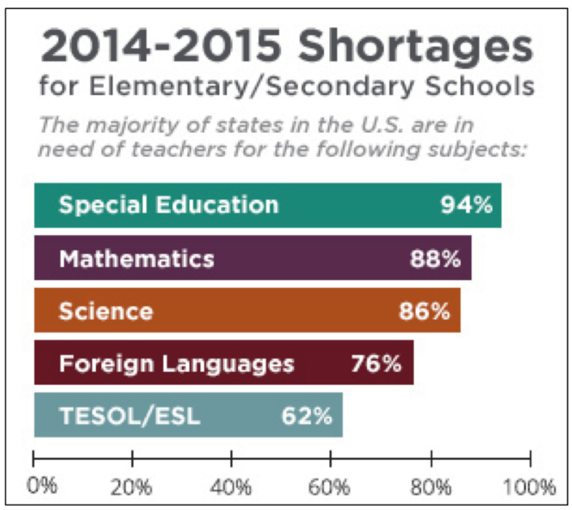


Figure 2: The shortage of teachers nationwide from 2014-2015 [66].

1.1 Effective Instructors

The benefits of effective instruction have been proven through a study conducted in the Boston Public School System [79]. In Figure 4, the learning of students who had the "least effective" teacher was compared to the students who were taught by the "most effective" teachers. The average student gains for students with the most effective teacher in math was 14.6, whereas the average student gains associated with the least effective teacher were negative. A negative learning gain can be the result

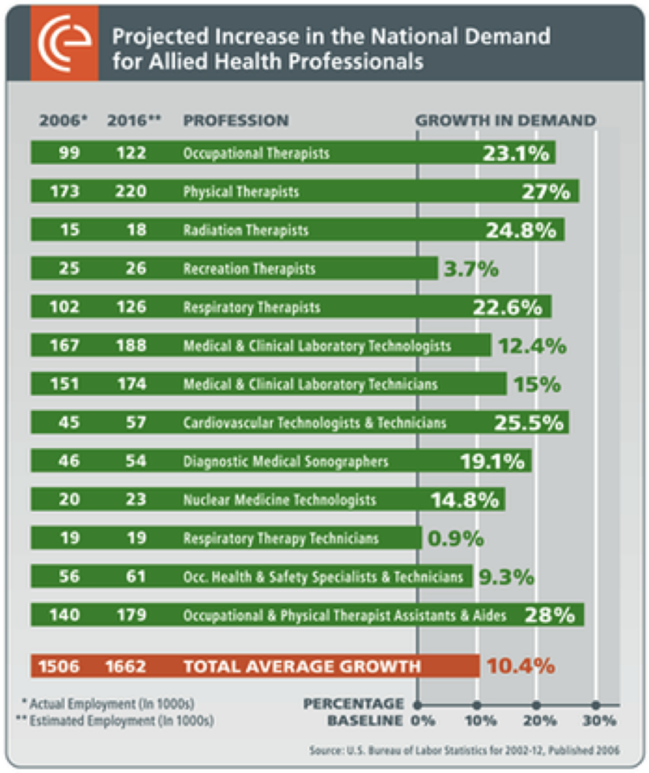


Figure 3: The projected increase in the national demand for allied health professionals [67].

of not properly engaging each student in the classroom.

Not having an effective teacher also has long term effects on student development as shown in Figure 5. Let us imagine that there is a student named Nicole who enters middle school in the 50$^{\text{th}}$ percentile of her class. By the time she matriculates through middle school with a high-performing teacher, she has the potential to reach the 90$^{\text{th}}$ percentile in her class. On the contrary, matriculating through middle school with a low-performing teacher could be detrimental to her academic development and result in her leaving in the 37$^{\text{th}}$ percentile.

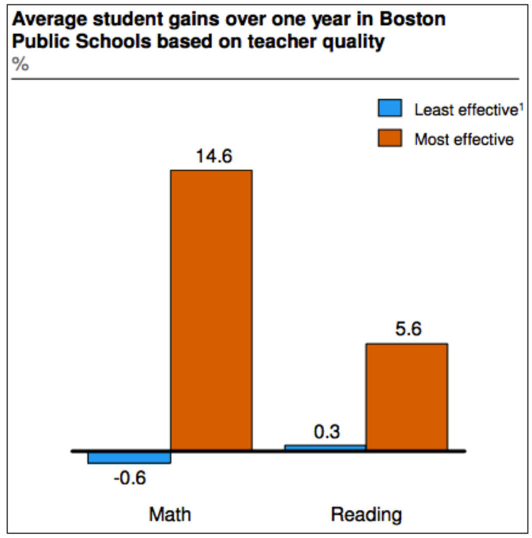


Figure 4: A study conducted in Boston City Public Schools to analyze the cumulative and residual effects of ineffective teaching on student achievement within one academic year. [79]

1.2 Learning Technology

Learning technologies have been a vastly growing option to combat the shortage of effective instructors and lack of motivation in the traditional learning environment. This is where traditional learning is integrated with technologies such as computers. These technologies have the potential to increase performance and motivation to learn due to the deviation from traditional, mundane classroom instruction. Learning technology is a broad term used to describe any "computer" use in an "educational setting." For the context of this dissertation "computer" includes laptops, desktops, tablets, smartphones, video games, and virtual reality (VR) systems. In addition, an "educational setting" includes any environment where instruction is required to learn a new task, and a new task can range from a computational task such as solving a math problem or a motor task such as solving a movement problem. Although there are many advantages to using learning technologies, there are also a few disadvantages to consider. In particular, a major concern is that learning technologies

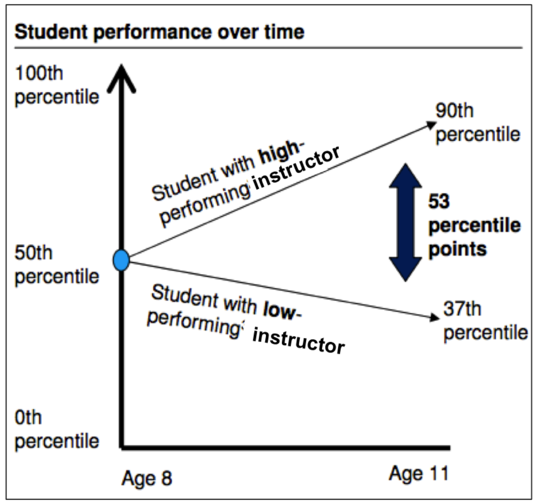


Figure 5: A study conducted in Boston City Public Schools to analyze the cumulative and residual effects of ineffective teaching on long-term student achievement. [79]

narrowly monitors engagement, which is a critical component to learning. As such, this dissertation will describe the benefits of developing a Robotic Educational Agent (REA) that uses both the principles of an effective instructor coupled with learning technology.

1.3 Robotic Educational Agent (REA)

In this dissertation, we aim to enhance learning through development of a robotic educational agent (REA) as shown in Figure 6. The REA is able to optimize learning by blending strategies to increase 1) engagement in a similar manner as an effective human instructor and 2) performance through a variety of learning technologies.

As such, we detail a system that integrates a real-time engagement model into a task learning scenario, discuss the processes employed on a robotic educational agent to re-engage the student using behaviors comparable to that of a human instructor, and develop a learning model that proves learning and retention is achieved. Chapter 2 provides a literature survey to discuss prior related work. Chapters 3-7 further discuss Contributions 1-5, respectively. Moreover, Chapter 3 provides an overview of

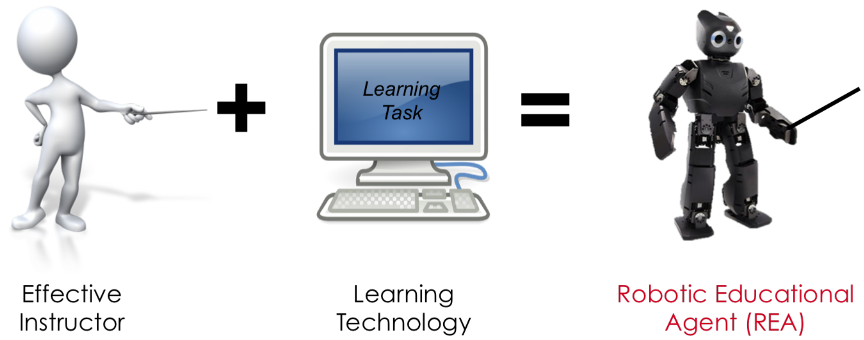


Figure 6: A Robotic Education Agent (REA), which is the result of integrating an effective instructor with learning technology.

the engagement model whereas Chapter 4 discusses the nonverbal and verbal behaviors used in the study. The behavioral strategies embedded on the humanoid robotic platform for motivational feedback in an active learning environment is discussed in Chapter 5, and the behavioral strategies used to provide guided instruction and corrective feedback are is discussed in Chapter 6. Chapter 7 discusses the learning model used to evaluate when retention has been achieved after learning a novel task. Finally, the overall conclusion is discussed in Chapter 8.

CHAPTER II

LITERATURE SURVEY

2.1 Monitoring Engagement

In successful classroom settings, teachers are able to observe the student's engagement in real-time and employ strategies to re-engage the student, which, in effect, improves attention, involvement and motivation to learn [84]. This is also true during one-on-one tutoring sessions because tutors are able to track engagement in real-time as well. In general, teachers are able to determine engagement by following behavioral cues from students such as direction of attention, posture, facial expressions, and responsiveness to instructional activity [30]. Understanding this behavioral engagement is a crucial component in education because it is often related to the academic achievement of a student [35, 44].

Currently, computer-based education (CBE), also known as learning technologies, is a widely used method of instruction inside the classroom and at home. Research has shown that CBE actually improves academic achievement [7] and student motivation [81] when compared to traditional classroom instruction. Using CBE reduces the amount of instructional time required and increases the students attitude towards learning [55]. Although research has shown CBE as being a highly effective learning tool, it pales in comparison to a human tutor [7]. Therefore, CBE should be used as a supplement to traditional instruction and not as a replacement [59]. In this dissertation, we will determine how our system can monitor student engagement in a manner comparable to that of real classroom teachers.

2.1.1 Common Methods

CBE mainly focuses on comprehension of material [8] and not real-time engagement, which is essential for optimal academic achievement. Comprehension of material is determined solely by the validity of answer selections. Many standardized tests today, such as the SAT and GRE, adapt to the students based exclusively on their responses. This type of evaluation is known as computerized adaptive testing (CAT) [87]. If the student answers a question correctly, he/she is given a more difficult problem. If the student answers a question incorrectly, he/she is given a problem of less difficulty. However, for an educational system to be optimum, it must ensure that the student is actively and continuously engaged. Computer-based tools mainly focus on comprehension because of the difficulty associated with determining cognitive states. Due to the variability of behavior, characteristics, and environment, computational methods with the capability of identifying the behavioral cues associated with engagement have yet to be developed [84].

As an alternative to measuring engagement in real time, scales have been created to evaluate motivation once the student has completed a system [74]. The problem with this method is that an educational agent will not be able to adapt to the educational needs of the student once the learning session has already been completed. The art of adaptation requires real-time information processing, which scales are unable to deliver.

A more promising alternative to measuring engagement is through electroencephalography (EEG) signal measurements. EEG signals are able to identify subtle shifts in alertness, attention, and workload in real-time [13]. Szafir and Mutlu used an EEG headset to monitor engagement in an educational setting through storytelling [84]. When the EEG signals would begin to drop during narration, adaptive behavioral cues (verbal and nonverbal) would be used to re-engage the students. EEG measurements have the advantage of being well studied and low cost [84]; however, wearing a headset creates a controlled testing setup, which does not convey a natural learning environment. This ultimately has the potential to cause unnecessary distractions and distort results.

2.1.2 Eye Gaze Methods

In efforts to create a non-invasive tool to monitor engagement in real-time and within a natural learning environment, a viable option would be to use eye gaze. Research has shown that there are many neural components related to vision. Posner et al. investigate how people are able to visually fixate on one location while mentally focusing on something else [71]. There is currently no way of determining what a person is thinking about while looking at something – this would require recording eye gaze and brain activity. An investigation was performed to obtain the relation between eye fixation and neural activity with monkeys. The results showed that the activity of the prefrontal (PF) neurons was influenced by the task being performed [71].

Smith et al. performed a study to evaluate the correlation of eye movements and 1) awareness and 2) hippocampus-dependent memory [83]. Participants viewed images that were novel, repeated, and manipulated. The participants either had no memory problems or were memory impaired patients (with damage to the hippocampus). There were three experiments – Experiment 1 assessed awareness of manipulation after all the images had been viewed; Experiment 2 assessed awareness after each scene was presented; Experiment 3 assessed the memory impaired patients.

The results of both Experiment 1 and 2 showed that participants made fewer fixations and sampled fewer regions when viewing familiar compared with novel images. In addition, fixations for familiar images were longer than fixations for novel images. This work suggests that simple repetition is enough to change viewing behavior. Also in Experiment 1 and 2, measured awareness was very similar. When participants were aware of a change, their eye gaze was directed more to the critical region that had been changed than the unchanged critical region in repeated images. However, when the subjects were unaware of a change, the eye gaze was similar to showing a repeated image. Lastly in Experiment 3, the subjects had trouble deciding whether images were novel, repeated, or manipulated. The patients exhibited the same level of confidence for incorrect and correct responses. Therefore, ability to correctly classify images is dependent on hippocampus-dependent, declarative memory.

Asteriadis et al. were able to use common patterns found in eye gaze to evaluate engagement while learning. More specifically, they developed a system using head pose and movement, direction of eye gaze, as well as measurements of hand gesture expressivity to determine six user-states in an e-learning environment: attentive, full of interest, frustrated/struggling to read, distracted, tired/sleepy, and not paying attention [9]. The developed system was able to effectively detect reading- and attention-related user states very well when subjects were asked to read/watch an electronic document (web page, multimedia presentation, video clip). However, this system was not tested in a complex problem solving or test-taking environment. As such, we aim to develop a method of monitoring engagement that is adequate for an environment where high-cognitive thinking is prevalent.

2.2 Effective Emotion in Learning

In addition to monitoring engagement, educational agents must also maintain or increase the student's level of engagement through use of various communication modalities. Because the interaction between the student and the teacher is best modeled as being a social dialog [78], it is essential to delve deeper into the methods used to improve the social interaction exhibited in various learning scenarios. Emotions yield a natural form of communication, in that they can be shown visually through facial expressions, vocal expression, and actions/body movements. When certain emotions are integrated into social settings they have the capability to create a comfortable, welcoming environment for all parties. This, in effect, will increase a person's willingness to engage in the social interaction. Moreover, in the realm of human-robot interaction (HRI), emotions have been shown to enhance the social interaction involved with education [78, 51, 84, 68, 61], motor-task learning [15, 48], play partners [69, 57], companions [52, 57], elderly care [85], and weight-loss [52].

One of the key uses of emotions in HRI scenarios is to build a bond between the two entities. Typically, this bonding mechanism can be enhanced by having the robot exhibit forms of empathy. Empathy is a key factor used to enforce socially supportive behaviors [86]. Smiling and showing sensitivity to the individual's emotions enhance the interpersonal relationship, which ultimately leads to increased enthusiasm and learning [86]. In [78], Saerbeck was able to implement empathy best by simply having a robotic agent smile (happy face) when a task was completed correctly and frown (sad face) when the task was completed incorrectly. This study showed that the appropriate expression of empathy in a social interaction scenario is best visualized through a happy-sad continuum as shown in the circumplex model of affect (Figure 7) [77]. Arousal is represented in the vertical axis (quiet-active continuum), whereas valence is represented in the horizontal axis (happy-sad continuum). The use of other emotions such as anger, surprise, and nervousness as feedback were shown not to be essential for active engagement.

2.2.1 Nonverbal Methods

Because body movement can be used to enable a robot to show forms of emotions, we focus of the impact of nonverbal cues that increase the quality of interaction in learning scenarios. During a case study involving a humanoid robot and children, [12] was able to analyze the effects of upward and downward head movement relative to positive and negative emotion. The humanoid robot was programmed to have six

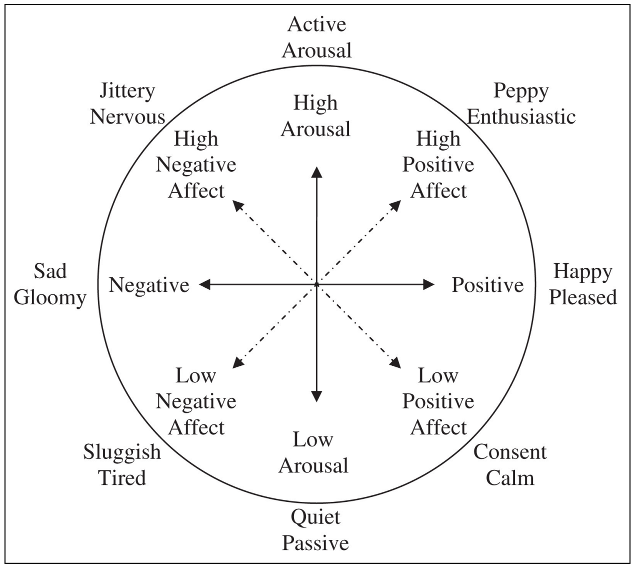


Figure 7: The circumplex model of affect. [77]

different base poses: anger, sadness, fear, pride, happiness, and excitement. Within each base pose, the head was positioned either up, down, or forward to make a total of 18 poses. The results showed that moving the head up improved the identification of pride, happiness, and excitement, while moving the head down improved the identification of anger and sadness. Fear was identified well regardless of the head's position. In general, moving the head up can enhance positive emotions, while moving the head down can enhance negative emotions.

In a similar study, Li and Chingnell analyzed how simple head and arm movements were able to communicate emotion in social robots [56]. Here, they used a teddy bear to implement various arm and/or head movements. They concluded that when head movements were compared to arm movements, arm movements overall were perceived to be more lifelike. They also stated that these simple gestures alone do not provide a lot of information and recognition is low, which suggests that another medium to communicate emotion is needed.

Schegloff discussed the different effects achieved when changing the upper body parts versus the lower body parts [80]. He categorized the body into nine different stances: stance pose, hip pose, torso pose, shoulder pose, head pose, hip torque, torso torque, shoulder torque, and head torque. He concluded that lower body movement suggests "dominant involvement," whereas upper body movement suggests "subordinate involvement." This could possibly mean that lower body movements have more extreme effects on emotions, whereas upper body movements have less extreme effects. Although extreme emotions can be thought of as being less natural, in the realm of robotics, interaction is actually enhanced with the use of exaggerated motions [41]. In particular, Gielniak and Thomaz were able to present evidence that engagement is increased along with perceived entertainment value by over emphasizing movements during social interaction [41].

In an investigation involving the communication of musical expression through robotic gestures [26], robotic movements were derived from a perceptual test done by Dahl and Friberg [33]. In [26], Burger and Bresin show how the use of variables such as amount, speed, fluency, regularity, and direction on a mobile robotic platform are able to convey happiness, anger, and sadness (Table 1). They also incorporated the work of [50], which stated that round shapes convey positive emotion and sharp/spiky objects oftentimes convey negative emotions. In result, Burger and Bresin had the robotic platform perform fluent, circular movements to convey happiness and jerky, sharp movements to convey anger [26].

Table 1: Implementation of the robot's movements [26]

Movement Cue	**Happiness**	**Anger**	**Sadness**
Amount of Gesture	Large	Large	Small
Speed	Fast	Fast	Slow
Fluency	Fluent	Jerky	Fluent
Regularity	Regular, circular	Irregular	Regular
Direction of arm Movements	Upwards	Fast up & down	Slow up & down

2.2.2 Verbal Methods

Another primary method used for engagement in the classroom environment is the use of verbal cues, which can be used to encourage the student, provide instruction, and give positive praise. As such, we focus on the impact of verbal cues in the learning environment. By changing acoustic characteristics such as tempo, pitch, intensity, voice quality, and articulation, verbal cues, or behaviors, can be used to evoke a range of emotions that impact student engagement [14, 65, 49, 64, 70, 27]. In [64], Murray and Arnott summarize how the human voice is affected by emotions such as anger, happiness, sadness, surprise, and disgust (Table 2).

Table 2: Effect of emotions on human speech [64, 70]

	Fear	**Anger**	**Sorrow**	**Joy**	**Disgust**	**Surprise**
Speech Rate	much faster	slightly faster	slightly slower	faster or slower	very much slower	much faster
Pitch Average	very much higher	very much higher	slightly lower	much higher	very much lower	much higher
Pitch Range	much wider	much wider	slightly narrower	much wider	slightly wider	
Intensity	normal	higher	lower	higher	lower	higher
Voice Quality	irregular voicing	breathy chest tone	resonant	breathy blaring	grumbled chest tone	
Pitch Changes	normal	abrupt on stressed syllable	downward inflections	smooth upward inflections	wide down- ward terminal inflections	rising contour
Articulation	precise	tense	slurring	normal	normal	

Prior research has shown that these human speech ideals can effectively be implemented on a robotic platform. For example, in [14], Breazeal was able to express emotion on a robotic platform by correlating human speech ideals (Table 2) into a robotic speech synthesizer. Participants were able to perceive the robots intended emotion in most cases; however, there were a few misclassifications when the emotions shared negative valence or high arousal (i.e. angry and disgust, happy and excitement).

In addition to acoustic characteristics, sentence structure, language markers, and vocabulary choice indirectly shape the social interaction between the agent and student [65]. For example, age appropriate vocabulary is needed to maintain the student's level of engagement, and by adding markers such as "please" and "thank you," the agent can be perceived as being very polite [24]. Mutlu performed an investigation where he studied human communication and explored how robots would be able to convey the same rich social outcomes of learning, rapport, and persuasion [65]. Through combinations of verbal, vocal, and nonverbal cues, Mutlu was able to observe how embodied communication cues can be useful in enhancing social interaction in HRI.

2.3 Learning and Retention

After the learning process, it is important to evaluate *if* and *when* information is being stored in one's memory for easy recall later. This concept is known as retention. Retention lays a foundation to build upon and enables deeper learning over a long period of time. Because of this, it is important to evaluate when retention has occurred as well as the steps needed to reach this goal. When an instructor is able to identify when and how retention has been obtained for each individual student, learning can, in effect, be optimized long-term.

To evaluate human task performance during the learning process, previous research uses the learning curve as a viable metric. During a learning scenario, the student's performance, P, is dependent on the number of instances or trials practiced, N. The initial performance of the student is described by B. In the instances where there exists prior knowledge or practice before the scenario, N_0 represents the number of initial trials. Therefore, $N+N_0$ describes the total number of practice trials. Lastly, the learning rate of the student is denoted by β, which gives great insight on the quality of the instruction. The learning curve has been described through both exponential and power-law functions. The exponential family of functions are shown in (1) and (2), and its respective curve is shown in Figure 8.

$$P(N) = A + Be^{\beta(N+N_0)} \tag{1}$$

$$P(N) = Be^{\beta N} \tag{2}$$

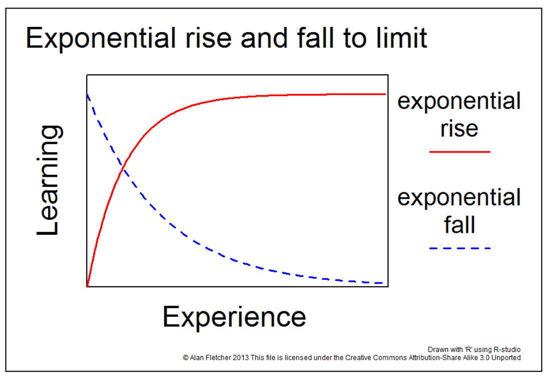


Figure 8: The learning curve described by the exponential family of functions [5].

The power-law family of functions is shown in (3) and (4), and its respective curve is shown in Figure 9. Although parameters A and N_0 allow these functions to be more accurate, it is common practice to equate both to zero, which results in the simplified equations (2) and (4).

$$P(N) = A + B(N + N_0)^{\beta} \tag{3}$$

$$P(N) = BN^{\beta} \tag{4}$$

The power-law learning curve is typically used in a classroom setting where instruction is dependent on what remains to be learned. This "power law of practice"

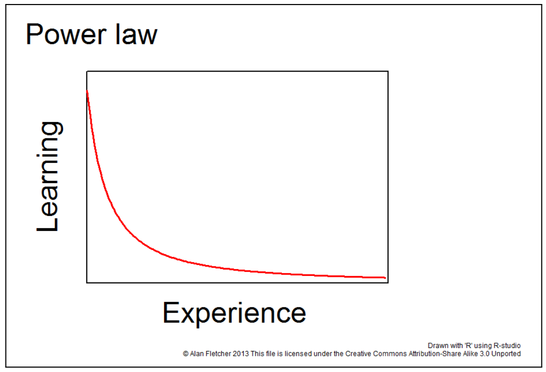


Figure 9: The learning curve described by the power-law family of functions [6].

theory is supported by the research performed in [75, 73]. However, the exponential learning curve is more useful when evaluating a single student's performance [47]. As such, since we aim to achieve individualized learning in this dissertation, learning and retention will be evaluated using equation (2).

In essence, the learning curve states that performance increases with experience or practice. In addition, after n trials, the student's performance begins to plateau and learning is achieved (Figure 8 and 9). It has been theorized that by reaching this "plateau," retention can be increased; however, more research needs to be conducted to reveal the exact correlation between learning and retention.

2.4 Robots in Education

The primary objective of tutoring, a practice designed to supplement classroom-based learning, is to assist and guide students to become independent learners. To be effective in this practice, a human tutor must provide direction maintenance – i.e. when the learner disengages from the task at hand, the role of the tutor is to keep him or her in pursuit of the specified objective [88]. Studies have shown that humans are more responsive to completing tasks when there is a physical robotic embodiment versus using CBE methods alone such as remote virtual agents [10]. In addition, it has been theorized that robotic-based education (RBE) methods can approach the effectiveness of human tutors by coupling instructional methods in CBE with human-equivalent behavioral cues of engagement. By using social cues, a long-term relationship between the robot and the subject can be fostered [52]. This relationship drastically increases the subject's motivation to complete a task and the subject's desire to spend time with the robot for a long period of time. In addition, ample studies have shown that the effect of being perceived as a social interaction partner can be enhanced by a physical robotic embodiment [72]. These characteristics are ideal for a student interacting with a robot tutor in a learning environment. As such, in efforts to analyze the RBE approach, research has been conducted on both implementing sociable [78, 52, 61, 68, 53, 57, 54] and educational robots [78, 61, 46, 68, 57, 90, 51, 54].

In the realm of education, robots are currently being used to teach math [51], history [68], new languages [78, 90, 54], motor tasks [89, 32, 31, 58, 25], and new tasks [61, 57]. Some studies vary the type of feedback (positive, negative, neutral) [68] and behavioral techniques [84] given from the robot, while others vary the type of learning adaptation or scaffolding [51] provided from the system. Generally speaking, students are more attracted to the robot when it exhibits positive feedback [78, 68], are more motivated to learn from the robot when there is individualized learning [61, 51], and have increased recall abilities when the robot uses appropriate behavioral techniques to re-engage [84].

Saerbeck et al. investigated whether or not social engagement with a robot interface could effectively by applied to education [78]. Their research was done with an interactive cat (iCat) whose goal was to teach a new language to a child. The iCat platform has the shape of a cat, and its height is approximately 40 cm. The study compared a socially supportive iCat (engaged in social dialog) to a neural iCat (unidirectional knowledge flow). The students involved in the socially supportive iCat case were more motivated, which is essential for any educational technology to have long-term effectiveness.

Kory and Breazeal developed a system that used a robotic learning companion to improve language development through social storytelling [54]. The DragonBot robotic platform was used in this study, which has "squish and squash" principles of animation that allow for expressive movements. In addition, a smart phone is used as the robot's face. learning companion to teach children oral language skills. Kory and Breazeal concluded that strategically matching or mismatching the robot's ability to the student's could potentially improve language learning outcomes.

Michaud et al. believed that mobility, appearance, interaction modalities, and behavior all influence a child's ability to sustain interest and learn [61]. They used Roball, a spherical robot with a diameter of approximately 15 cm, to evaluate autonomous motion in children 12-18 months old. An algorithm was used to adapt the robot's behavior to the child's interaction using proprioceptive sensors. Michaud et al. concluded that mobile robots as assistive technology are great for creating interplay and learning situations. Mobile robots allow adaptation to children and the environment, and they keep children engaged.

Han et al. of Korea developed the world's first e-learning home robot (IROBI) in March 2004 [46]. IROBI, a humanoid robot consisting of only a head and torso, demonstrated the prospect of robots as a new educational media. Users could interact with IROBI using voice and a touch panel, and the robot communicated with people by presenting voice, gestures, and multimedia contents on an LCD screen. During this investigation, Han et al. compared traditional media-assisted learning and web-based instruction (WBI) to Home Robot-assisted learning [46]. Han et al. concluded that IROBI was the most effective in promoting and improving the child's learning concentration, interest, and achievement when compared to other instructional media.

It is theorized that these learning scenarios can be further enhanced by integrating a virtual reality (VR) gaming component into the teaching sessions as a form of CBE [39]. Virtual reality refers to a computer technology that creates a three-dimensional (3D) virtual context that allow for interactions by the user [29]. In general, gaming with instruction allows for flexible adaptation to each user's learning level, real-time feedback from the system, and increased engagement through the use of hidden teaching tactics. Studies have been conducted that couple VR games with REAs focused on teaching motor tasks [32, 31, 58, 25]. Each of these studies compared robotic-assisted instruction alone to a combination of robot-assisted instruction and VR gaming. Overall, the results showed that the addition of interactive VR gaming was able to increase motivation, but not necessarily improve the user's learning or performance.

2.5 Summary

Prior work does not effectively monitor student engagement in real-time, lacks a consistent framework for developing gestural behaviors for robotic agents, and does not consistently improve both motivation and learning. As such, in this dissertation, we detail a system that addresses each of these concerns with the development of the Robotic Educational Agent (REA).

CHAPTER III

ENGAGEMENT MODEL

In this dissertation, we seek to develop a robotic educational agent that can interactively function in an equivalent manner as a human tutor. The preliminary work to achieve the objective of monitoring student engagement is detailed in the following sections. We begin by developing an engagement model based on the interactions between the student and the teaching device (tablet, computer, or virtual reality game) in Section 3.1 [18, 16]. We later expand the model by developing an eye gaze algorithm based on student fixations and saccades in Section 3.2. As such, this chapter leads to our first contribution:

> *1. Develop an engagement model, which yields a non-invasive method of monitoring student engagement when performing high-level cognitive tasks.*

3.1 Human-task Input

In this section, we discuss our preliminary engagement model which uses techniques that determine behavioral user state and correlate these findings to human-task inputs, measured by mouse and keyboard events [18, 16]. Event processes are observed to identify a common pattern associated with an engaged versus a disengaged student. We evaluate the correctness of our model based on an investigation involving a middle-school after-school program in which a 15-question math exam that varied in cognitive difficulty was used for assessment. The eye gaze technique described in [9] is referenced for the baseline comparison model for engagement. We conclude the investigation with a survey to gather the students' perspective of their mental state.

3.1.1 Engagement Metrics

We define a learning task as a series of problems that must be solved by a student. An engagement model is thus defined as a model that can correctly assess the level of engagement of the student while involved in the learning task (Figure 10). The level of engagement is defined by a series of *on-* and *off-task* events. The human-task inputs consist of all physical input events that each student uses during the learning task. We observe all the events post-task and derive a model based on the common patterns present for all students.

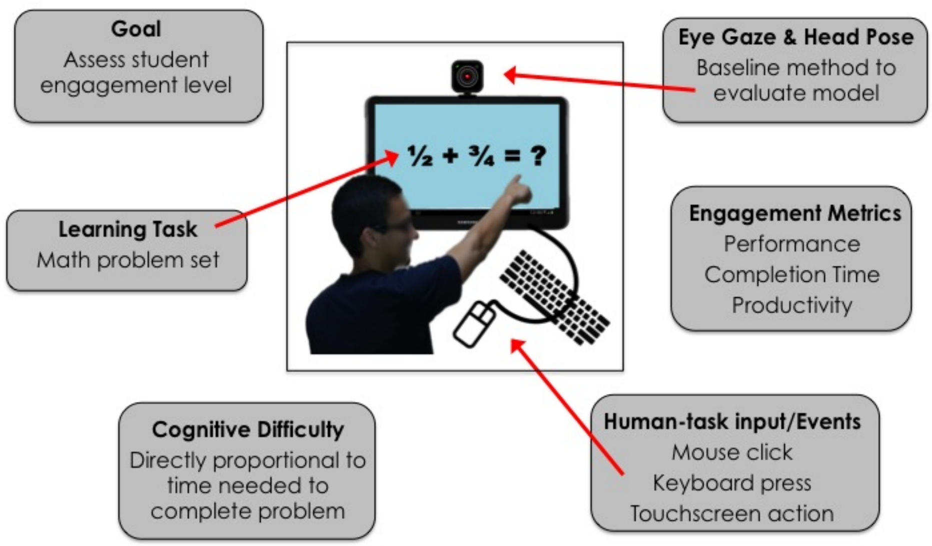


Figure 10: Diagram of the entire system with the associated engagement metrics.

We first conduct a pilot test to derive a baseline model. We then define the difficulty level of the problems to allow for future adaptation to a variety of problem sets. The difficulty of a problem is directly proportional to the amount of time needed to submit a response. Based on the results of the pilot studies, we are able to categorize the student's response as either slow, average, or fast. The problems that require less time to complete are categorized as low-level cognitive problems, whereas the problems that require more time are categorized as high-level. We also monitor response accuracy, which is defined as the correctness of the submitted response.

In this investigation, our learning task is focused in a math task (i.e. cognitive learning task). Specifically, we focus on achievements on a math test. We implement the engagement model by first identifying a set of events needed to effectively navigate through the math test. If an input event or combination of input events fall within the approved list needed to properly execute a function, the student is classified as being *on-task*. Otherwise, the student is classified as being *off-task*. We monitor the input events over a period of samples, which consists of n = *8* events. If more than p = *25%* of the sample is classified as being off-task, then the entire sample will be classified as off-task. The inequality used to determine when a series of input events is on-task is shown in (5).

$$\frac{1}{n}\sum_{i=m}^{m+(n-1)} x_i < p \quad OR \qquad \frac{1}{n}\sum E < p, \qquad \forall m = 1, 2, ..., n \tag{5}$$

The subset of n sequential events is categorized by $\boldsymbol{E}$ which is defined as $\{x_m + x_{(m+1)} + ... + x_{(m+(n-1))}\}$, where x is an event, m is the initial event in the series, and n is the total number of events observed in the series. Each event that effectively executes a function yields a value of one, and each event that does not execute a function yields a zero. Equation 5 is computed for every sample until there is no longer a subset of n events to evaluate.

In [9], six user-states were defined based on eye gaze for an e-learning environment to categorize if the user was attentive, full of interest, frustrated, distracted, sleepy, and not paying attention. We combined these categories to form two basic user-states – engaged and disengaged. Attentive, full of interest, and frustrated are classified as engaged, while distracted, sleepy, ot not paying attention are classified as disengaged. While it might not seem that a user-state of frustrated should be classified as engaged, one is only frustrated when he or she is dedicating attention to a particular task. However, frustration typically leads to being disengaged if the focus of frustration is not resolved in a timely manner.

The amount of time that is classified as disengaged, $T_{disengaged}$, will be recorded along with the total time needed to complete the math test, T_{total}. All of this data is used to derive the percent error associated with Asteriadis et al.'s eye gaze and head pose model [9] as shown in (6).

$$Percent\ Error = \frac{T_{disengaged}}{T_{total}} x100\% \tag{6}$$

3.1.2 Hypotheses

Two hypotheses were developed for our system:

1. The student is engaged if his or her series of events are classified as: on-task and correct (regardless of speed) or on-task, slow or average, and incorrect.

2. Eye gaze and head pose will not be an accurate measure of user state/engagement for high-level cognitive questions.

3.1.3 Experimental Design

To explore the trends developed over time associated with engagement in computer-based education (CBE), we designed and conducted a pilot study in which students completed a computer-based math test of varying difficulty. A total of 13 students took part in this experiment and all were recruited middle school students from an afterschool program in Atlanta, GA. The population consisted of both females and males in the age range of 10-14 years old (Male: 6, Female: 7; Sixth grade: 2, Seventh grade: 5, Eighth grade: 6).

The evaluation consisted of two segments to assess how well the engagement model performed when compared to eye gaze techniques. The initial validation of the engagement model's performance consisted of analyzing nine questions of low difficulty, which required low-level cognitive thinking. This segment was directly followed by analyzing six questions of high difficulty, which required high-level cognitive thinking. The questions were taken from Georgia's Criterion-Referenced Competency Tests (CRCT) [2].

The students were placed in a normal testing environment within a school, as shown in Figure 11. Due to the size of the classroom, all 6 males were tested as the first group followed by 6 of the females. An additional female was tested alone. The instructions provided to the student were the following:

> *"You will take a 15-question math test. It does not matter how well you perform, and I do not expect you to know all of the answers. However, it is important that you stay focused on each question, give it your best effort, and avoid being sidetracked."*

Each student was provided pencil and paper, which was initially placed next to the laptop (Figure 11). As the student navigated through the test, the mouse and keyboard events were recorded to determine total time, response accuracy, and proper event execution. We also used a web camera to monitor eye gaze and head pose, which was used to estimate behavioral user state throughout the test. The camera was also used to perform video observations once the test had been completed.

We designed three 15-question math tests to assess our hypotheses one for each grade level. The basic layout of each test is as shown in Figure 12. Boardmaker Plus is the software that was used to create the math program [1]. The tests were designed for students between the 6^{th} and 8^{th} grade. Nine questions on the test were low difficulty and required low-level cognitive thinking to complete. Most, if not all, of those problems can be computed quickly using mental math because they only require one processing step to answer. However, six questions were high difficulty and

Figure 11: This is the small classroom where all the testing took place. Each student had their own laptop, pencil, and scratch-paper.

required high-level cognitive thinking to complete. Most, if not all, of those problems cannot be computed quickly using mental math because they require multiple steps to answer. In many cases, pencil and paper may be needed to develop an answer.

3.1.4 Results

When analyzing the results, we took all the engagement metrics into consideration, total time, response accuracy, and proper function execution, as well as eye gaze and the exit survey. Table 3 summarizes the data collected using the preliminary engagement model as compared to the eye gaze technique. The total time needed to complete each question was calculated and shown in Figure 13 for the 7th and 8th grade. Because there were only two students in the 6th grade, all of their responses were automatically classified as being of average speed. Using a boxplot, we were able to properly divide the remaining data into its respective quartiles and categorize any outliers as slow or fast. Due to the nature of the box and whisker plot, there will always be a similar distribution between the average, slow, and fast categories as shown in Figure 14(c). In addition, the students answered 45% of the question correctly and 55% of the questions incorrectly as shown in Figure 14(d).

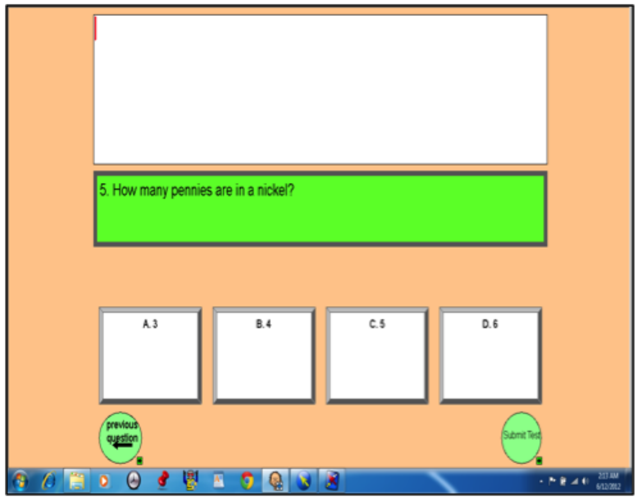


Figure 12: The basic layout of each question on the math test is shown. At the top of the interface, the student was able to type the answer into a textbox. The question was stated in the center of the screen within a green box. Below the question, the multiple-choice selections were displayed as rectangular buttons.

We also observed that 4% of the keyboard and mouse input was classified as being off-task as shown in Figure 14(b). Through use of the list of approved events, we were able to determine if the mouse clicks and keystrokes occurred within the necessary constraints to successfully navigate through the test. Figure 14(a) shows the combinations of events that model engagement and how often each combination occurred during this study.

The term error refers to the amount of time the head pose and eye gaze technique discussed in [9] categorized the student as either distracted, sleepy, or not paying attention. Through use of (6), the eye gaze and head pose technique had an average of 24.2% error for the 6^{th} grade, 41.1% error for the 7^{th} grade, and 34.8% error for the 8^{th} grade. This error suggests that the eye gaze and head pose is not the best measure of engagement. In addition, for the students who scored considerably higher than their peers, they exhibited up to a 65% eye gaze error. Figure 15 shows the relationship between the students test score and the amount of time his or her gaze

Table 3: Summary of results

Grade & Difficulty	**Avg. Score**	**Avg. Time**	**Developed Model**	**On-task (Eye-Gaze)**
6^{th}	*57%*	*13min*	*93%*	*76%*
Low	50%	36s	100%	
High	67%	81s	83%	
7^{th}	*39%*	*10min*	*96%*	*59%*
Low	44%	29s	100%	
High	30%	59s	90%	
8^{th}	*46%*	*15min*	*98%*	*65%*
Low	46%	33s	100%	
High	45%	101s	94%	

was not directed towards the screen.

Following the math test, 5 questions were asked about each question. Three of the questions were based on a 5-level Likert scale, one required a yes/no response, and the last was multiple-choice. Table 4 shows the results of the 3 Likert questions, and the mean and standard deviation are computed based on 195 samples (13 students x 15 questions). Overall, the students agreed that they were engaged for each question in the complete test with an average score of 3.96 (Agree = 4, SD = 1.40). They agreed that they understood the questions with an average score of 3.71 (Agree = 4, SD = 1.38). Lastly, the students agreed that they knew how to solve the problems with an average score of 3.56 (Agree = 4, SD = 1.41).

Table 4: Statistical analysis of exit survey

Statement	***m***	**SD**
I was engaged.	3.96	1.40
I understood the question.	3.71	1.38
I knew how to solve the problem.	3.56	1.41

Table 5 shows the results of the multiple-choice question that asked how each answer selection was decided. The options were either that the student made a

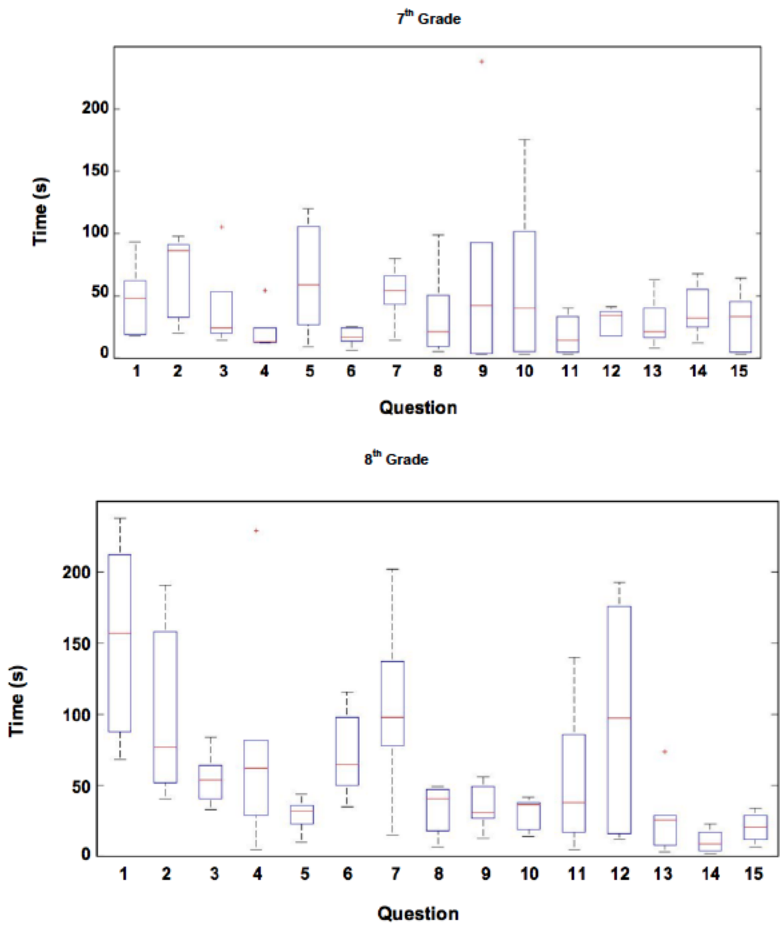


Figure 13: (a) Total time required per question for 7$^{\text{th}}$ grade (top) and 8$^{\text{th}}$ grade (bottom).

random guess, an educated guess, or no guess/solved the problem. This may also give some insight on how well the students believed they understood each question and, furthermore, reflect their confidence level. Lastly, Table 6 shows that the students on average used pencil and paper to solve the problems 56% of the time.

3.1.5 Discussion and Conclusion

Across all students/tests, less than 5% of the samples were classified as being off-task, which is statistically significant. This suggests that there is a direct correlation between an engaged student and our method of calculating on-task events. Moreover,

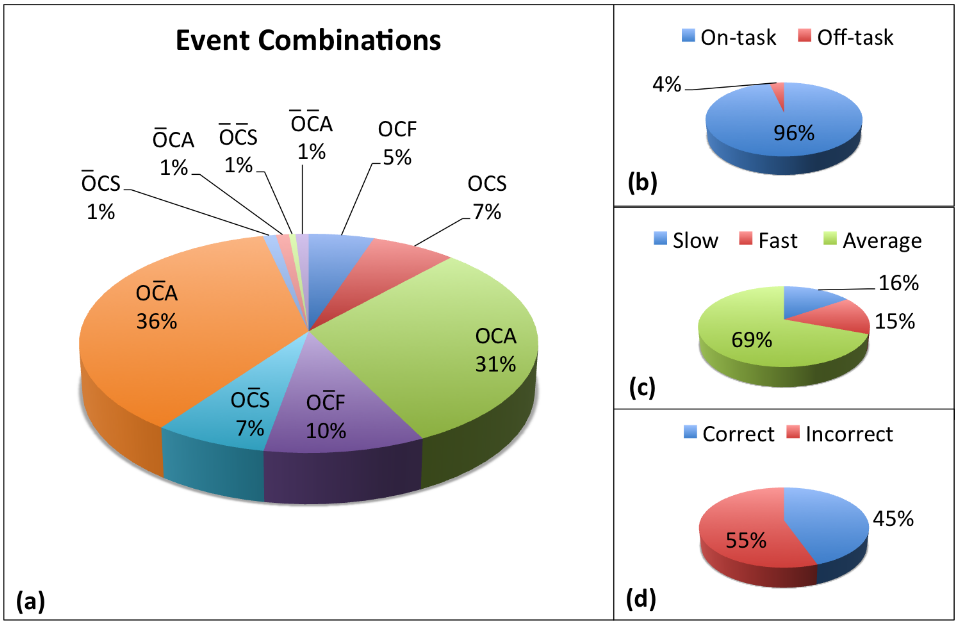


Figure 14: (a) This chart shows the how often we received each combination of events (S=slow, A=average, F=fast, C=correct, C̄=incorrect, O=on-task, Ō=off-task). (b) O vs. Ō events. (c) Speed of responses. (d) C vs. C̄ responses.

Table 5: Student's confidence of response

Selected Response	**Total**	**Percentage**
Solved	101	52%
Educated guess	53	27%
Random guess	41	21%

if a student is classified as being on-task, he or she is engaged (regardless of speed or response), which proves Hypothesis 1.

Furthermore, validity of responses alone is not enough information to determine user-state as exhibited in Figure 14(d). Speed coupled with the validity of responses can help to determine more information about the engaged student. If the student is on-task and has a series of fast responses with a series of correct answers (OCF), the student may need questions of higher difficulty. The results show that 6% of the sample was OCF. If the student is on-task and has a series of slow responses

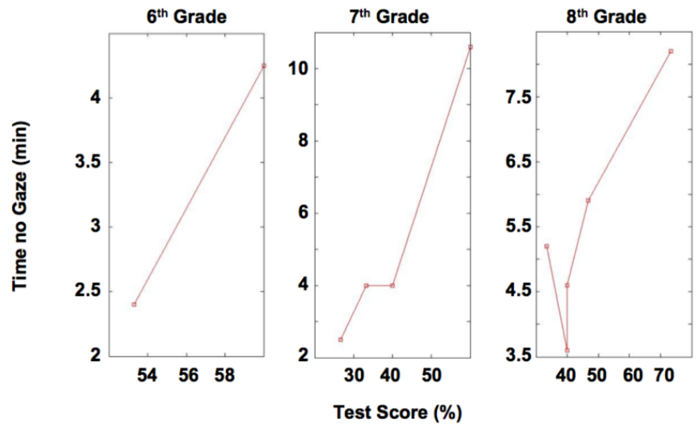


Figure 15: This graph shows the relationship between the students' test scores and the amount of time that eye gaze was not on the computer screen.

Table 6: Student's use of pencil & paper

Needed Pen & Paper?	Total	Percentage
Yes	109	56%
No	86	44%

with a series of correct answers (OCS), the student may understand the material and require more time to think. The results show that 7% of the sample was OCS. If the student is on-task and has a series of slow responses with a series of incorrect answers (O$\bar{C}$S), the student may lack understanding and need questions of lesser difficulty. The results show that 7% of the sample was O$\bar{C}$S. This additional information will be used in the future to better integrate instructional scaffolding and adaptation within a learning scenario in Chapter 6.

This work also suggests that eye gaze and head pose technique is not an effective measure of engagement when high-level cognitive thinking is required, which supports Hypothesis 2. Based on the video observations performed post-testing, the students consistently looked down at the paper to write out the multistep problems and calculate the answers by hand. The use of pencil and paper was further documented by the students in the exit survey (Table 6). We also observed that other students looked at random objects in space to perform mental math. In fact, we observed that the longer that the student looked away from the computer screen, the higher he or she performed on the test. The 8$^{\text{th}}$ grade student who scored the highest looked away from the screen for 8.2 minutes, which was 47.4% of the entire test time. The 7$^{\text{th}}$ grade student who scored the highest looked away from the screen for 10.6 minutes, which was 65.4% of the entire test time. The large time delay associated with the lack of eye contact from the human to the computer screen caused Asteriadis et al.'s eye gaze technique to incorrectly declare the students as being distracted or disengaged. However, using our engagement model, we were able to correctly categorize the students as being engaged. By monitoring the time delay/speed, accuracy of responses, and proper event execution associated with each question, we are able to expand the eye gaze model proposed by Asteriadis et al. and apply it in a complex problem solving environment [13].

3.2 Eye Gaze Input

Generally speaking, monitoring eye gaze has been used as a viable metric for measuring attention, and this is only one of many domains where eye tracking is beneficial. In Section 3.1, we observed that the eye gaze and head pose method of monitoring engagement as described in [9] worked well for low-level cognitive problems, but failed more often for the higher level problems. Because we want to target this system in the math domain, it is important that the engagement model is able monitor the student's attentiveness regardless of the problem difficulty. The preliminary engagement model takes time, response, and function execution into consideration, and it is able to successfully monitor engagement for all-level cognitive tasks. However, we hypothesize that we can create a richer and more adaptive learning environment by developing an eye gaze algorithm that is suitable for all-level cognitive tasks.

We began this exploration by creating a low-level cognitive environmental setup, such as viewing artwork (Figure 16). Future work will involve a high-level cognitive environmental setup, such as complex math problems. A camera is used to gather the necessary information such as gaze locations, fixations, and saccades while each user navigates through these environments. Once the data was collected, we applied various pattern recognition techniques to develop our eye gaze engagement algorithm. We expect to see trends that indicate future salient points in the environment, current areas of interest, level of cognitive load, and, ultimately, level of engagement.

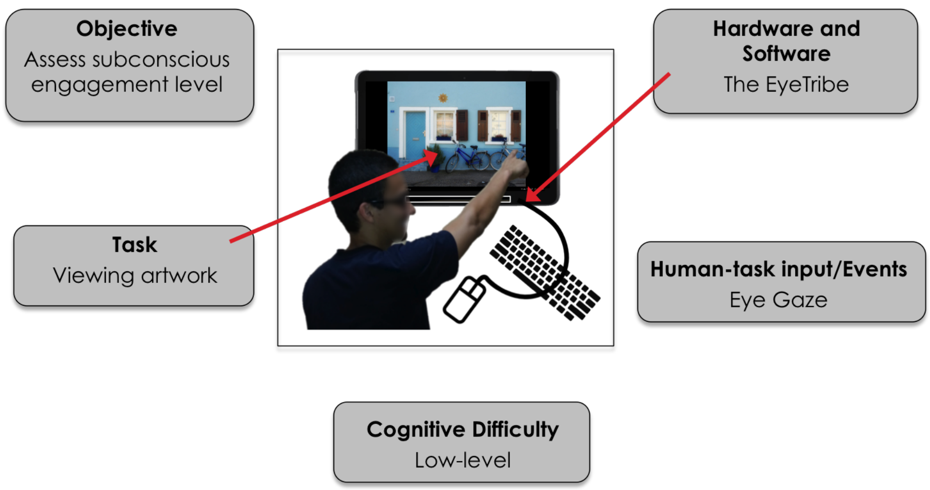


Figure 16: The entire eye gaze system – the Eye Tribe is placed below the monitor where the images are displayed to the user.

3.2.1 HumAnS Gaze Tracker

For this study, we have developed the HumAnS Gaze Tracker, which is a system designed to monitor a participants gaze while viewing various artwork displayed on a monitor. The Eye Tribe is the hardware used to track the eye gaze coordinates, pupil size, pupil centers, saccades, and fixations during the experience [3] (Figure 16). The administrator interacts with the graphical user interface (GUI) shown in Figure 17. Through use of this GUI, we are able to upload various image sets, input the participant's information, calibrate the Eye Tribe for each participant, display the images, and store the eye gaze data in a log file.

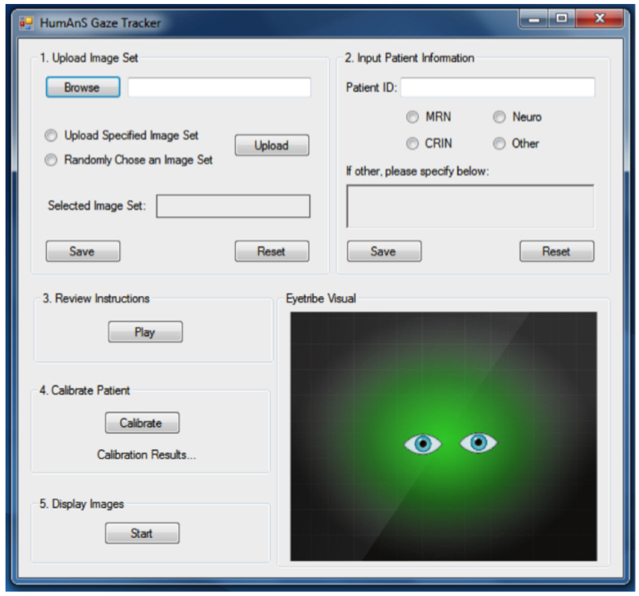


Figure 17: HumAnS Gaze Tracker GUI.

3.2.2 Hypothesis

To have a global impact with eye gaze devices, an accurate, low-cost system is needed for personal use in homes. As such, in this paper we focus on using a low-cost platform, the Eye Tribe, to develop software able to track gaze and further detect novelties in an environment. Individuals encounter novel environments on a daily basis – being able to predict salient points in an environment has the capability to enhance learning and understanding, increase engagement when completing a task, as well as provide insight on an individuals hippocampus-dependent memory. We hypothesize that by monitoring the amount of eye fixations along with the associated time durations, salient points in environments can be predicted before the individual becomes fully cognizant of them.

3.2.3 Experimental Design

To evaluate the correlation between eye gaze and subconscious awareness of salient points in an environment, we employed a single group design for this pilot study. A total of 9 participants took part in this experiment and all were recruited students from undergraduate and graduate studies at Georgia Institute of Technology in Atlanta, GA. The population consisted of both females and males in the age range of 20-30 years old (mean = 24.66, standard deviation = 2.78; Male: 5, Female: 4; undergraduate: 2, graduate: 7). During the study, the participant sat at a desk where a laptop and Eye Tribe were placed in front of him or her as shown in Figure 18.

Figure 18: The actual experimental setup.

To begin, the administrator uses the HumAnS Gaze Tracker GUI (Figure 17) to record the participants identification information and calibrate the Eye Tribe. We structure the design similar to a study conducted by Smith et al. [83]. Here, they used a combination of novel, repeated, and manipulated images to evaluate any patterns in eye movements. For the context of our study, we used 12 novel images and paired them with a slightly manipulated version of the image (Figure 19). The administrator used the GUI to upload a specific image set with 24 images in total (12 pairs). The system first displayed one image from all the pairs. Once one image from each pair has

Figure 19: An image pair used in the study. The image on the left has been manipulated, and the image on the right is the original.

been displayed, the system then displayed the second image of the pair in the same order. (We alternated between first displaying the original image and manipulated image to tease out the effects of photoshop in the manipulated image.) Each image is displayed for a total of ten seconds, and a black screen is shown between images for one second.

At the completion of each presentation, each participant recorded their perception of the images in an exit survey (Figure 20). For each image, we asked each participant if they noticed a change in the image and to give details about this change if applicable. When asked these questions, the last or most recent version of the image pair would be shown. The entire study was completed in approximately 10-15 minutes.

3.2.4 Results

Each participant viewed a total of 24 images; therefore, we were able to collect 216 instances of eye gaze data for all 9 participants. For each image, we monitored the time stamp, x- and y-coordinates of both left and right eyes, and pupil size. With this information we were able to calculate gaze duration, fixations, and saccades. The Eye Tribe was able to capture 30 frames per second, so we were able to collect

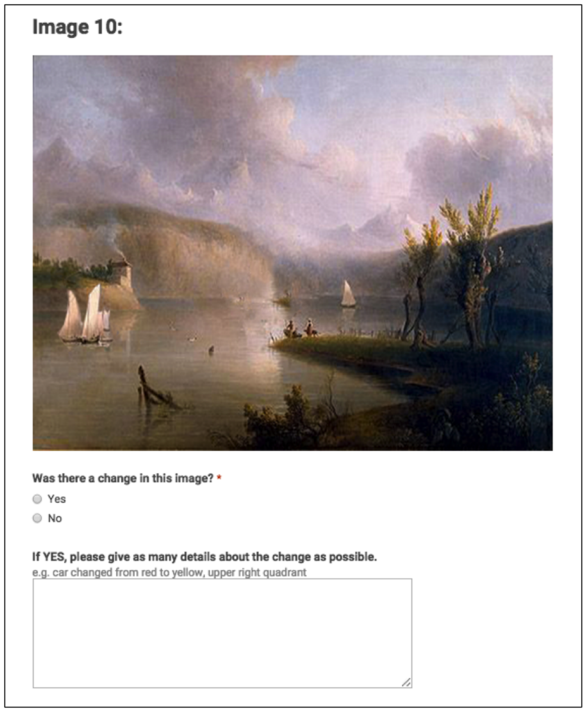


Figure 20: Exit Survey.

approximately 300 frames of gaze data for each image per participant. From there, the gaze coordinates were overlaid onto the images to have a visual of each users experience as shown in Figure 21.

The exit survey showed very interesting results. When we asked the participants if they were aware of changes in the image pairs, the participants did not notice a change for 46 of the 108 instances as shown in Table 7. For the 62 instances where the participants were able to notice a change, their responses are recorded in Tables 28 - 39 found in Appendix A. In general, if the participant noticed a change in the image, he or she was also able to recall the details of the change. A few participants were aware that the image changed, but were not able to note the correct change.

Because our hypothesis focuses on using eye gaze to observe if participants are subconsciously aware of salient points in environments, it is valuable to analyze the scan paths of the participants who noted that there was no change in the image pair

Figure 21: Participant 7 scanpath results for Image Pair 10. This participant noted on the survey that he or she was NOT aware that the image changed; however, it is shown here that there is fixation on the boat's new and original location in the right image.

(46 instances). Of these 46 instances, 10 instances show that the participant actually gazed at the changed object's new location, 3 gazed at the objects original location, and 29 gazed at both the new and original location. Only 4 participants completely overlooked the object of interest as shown in Table 8. In particular, Participant 7 noted that there was no change in Image Pair 10; however, the scan paths show that the area of interest was noticed and the participant spent most of the time fixated on the new and original location of the boat as shown in Figure 21. In the same breath, Participant 8 did not notice the new or original area of interest as shown in Figure 22.

3.2.5 Discussion

The preliminary engagement model used in Section 3.1 is able to use physical input from the student, such as mouse and keyboard events, to determine if the he or she is engaged during a math exam. The preliminary model uses the metrics of time, response/performance, and proper function execution when evaluating engagement

Table 7: Total number of yes/no responses from the participants when asked if they noticed a change in each image pair.

Image Pair	"No" Response	"Yes" Response
1	3	6
2	3	6
3	0	9
4	0	9
5	7	2
6	2	7
7	5	4
8	6	3
9	1	8
10	4	5
11	9	0
12	6	3
Total:	**46**	**62**

levels. However, by integrating an eye gaze metric into the model, a richer learning environment can be created by monitoring where the student is subconsciously directing his or her attention on the teaching device (computer/tablet).

With subconscious eye gaze as a metric, learning and understanding can be enhanced by integrating scaffolding methods in the system based on gaze location. Through gaze, the student is passively articulating areas of interest/salient points in their learning environment. By providing scaffolding based on these salient points, the student's engagement levels are further increased because the student has already deemed this an area of interest. For instance, during a multiple choice math test, if a large amount of a student gaze's is directed towards the incorrect answer response, an intelligent tutoring system can interject with information that would assist the student with rethinking their initial analysis. This idea is further displayed in Figure 23.

Figure 22: Participant 8 scanpath results for Image Pair 1. This participant noted on the survey that he or she was NOT aware that the image changed, and the gaze data displayed here further corroborates his or her assessment.

3.2.6 Conclusion

Of the 46 instances when the participants were not consciously aware of any changes in the image pairs, 91% of their eye gaze data showed that they were subconsciously aware of the novel changes in their environment (Table 8, Figure 21). This alone proves our hypothesis to be true, in that we are able to predict salient points in an environment via eye gaze before an individual becomes fully cognizant of them. Furthermore, monitoring eye gaze in a novel environment has the potential to enhance learning and understanding as well as engagement when completing a task.

Table 8: Of the 46 instances that the participants did NOT notice a change, 42 of them subconsciously gazed at the new and/or original area of interest. Only 4 participants completely overlooked the areas of interest.

Image Pair	New	Original	Both	Neither	Total
1	2	0	0	1	3
2	3	0	0	0	3
3	0	0	0	0	0
4	0	0	0	0	0
5	0	1	5	1	7
6	1	0	1	0	2
7	3	0	1	1	5
8	0	2	4	0	6
9	1	0	0	0	1
10	0	0	4	0	4
11	0	0	8	1	9
12	0	0	6	0	6
Total:	**10**	**3**	**29**	**4**	**46**

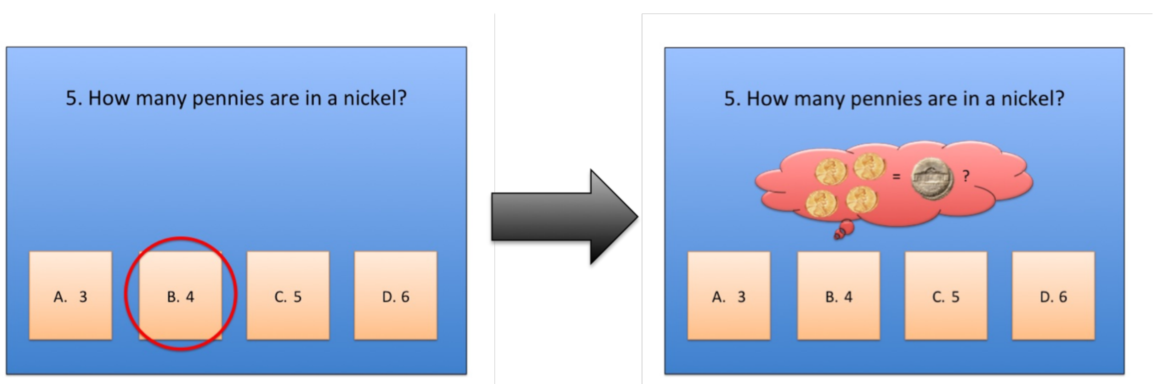


Figure 23: On the left, the student's gaze is directed towards the incorrect response. On the right, an intelligent tutoring system is able to adapt and interject additional information to assist the student before the he or she decides to select and submit this response.

CHAPTER IV

SOCIAL INTERACTION MODEL

The role of emotions in social scenarios is to provide an inherent mode of communication between two parties. When emotions are properly employed and understood, people are able to respond appropriately, which further enhances the social interaction. Ultimately, effective emotion execution in social settings has the capability to build rapport, improve engagement, optimize learning, provide comfort, and increase overall likability. In this chapter, we discuss associating dominant emotions of effective social interaction to verbal and nonverbal behaviors on a humanoid robotic platform. The chapter addresses our second contribution:

> *2. Develop a framework for creating re-engagement strategies, verbal and nonverbal cues, on a humanoid robotic platform that aid in enhancing learning.*

4.1 Nonverbal Behavioral Strategies

Because empathy is a key factor used to enhance interpersonal relationships, which ultimately leads to increased enthusiasm and learning [86], we derived a framework for implementing happy and sad emotions on a humanoid robotic platform [17]. Anger can be detrimental to building rapport and establishing a level of comfort in social settings, so this emotion was not investigated. The major areas of interest when developing gestural behaviors are the head movements [12, 56, 80], arm movements [56, 80, 26, 33], and the overall size [80, 26, 33] and speed [26, 33] of the gesture.

4.1.1 Gestural Behavior Framework

Based on prior studies, it is noted that moving the head in an upward position should convey happiness, while moving the head in a downward position should convey sadness [12]. Moving the arms in an upward position should convey happiness, while moving the arms slowly up and down should convey sadness [26]. As such, the framework that has been developed and validated in this investigation to create empathetic gestures on a humanoid robotic platform is displayed in Table 9.

Table 9: Framework for gestural behavior implementation

Key Principle	**Happy Characteristics, *HC***	**Sad Characteristics, *SC***
Head Direction	Upward	Downward
Arm Direction/Movement	Upward	Slow up & down
Gesture Size, ***S***	Large	Small
Gesture Speed, ***P***	Fast	Slow

For this investigation, size $\boldsymbol{S}$ of the gesture is determined by the number of body parts in motion coupled with the range of motion of the movement as shown in (7) - (9).

$$S_{large} = AB \tag{7}$$

$$S_{medium} = A\overline{B} \; or \; S_{medium} = \overline{A}B \tag{8}$$

$$S_{small} = \overline{A}\;\overline{B}, \tag{9}$$

where $\boldsymbol{A}$ is the number of active servos/joints and $\boldsymbol{B}$ is the range of motion. Based on this definition of "size," a large gesture should convey happiness, while a small gesture should convey sadness [26]. The speed $\boldsymbol{P}$ of the gesture is determined by the rate of change in the movement (not the total length of time), as shown in (10).

$$P = B/t, \tag{10}$$

where $\boldsymbol{t}$ is time. Based on this definition of "speed," a fast gesture should convey happiness, while a slow gesture should convey sadness. (Note: The actual high/low thresholds associated with speed and size of the gestures were determined through empirical studies.)

By using this framework, gestures that effectively convey empathy, happy and/or sad emotion, can be developed with ease. These gestures are developed by comparing the sum of happy characteristics, $\boldsymbol{HC}$, to the sum of sad characteristics $\boldsymbol{SC}$. More specifically, a happy gesture is the result when

$$\sum HC > \sum SC. \tag{11}$$

A neutral gesture is the result when

$$\sum HC = \sum SC. \tag{12}$$

A sad gesture is the result when

$$\sum SC > \sum HC. \tag{13}$$

4.1.2 Gesture Implementation on Robotic Platform

For the robotic social agent, we utilize the DARwIn-OP platform (Darwin) [45], a humanoid robot with 20 actuators, resulting in 6 DOF for each leg, 3 DOF for each arm, and 2 DOF for the neck (Figure 24). To enable interaction with the human, Darwin was programmed with 15 gestural behaviors created by applying the gestural framework in Table 9. Of the 15 gestures created, 8 were happy, 3 were neutral, and 4 were sad gestures. A brief description of each gesture is given in Table 10. These gestures were programmed using Darwin's default program ActionEditor in which we programmed a sequential set of actuator positions, with speed and timing constraints, to affect an appropriate gesture. Figure 24 - 26 displays an example of a happy, neutral, and sad gestural behavior, respectively.

The 15 gestures are further broken down into the key principles of the framework in Table 11. By doing this, it is evident how each gesture is associated with specific

Table 10: Description of gestural behaviors from the robotic agent

Gesture	**Description**
H1	Robot looks upward while raising his arms in the air and bringing them together, as if he were clapping.
H2	Robot moves his head only in an up and down motion while raising both arms to form a 90°angle with the ground.
H3	Robot raises his left arm in a 90°angle, then pulls it down at a rapid speed.
H4	Robot bends his knees, then straightens his legs while raising both arms. Darwin is moving his head up and down.
H5	Same description as H4, however, there are very subtle differences.
H6	Robot bends his knees, then straightens his legs while raising both arms and looking upward.
H7	Robot simply moves his head only in an up and down motion.
H8	Robot raises both of his arms simultaneously and forms a 90°angle while his head is upward ("field goal" sign).
N1	Robot bends his knees then straightens his legs repeatedly while moving his head up and down simultaneously.
N2	Robot raises his left arm towards his head. He then moves his arm up and down next to his face (scratches head).
N3	Robot nods his head up and down while simultaneously moving his arms back in forth (engaging in conversation).
S1	Robot lowers his head to the ground and then raises his hands to his head, as if they were holding his head
S2	Robot simply lowers his head to the ground.
S3	Robot lowers his head and raises his hands to his head (holds head). He then slowly shakes his head from side to side.
S4	Robot lowers his head to the ground and then slowly shakes his head from side to side.

characteristics for depicting emotion. In particular we highlight the position of the head [12], the direction/movement of the arms [26], the movement of the legs [80], and the overall size and speed of the gesture [26]. We also implement smooth, fluent, and regular movements for both the happy and sad gestures [26]. In addition, the purpose of highlighting the movements of each body part is to observe if dominant involvement (lower body) yields any significant differences when compared to subordinate involvement (upper body).

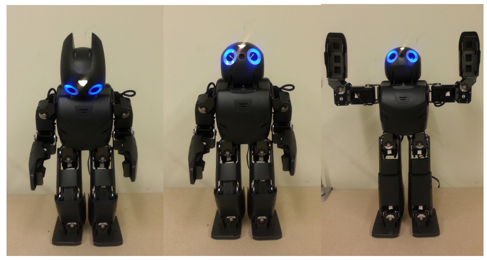


Figure 24: A *happy* gesture (H4) is broken down into three parts. Refer to Table 10 for each gesture's description.

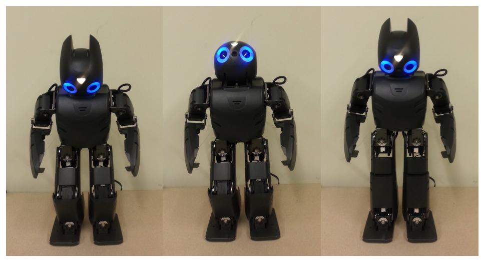


Figure 25: A *neutral* gesture (N1) is broken down into three parts. Refer to Table 10 for each gesture's description.

4.1.3 Hypothesis

Previous studies have shown that empathy is a key factor used to enhance interpersonal relationships, which ultimately leads to increased enthusiasm and learning [86]. Therefore, we have derived a framework for implementing happy and sad emotions on a humanoid robotic platform. Our hypothesis states that by applying the framework outlined in Table 9, individuals will be able to perceive the correct resulting emotion (happy, neutral, or sad) implemented on a humanoid robotic platform.

When the null hypothesis is accepted, the predicted resulting emotion and the actual resulting emotion will not be equivalent and/or the sensitivity of the resulting emotion will be less than 75%. When the null hypothesis is rejected, the predicted resulting emotion and the actual resulting emotion will be equivalent and the sensitivity of the resulting emotion will be greater than 75%. Sensitivity is the true positive

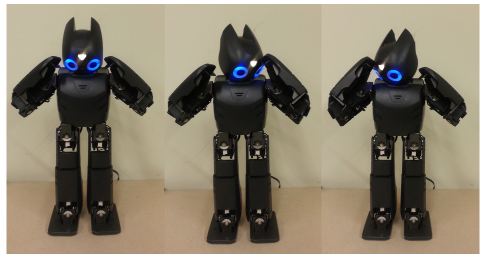


Figure 26: A *sad* gesture (S4) is broken down into three parts. Refer to Table 10 for each gesture's description.

rate (TPR) and (14) will be used to test the hypothesis.

$$TPR = TP/P = TP/(TP + FN), \tag{14}$$

where P is the number of positive instances, TP is the number of true positives, and FN is the number of false negatives.

4.1.4 Experimental Design

To evaluate the perception of gestural behaviors implemented on the robotic social agent, we employed a single group design for this study. A total of 13 participants took part in this experiment and all were recruited students from undergraduate and graduate studies at Georgia Tech in Atlanta, GA. The population consisted of both females and males in the age range of 18-34 years old (m = 25.8, SD = 3.9; Male: 8, Female: 5; undergraduate: 1, graduate: 12). During the study, the participant sat at a desk where Darwin stood 2 feet away as shown in Figure 27. Once the participant was positioned, Darwin performed a gesture (Table 5), and then returned to a standing rest position. The gestures were selected at random to ensure that order of the gestures presented did not have an effect on perception. If the participant did not see a gesture fully or asked to view it again, Darwin was tasked to perform it again until the participant was ready to move forward to the next gesture. At the completion of each gesture, the participant recorded their perception of Darwin's behavior on a

Table 11: Key principles and associated emotion

Gesture	Head	Arms	Legs	Size	Speed	$\sum$ HC	$\sum$ SC	Resulting Emotion
H1	Up	Up	–	Medium	Fast	3	0	Happy
H2	–	Up	–	Medium	Fast	2	0	Happy
H3	–	Up/Down	–	Medium	Fast	1	0	Happy
H4	–	Up	Bend	Large	Fast	3	0	Happy
H5	–	Up	Bend	Large	Fast	3	0	Happy
H6	Up	Up	Bend	Large	Fast	4	0	Happy
H7	–	–	–	Small	Moderate	1	0	Happy
H8	Up	Up	–	Medium	Fast	3	0	Happy
N1	–	–	Bend	Medium	Moderate	0	0	Neutral
N2	–	Up/Down	–	Medium	Moderate	0	0	Neutral
N3	–	Midway	–	Medium	Moderate	0	0	Neutral
S1	Down	Midway	–	Medium	Slow	0	2	Sad
S2	Down	–	–	Small	Slow	0	3	Sad
S3	Down	Midway	–	Large	Slow	1	2	Sad
S4	Down	–	–	Medium	Slow	0	2	Sad

5-point Likert scale (very happy - very sad) (Figure 28). This is repeated until all 15 gestures had been performed by Darwin and evaluated by the participant. The study was completed in 10 minutes.

4.1.5 Results

To prove or disprove the hypothesis that the perception of emotion implemented on a robotic social agent can be determined by the key principles outlined in Table 11, we analyze the results of the Likert scale and confusion matrix. First we look at the results of a 5-point Likert scale, where 1 is “Very Happy” and 5 is “Very Sad.” These results are shown in Figure 29. There were a total of 104 happy, 39 neutral, and 52 sad gesture instances. Because the data set is unbalanced, we evaluate the sensitivity and the specificity of each emotion in Table 12.

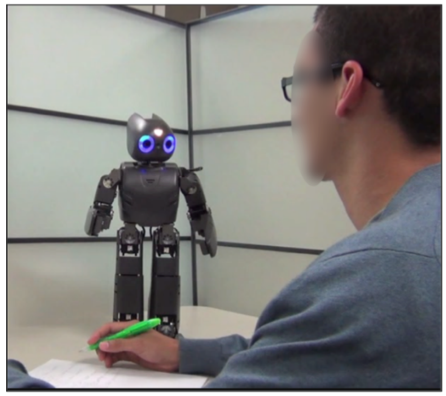


Figure 27: The experimental setup.

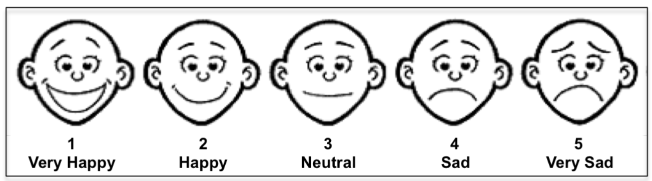
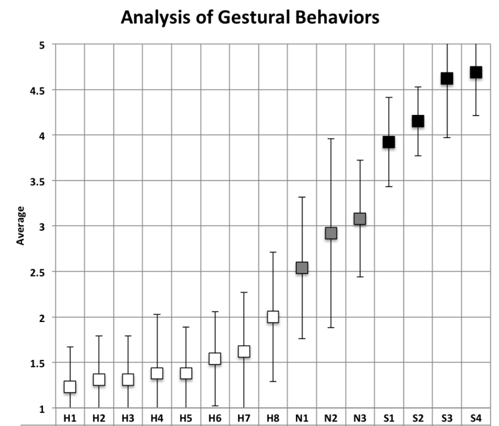


Figure 28: The scale that the participants used to rank each gesture's perceived emotion.

4.1.6 Discussion

The standard deviations in Figure 29 were all less than 0.8 excluding one, and it was a common trend that the "neutral" emotions had higher standard deviations than the "happy" and "sad" emotions. An explanation for this is that the lack of dominant characteristics in the gesture caused confusion for the participants. Table 11 shows that N1, N2, and N3 all have no sad principles and no happy principles, so it is not a surprise that participants were confused with these gestures. Even during the actual testing, these gestures where oftentimes asked to be repeated for clarification. This result suggests that there must be a distinguishable amount of happy and sad principles for accurate perception of gestures.

Similarly, Table 12 clearly shows that the participants were not confident in predicting when the intended emotion was neutral. There were 16 instances of false



	H1	H2	H3	H4	H5	H6	H7	H8	N1	N2	N3	S1	S2	S3	S4
Mean	1.23	1.31	1.31	1.38	1.38	1.54	1.62	2	2.54	2.92	3.08	3.92	4.15	4.62	4.69
SD	0.44	0.48	0.48	0.65	0.51	0.52	0.65	0.71	0.78	1.04	0.64	0.49	0.38	0.65	0.48

Figure 29: Each gesture's average perceived emotion is shown. The upper and lower error bars are equivalent to one standard deviation.

Table 12: Analysis of sensitivity & specificity (A = actual, P = predicted)

		Positive (P)	**Negative (P)**	**Sensitivity/ Specificity**
Happy	**Positive (A)**	99	5	95.19% (TPR)
	Negative (A)	10	81	89.01% (TNR)
Neutral	**Positive (A)**	3	16	58.97% (TPR)
	Negative (A)	8	148	94.87% (TNR)
Sad	**Positive (A)**	49	3	94.23% (TPR)
	Negative (A)	6	37	95.80% (TNR)

negatives and 23 instances of true positives, which resulted in a sensitivity of only 58.97%. Because the TPR is less than 75% for the neutral intended emotion, we are not able to reject the null hypothesis. However, the participants were very confident predicting when the intended emotion was not neutral. There were 8 instances of false positives and 148 instances of true negatives, which resulted in a specificity of 94.87%.

For both the happy and sad intended emotions, the participants were confident predicting both when the emotion was and was not present. For happy, there were

5 instances of false negatives and 99 instances of true positives, which resulted in a sensitivity of 95.19%. This high TPR for the happy intended emotion allows us to reject the null hypothesis. There were 10 instances of false positives and 81 instances of true negatives, which resulted in a specificity of 89.01%. For sad, there were 3 instances of false negatives and 49 instances of true positives, which resulted in a sensitivity of 94.23%. This high TPR for the sad intended emotion allows us to reject the null hypothesis. There were 6 instances of false positives and 137 instances of true negatives, which resulted in a specificity of 95.80%.

Figure 29 illustrates that the participants were able to distinguish seven gestures as extreme instances. H1, H2, H3, H4, and H5 were on average "very happy," whereas S3 and S4 were on average "very sad." In addition, the gesture with the smallest standard deviation of 0.376 was S2 with an average of 4.154 (sad). Once the range of happiness is combined into one category and the range of sadness is combined into one category, the participants completely agree on their perception of the gestures. In particular, H1, H2, H3, H5, H6, S2, and S4 have no deviation across all participants (SD = 0). This suggests that implementing these gestures into a social scenario would be ideal to enhance engagement and motivation.

Lastly, the movement of the upper body versus the lower body as discussed in [80] did not reveal any trends necessary for distinguishing extreme emotion (very happy/very sad). All 15 of the gestures had some type of upper body movement, but 4 of the gestures incorporated lower body movement as well. Of the 4 gestures that incorporated lower body movement, 2 were classified as an extreme emotion (50%). However, of the 11 gestures that did not incorporate lower body movement, 5 were still classified as an extreme emotion (45.45%), while 6 were not classified as extreme emotion (54.54%).

4.1.7 Conclusion

This study revealed that by altering head direction, arm direction, gesture size, and gesture speed on a humanoid robotic social agent, participants are able to achieve accurate perception when the intended emotion is happy or sad. By using these key principles to categorize the gestures, the standard deviation was kept consistently at a minimum when identifying emotion. In fact, seven of the gestures yielded no standard deviation across all the participants. When using this framework, the participants are very confident in identifying when the intended emotion is happy, not happy, sad, not sad, and not neutral. However, participants are not confident identifying when the intended emotion is neutral. This work suggests that engagement and motivation during social interaction can be optimized through the use of happy and sad gestures derived using the described framework.

4.2 Verbal Behavioral Strategies

Studies have shown that the use of verbal encouragement strategies in education is able to maximize learning. This idea is derived from traditional classroom settings where teachers use a multitude of behavioral strategies to maintain the students level of engagement. Motivated by these educational practices, we developed a number of socially-supportive phrases to embed on the robotic educational agent.

4.2.1 Verbal Behavior Implementation

Open dialogue integrating socially supportive phrases between teacher and student is ideal for optimal learning [78]. The use of verbal cues has the ability to encourage the student, provide instruction, and give positive praise. To support this theory, we have embedded verbal behaviors that enable the educational agent to provide socially supportive phrases for re-engagement as the student navigates through the learning task. During the utterance of verbal phrases, the robotic platform turns its gaze towards the student; otherwise, the robot remains looking at the teaching device. The goal of the verbal phrases is to encourage the student based on their current learning performance. It is very important that the phrases are socially supportive and convey the idea of teamwork. There is a dialogue established between the student and the robot, and not a unidirectional knowledge flow (i.e. the robot is not giving instructions or issuing commands to the student). A sample of these socially supportive phrases is shown in Table 13. For implementation purposes, the phrases were recorded and stored on Darwin's external SD card as mp3 files. Validation of the verbal behaviors employed on the robotic educational agent is discussed in Chapter 5.

Table 13: Sample of verbal responses

Answer	**Speed**	**Phrase**
Correct	Fast	"You really know your stuff!"
		"You're a genius!"
		"Fantastic!'
	Slow	"This is hard, but we're doing great."
		"Thanks for all your hard work."
		"This is really making me think."
Incorrect	Fast	"Wait, I didn't get to read that one."
		"Hang in there. We're almost done."
		"I'm lost. We're going too fast.'
	Slow	"Don't worry. I had trouble with that one too."
		"That one was very challenging."
		"Don't sweat it. We'll get the next one."
None	Inactive	"Let's make an educated guess."
		"I was completely stumped on this one."
		"Don't forget about me over here."

CHAPTER V

ENGAGEMENT WITH ROBOTIC EDUCATIONAL AGENTS

Our studies have shown that teaching processes, which incorporate robotic-based *engagement* methods, can approach the effectiveness of human instructors [20, 22, 23, 38]. We discuss the overall system approach, which consists of the engagement model (Chapter 3) and the forms of multi-modal nonverbal and verbal cues used by the robotic agent (Chapter 4). The result of this study fulfills our third contribution:

> *3. Develop a system that uses the physical input-based engagement model and baseline re-engagement strategies embedded on a robotic platform to re-engage students during a learning task.*

5.1 Computation Tasks

In this section we elaborate on the process of embedding social interaction within a humanoid-student *math*-learning scenario in order to re-engage students during high-demand cognitive tasks (Figure 30). Results derived from 44 students engaging with a robotic educational agent during a tablet-based math exam show that, while various forms of social interaction increase test performance, combinations of verbal cues result in a slightly better outcome with respect to test completion time.

5.1.1 The Learning Environment

In traditional learning scenarios, active engagement is an important goal for both students and teachers [4]. One of the most non-engaging, yet necessary, elements of the current learning environment is the process of testing [62]. As such, in this

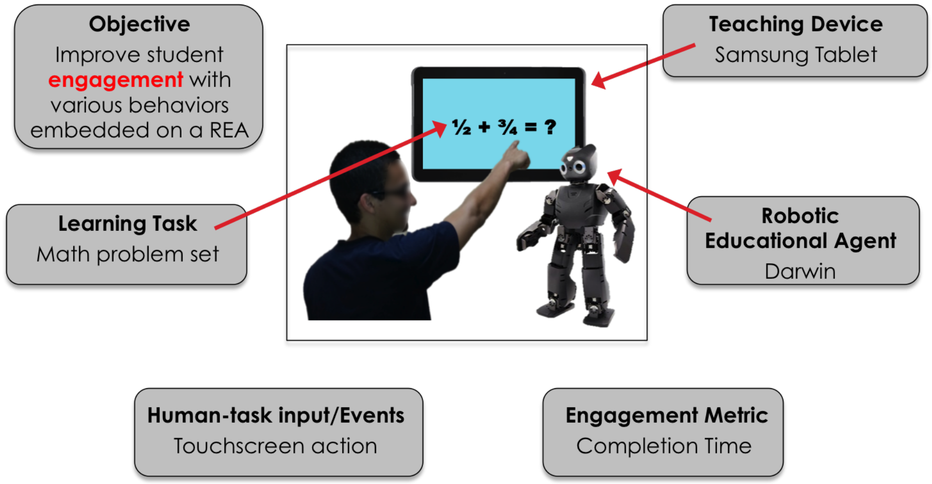


Figure 30: Diagram of the entire system with the REA.

chapter, we focus on the math-testing scenario to evaluate the role and effectiveness of engagement using a robotic agent. For our work, we employed a 15-question multiple-choice algebra or calculus test, which was proctored using a Samsung Galaxy Tablet (Figure 31). There were three distinct display screen layouts throughout the test: the welcome screen, the multiple-choice test screen(s), and the completion screen. The first screen, the welcome screen, introduces the student to the system and enables initiation of the test. The test, itself, is composed of a sequential set of screens that highlights a single question, with an associated image when applicable, and a set of multiple-choice answers with button choices A, B, C, and D (Figure 31). The application only allows the student to make one selection, and then he/she will press the Next button located at the bottom of the screen to move forward to the next test screen. The application does not allow students to navigate backwards during the test. Once the student reaches the last test screen, a 'Submit Test' button replaces the 'Next' button. Once pressed, the completion screen is displayed, and the test has been completed.

As each student progresses through the test, their interactions with the tablet are communicated to the robotic educational agent Darwin [45] via Bluetooth. To enable

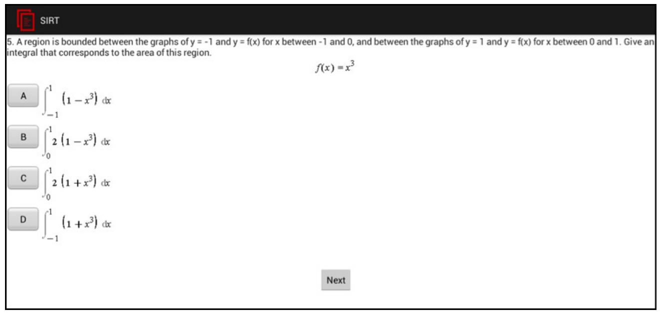


Figure 31: The Learning Environment - Calculus Test Question Screen.

real-time performance, only the numbers 0-9 are transmitted from the tablet to the Darwin. Each number conveys a different message to Darwin about the interaction between the student and the tablet. Basically, every button that is pressed is sent to Darwin, as well as the time intervals taken to navigate through the test. Table 14 defines what each number represents to Darwin.

Table 14: Bluetooth Communication Protocol between the Tablet and Darwin

Message Sent from Tablet	**Time (s)**	**Button Pressed on Tablet**	**Answer**
0	n/a	Start App Icon	n/a
1	n/a	A, B, C, or D	n/a
2	t<30	Next	Correct
3	t>90	Next	Correct
4	30<t<90	Next	Correct
5	t<30	Next	Incorrect
6	t>90	Next	Incorrect
7	30<t<90	Next	Incorrect
8	t =90n, n>0	n/a	n/a
9	n/a	Submit Test	n/a

Upon opening the tablet-based math test, 0 is sent to Darwin and he then begins his introduction on the welcome screen. If a multiple-choice answer is selected (A, B, C, or D), a 1 is sent to Darwin and he will respond appropriately based on the engagement type (verbal, nonverbal, or both). An answer is classified as either being *fast*, *slow*, or *average* based on the time elapsed on each question: if the student submits a response in less than 30 seconds this is fast; if the student submits a response in between 30 and 90 seconds this is average; if the student submits a response in more than 90 seconds this is slow. (The thresholds for these time intervals are based on the results from pilot testing as discussed in Chapter 3.) The answers are also classified based on whether or not the answer is correct. Messages 2-7 are the numbers sent to Darwin based on the answers submitted on the tablet.

To improve human-robot team performance, Shah et al. were able to reduce a subject's idle time by monitoring the beginning and end of tasks [82]. Based on the results from this study, we focused on decreasing idle time by monitoring task or question duration. Therefore, when there are long time intervals where there is no interaction between the human and the tablet, '8' is sent to Darwin. A long time interval is defined as 90 seconds; therefore, every 90 seconds of inactivity or idle time, Darwin is notified and he will respond appropriately. Lastly, 9 is sent to Darwin at the completion of the test.

Depending on user-state, Darwin provides the users cues that are either verbal, nonverbal, or a combination of the two (depending on the experimental group). For both verbal and nonverbal behaviors, the behavior was selected at random based on the message sent to Darwin from the tablet. For the engagement type that incorporates both verbal and nonverbal cues, the gestures and phrases were scripted and paired prior to Darwin's random selection. As such, we were able to expand Darwin's library of verbal and nonverbal cues by pairing the same phrase with multiple gestures. Although a phrase when it stands alone may mean one thing, by adding a gesture, the underline meaning of the message can be altered. Upon execution of a pair, both the gesture and the phrase are performed simultaneously. For example, if 3 (Slow correct answer submitted) is sent to Darwin, he may say, "This is hard, but we're doing great," while nodding his head.

5.1.2 Hypothesis

In this study, we look to validate the hypothesis that the use of a robotic educational agent can increase test performance by adaptively engaging with the student. Adaptive engagement is based on the concept that the engagement model is driven by identification of the behavioral state of the student.

5.1.3 Experimental Design

To evaluate the effectiveness of the robotic educational agent engaging students during the learning process, we employed a between-groups design for this study. To guarantee that the skills are evenly distributed between the groups, the students were selected at random. A total of 24 college students took part in Trial 1 of the experiment; this consisted of both females and males in the age range of 18-33 years old (m = 24.6, SD = 4.9, Male: 18, Female: 6). A total of 20 high school students took part in Trial 2 of the experiment; this consisted of both females and males in the age range of 15-16 years old (m = 15.5, SD = 0.51, Male: 12, Female: 8). Our experiment involved one factor – type of re-engagement:

- **No Agent -** Represents the control group. No agent is present.
- **Verbal -** The agent will say socially supportive phrases for re-engagement.
- **Nonverbal -** The agent will use only gestures for re-engagement.
- **Mixture of Both -** The agent will use both gestures and phrases for re-engagement.

The experimental setup (Figure 32) in this study involves a test-taking learning scenario. A Samsung Galaxy Tablet is used as the teaching device for displaying questions and recording the students answers. The tablet is placed on an adjustable stand at eye level. We utilize the humanoid robot Darwin as the platform for our robotic educational agent [45]. For experiments with the robot agent present, Darwin is positioned to the right of the tablet, yet between the tablet and the student. The robot is placed in a position such that the robot is always able to see and interact with both the tablet and the student.

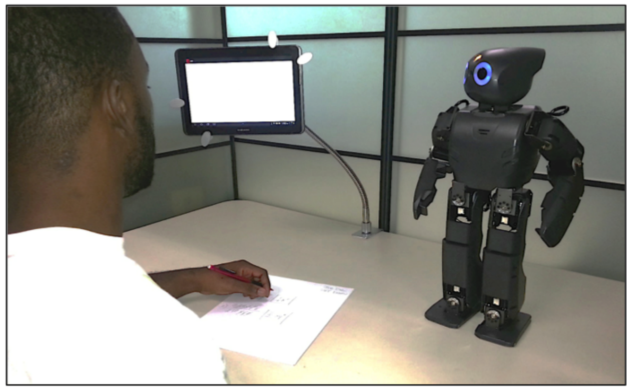


Figure 32: The experimental setup.

For experiments with the robot present, at the start of the test-taking learning scenario, Darwin gives a verbal introduction and discusses the activity that the students are about to perform. The script of this verbal introduction is shown below:

> *"Hello. My name is Darwin. We will be going through a series of 10 math questions to learn the material together. I appreciate you taking the time out of your busy schedule to work with me. Get your pencil and paper ready so we can start. Press begin when you're ready."*

The purpose of this introduction is to eliminate the novelty of the robot from the investigation and prepare the students for the test by instructing them to gather their materials. The students then navigate through the test questions until they reach the completion screen. The test questions consists of multiple-choice math questions of varying difficulty levels. As each student progresses through the test, the interaction with the learning device is communicated to Darwin. Every button that is pressed is sent to the robot, as well as the time intervals taken to navigate through the test (Table 14). After test completion, a message is also sent to the robot, at which point Darwin shows its gratitude and gives a farewell.

We focus on increasing engagement while decreasing idle time by monitoring task or question duration with the engagement model discussed in Chapter 3 [18, 16]. Depending on the estimated user-state determined by this model, Darwin provides the students cues that are either verbal, nonverbal, or a combination of the two. For both verbal and nonverbal behaviors, the behavior was selected at random based on the message sent to Darwin from the teaching device. We were able to expand Darwin's library of verbal and nonverbal cues by pairing the same phrase with multiple gestures. A stand-alone phrase can have one meaning, but by adding a gesture, the underline meaning of the message can be altered. Upon execution of a pair, both the gesture and the phrase are performed simultaneously.

For the experimental design, we utilize the same test, environmental setup, and engagement model across all students. The only change between groups is the type of cues that Darwin provides. For the control group, Darwin is removed from the table.

5.1.4 Results

To prove or disprove the hypothesis, we examined test completion time, the Likert scale questions that were asked in an exit survey, and the comments that students left at the end of the survey. We logged the total test time for each student in each trial, and the means for the four groups are shown in Figure 33. The results of Trial 1 and 2 are compared and contrasted in Figure 34(a), and the results averaged together from both trials are shown in Figure 34(b). The statistical analysis of each group and trial is shown in Table 15.

After the students completed the test, we asked them to rate their agreement with a series of statements on a 5-point Likert scale that ranged from "Disagree" to "Agree". For each of the questions on our survey, we performed an ANOVA test to

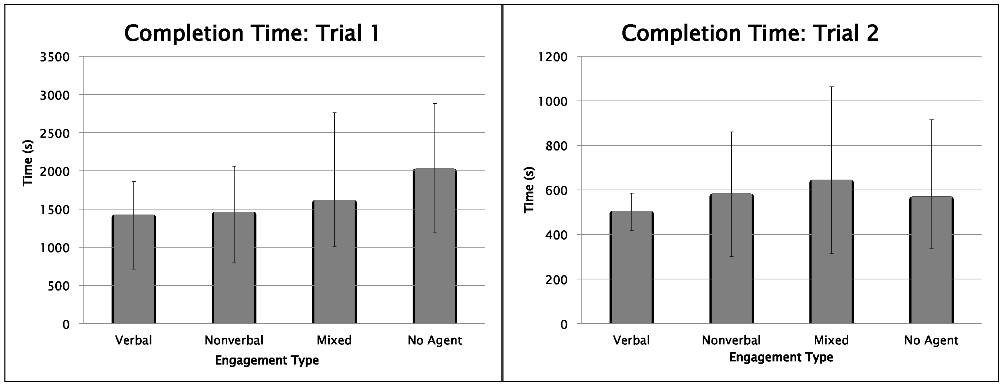
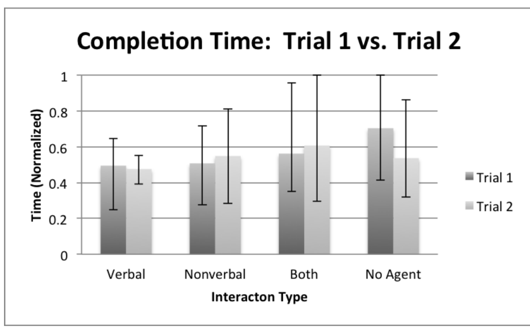
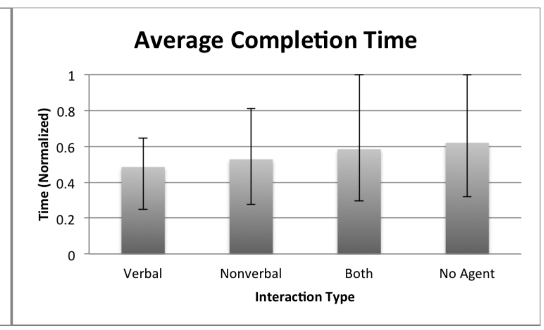


Figure 33: The average completion times and ranges for Trial 1 & 2.



(a) Trial 1 vs. Trial 2.



(b) Trial 1 & 2 combined.

Figure 34: Normalized results for Trial 1 & 2.

see if the differences between groups were significantly different. Table 16 depicts the question, the average response, and the p-values, which are separated by trial and test groups. In addition to the survey questions, we left room on the survey for students to provide freeform comments that reflected their overall experience. These are highlighted in the following section.

5.1.5 Discussion

For both trials, the verbal group was able to decrease idle time and maintain the students attention best with the lowest average test times. In Trial 1, when compared to the control group (No Agent), the verbal groups average test time was 30% lower, while in Trial 2, the verbal group was 11% lower than the control group (Figure

Table 15: Total time(s) statistical analysis

	Group	**Mean**	**SD**	**p-value**
Trial 1	*Verbal*	*1424*	*405*	0.28
	Nonverbal	1461	533	
	Both	1618	673	
	No Agent	2026	673	
Trial 2	*Verbal*	*505*	*72*	0.80
	Nonverbal	583	240	
	Both	645	290	
	No Agent	570	230	

34(a)). Across both trials, the verbal groups average test time was on average 22% lower than the control group (Figure 34(b)). The verbal group also presented the lowest standard deviations. In Trial 1, when compared to the control group, the verbal groups standard deviation was 40% lower, while in Trial 2, the verbal group was 69% lower than the control group. Across both trials, the verbal groups standard deviation was on average 51% lower than the control group (Figure 34(b)). This not only shows that the verbal cues were able to decrease time, but they were also able to do so uniformly throughout the groups. This small range, test time, and standard deviation values make it easier to guarantee a lower test completion time.

In Trial 1, there was a significant variance in how nervous the students deemed themselves to be during the test with and without Darwin. The control group was the least nervous during the test with a score of 1.00 (Disagree = 1; SD = 0), while the remaining groups with Darwin had an average score of 2.72 (Neutral = 3; SD = 1.4). This may be attributed to the students' fear of letting Darwin down during the test (Slightly Agree = 4; Avg = 3.9; SD = 1.2). This fear in effect makes the students nervous, which is only natural. In addition, the fact that the students have fear of disappointing Darwin supports that idea that a personal relationship was built between Darwin and the human.

In Trial 2, there was a statistically significant variance in how appropriate the

Table 16: Statistical analysis of survey responses (Trial 1/Trial 2)

Question	Verbal	Nonverbal	Both	No Agent	p-value
I found this test difficult	4.00	4.67	4.67	3.83	0.25
	2.00	2.00	1.60	1.20	0.43
I performed better than I had anticipated	3.00	2.67	3.33	3.17	0.74
	2.80	2.40	2.40	2.80	0.92
I was nervous during this test	2.50	3.33	2.33	1.00	0.03*
	2.20	3.00	2.00	2.00	0.51
I finished this test quickly	3.00	1.50	2.67	2.67	0.08
	3.40	3.20	3.40	2.80	0.79
I was frequently bored during this test	2.00	2.83	1.50	2.33	0.25
	1.80	3.40	1.80	4.60	0.002*
This test was an appropriate level for my skills	3.33	2.50	3.00	4.33	0.13
	2.60	2.60	4.00	1.80	0.19
I enjoyed taking this test	3.00	2.83	3.33	4.17	0.31
	4.00	3.20	4.40	2.20	0.07
I performed better on the test with Darwin	3.33	2.00	3.00		0.14
	2.80	2.00	3.00		0.19
Darwin distracted me during the test	2.17	2.50	2.50		0.90
	2.40	2.80	1.80		0.53
I was comfortable with Darwin's presence	3.50	3.33	4.33		0.27
	3.20	4.20	4.40		0.14
Darwin made me work quicker than usual	3.50	2.83	3.83		0.25
	2.40	2.40	3.80		0.14
Darwins feedback was helpful	3.33	3.00	3.50		0.74
	2.80	2.40	4.20		0.14
I was afraid of letting Darwin down	3.50	4.50	3.67		0.31
	2.60	1.80	3.40		0.15
Darwin always reacted appropriately	3.50	2.33	3.33		0.17
	4.20	1.80	4.40		0.002*
Darwin made me less nervous during the test	2.83	2.33	2.83		0.57
	3.40	2.60	3.20		0.62
Darwin helped me to stay focused on the test	3.33	3.67	3.17		0.69
	3.40	2.40	4.00		0.11
I like Darwin	4.17	4.50	4.67		0.62
	3.20	4.20	4.60		0.20
Interested in taking Darwin to a real test?	2.17	2.17	2.00		0.92
	1.8 0	2.40	2.40		0.48

students deemed Darwin's reactions to be during the test. The nonverbal group thought Darwin's actions were not appropriate with a score of 1.8 (Slightly Disagree = 2; SD = 0.84), while the remaining groups had an average score of 4.3 (Slightly Agree = 4; SD = 0.99). This supports the nonverbal freeform responses about how "weird" Darwin's movements were during the test. The lack of understanding of Darwin's actions was interpreted as him not giving any feedback at all, which resulted in a more unpleasant experience.

Because boredom is often associated with poorer learning and behavior problems [11], it is important to note that there was a statistically significant variance in how bored the student deemed him- or herself to be throughout the test in Trial 2. For both the verbal group and the group with a mixture of verbal and nonverbal cues, the average response to the question on boredom during the test was 1.8 (Slightly Disagree = 2; SD = 1.07). The nonverbal group followed with a score of 3.4 (Neutral = 3; SD = 1.52), while the group with no agent was the most bored with a score of 4.6 (Agree = 5; SD = 0.55). This shows that the verbal group and the group with both verbal and nonverbal cues were able to minimize boredom the best when compared to the other groups.

The freeform responses yield a range of responses – some students felt like the robotic platform was wasting space, while others enjoyed the robot's presence. In particular, the students said that the robot was a "friendly looking robot with a friendly voice." Similarly, another student said robot was "cute...and friendly." Lastly, a student stated that he or she felt more confident when the robot was assisting with the learning scenario. Although there were a lot of positive freeform responses, we would like to make improvements in the system in the near future to decrease the amount of negative responses received from students.

5.1.6 Conclusion

Across all interaction types, verbal, nonverbal, and both, the students enjoyed Darwin and were not distracted by his presence during the test. They were able to build a relationship with Darwin and did not wish to disappoint him with their performance. When compared to having no educational agent present, every interaction type that Darwin implemented was successfully able to maximize the time used in the learning environment. This was achieved by using the engagement model to monitor progression through the test and effectively eliminate idle time. In particular, the verbal engagement implemented on Darwin was able to reach this goal best, although by a small margin. In addition to minimizing idle time, the standard deviation was also extremely low when compared to the control group. Lastly, a mixture of verbal and nonverbal cues tends to have the least amount of boredom associated with it, which is ideal for a richer learning environment. Overall, the use of only nonverbal cues such as gestures shows no significant trends when compared to verbal cues; therefore, this works suggests that verbal engagement is ideal for enhancing test performance with computational tasks.

5.2 Motor Tasks

In this section we elaborate on the process of embedding social interaction within a humanoid-student *motor*-learning scenario in order to re-engage students during low-demand cognitive tasks. For individuals with a motor skill disorder, repetition of recommended motor tasks is essential for learning. For the context of this study, the tasks that will be discussed are comparable to the tasks performed in a physical therapy session. Moreover, external motivational feedback is an important component of motor-task learning such that individuals can remain engaged over an extended period of time and, ultimately, improve their performance. In order to promote the repetition of recommended motor tasks, several serious games have been developed to promote compliance with motor-task learning interventions.

To advance this work, we have developed a novel framework to couple serious games with a robotic educational agent (REA) that provides motivational feedback during interaction. The REA continuously tracks the student's kinematic performance and autonomously provides motivational cues to increase the engagement and performance. The details of the student's kinematic performance will be discussed in the following chapter. However, in this section, we will focus on student engagement and acceptance of Darwin during the learning session. To determine how motivational cues affect an individual's engagement level, we have tested the complete system with 20 able-bodied adults. Namely, we varied the type of motivational feedback given from Darwin while completing a reaching task. The results show that when motivational cues are not provided during interaction, the robotic platform is perceived to be unpleasant and the students are disengaged from completing the remaining tasks. The exit surveys also suggest that when motivational cues were provided, the students had a more pleasurable experience.

5.2.1 Super Darwin Pops (SDP)

In this section we introduce a system that can potentially be used as part of the intervention protocol for individuals with motor skill disorders. We have developed a novel framework to couple the virtual reality game Super Pop VR^{TM} [39] with a robotic educational agent Darwin [45] that can autonomously provide the student motivational feedback during game play and, ultimately, increase the efficacy of the corresponding intervention protocol. We will refer to the integration of these two platforms as ***Super Darwin Pops (SDP)***. A description of the overall SDP system is shown in Figure 35 along with all of the major components.

For simplicity, we focus on upper-body motor skills, of which the most dominant form is reaching movements. The ability to reach is critical for most, if not all,

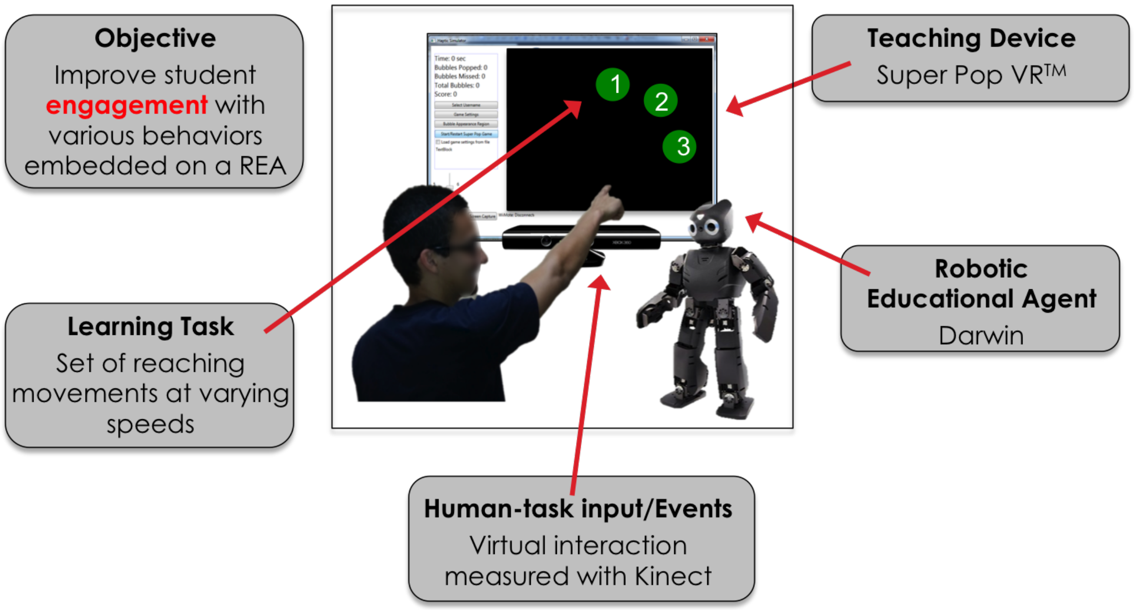


Figure 35: This shows the system diagram with the main components. 1) Darwin is the robotic platform which provides motivational feedback to 2) the student who is interacting with 3) the Super Pop VR^{TM} game [39].

activities of daily living such as feeding, grooming, and dressing [43]. Moreover, failure to substantially recover upper-extremity function can lead to depression [60]. As such, reaching movements, correlated to reaching exercises, are of interest in various rehabilitation scenarios. These exercises require a student to move from a defined initial position to a selected target position (Figure 36). In the proposed system, we evaluate the student's performance, and provide the corresponding feedback, with respect to these movement types.

5.2.1.1 Virtual Environment

In order to enable collection of a non-biased data collection process for the randomized trials, we employed a platform called Super Pop VR^{TM} [39, 38], a motivating virtual reality game used to track upper-body movements using the Kinect camera from Microsoft. The objective of the Super Pop VR^{TM} game is to interact with the virtual

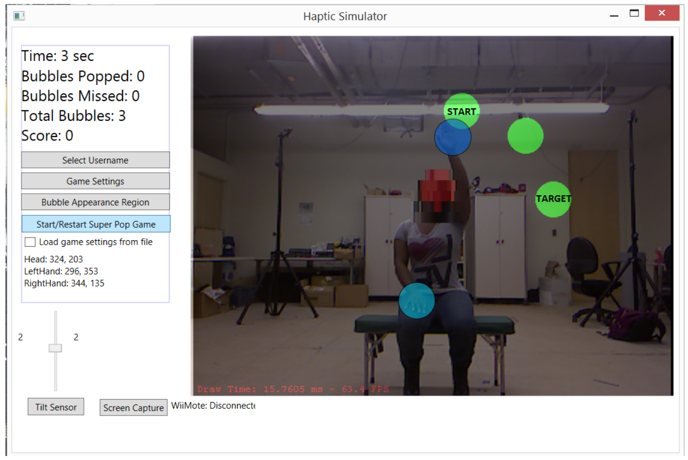


Figure 36: During game play, the student interacts with the motor-task learning game interface displayed on the projector screen. The reaching movement is shown from START to TARGET.

environment and 'pop' the virtual bubbles that appear on screen. For the student, the execution of a reaching movement is necessary to successfully reach for and 'pop' a bubble. Figure 36 is an example of popping a bubble.

For this study, students interacted with the system in a lab environment as shown in Figure 37. The environment settings were maintained constant in order to maintain consistency. The virtual reality game screen was projected onto a large screen via a projector connected to a PC laptop. The chair height upon which the student sat was 39cm tall, the distance between the student's chair and the Kinect camera was 200cm, and the distance between the projector and the screen was 470cm. For each student, a total of six reaching movements were collected per arm during game play.

5.2.1.2 The Robotic Platform

To enable interaction with the virtual environment and feedback to the human, Darwin was pre-programmed with a library of verbal and nonverbal behaviors. The nonverbal behaviors demonstrate the reaching task, whereas the verbal behaviors enable the robotic agent to provide socially supportive phrases for re-engagement as

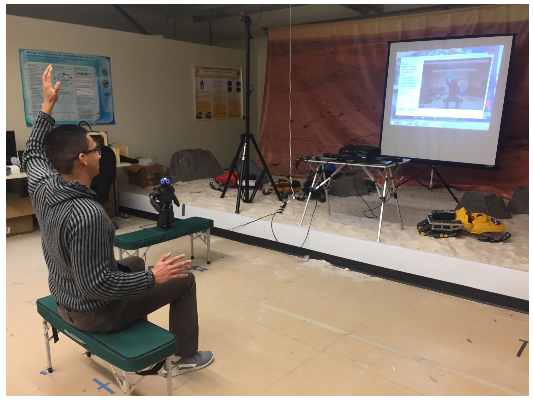


Figure 37: Actual experimental setup.

the student navigates through the learning task. These behaviors enabled Darwin to work together with the student in the virtual environment as a teammate, yet also provide motivational feedback when the task was performed, regardless of performance (i.e. how well they are completing the reaching task). The 'teamwork' aspect that is conveyed through these socially supportive verbal and nonverbal behaviors is ideal for optimal learning [78]. The behaviors used in this investigation to provide motivational feedback to the student are described previously in Section 4.2.

5.2.2 Hypothesis

Although studies have been conducted that show engagement levels can be increased through use of verbal motivational feedback when learning *computational* tasks [23], we hypothesize that similar trends can be found using verbal motivational feedback when learning *motor* tasks.

5.2.3 Experimental Design

To evaluate the effectiveness of the robot providing motivational feedback while learning a motor task, we employed a between-groups design for this study. A total of 20 participants took part in this experiment and all were recruited from the Georgia Institute of Technology in Atlanta, GA. The population consisted of both males and females in the age range of 18-45 years old (mean = 27.4, SD = 4.8 years; Male: 11, Female: 9). The participants signed the IRB (Institutional Review Board) approved consent form before engaging in the testing sessions.

For this study, the students were randomly assigned to a group:

- **Group 1: (Control)** The REA gives instruction to complete a reaching task in the virtual environment, but does NOT give motivational feedback upon task completion.

- **Group 2:** The REA gives instruction to complete a reaching task in the virtual environment, and gives motivational feedback upon task completion.

Darwin is positioned to the left of the projection screen, yet between the student and the screen. The robot is perpendicular to both the screen and the student, so that it is able to see and interact with both entities (Figure 38). During game play, the students' movements are mapped into the virtual environment where they are surrounded by the virtual bubbles as part of the game. At the start of the game, Darwin gives a verbal introduction and discusses the activity that the student will perform. The purpose of this introduction is to eliminate the novelty of the robot from the testing sessions and provide some low-level instructions. The following is the script of the verbal introduction:

> *"Hello. My name is Darwin, and I will be playing Super Pop with you today. I will ask you to complete a series of tasks, and I would love it if you would follow my instructions. When you're ready, please raise both of your hands as high as you can."*

Darwin then turns his head around and directs his eye gaze towards the projecting screen. During game play, a set of Super Bubbles (SB) are presented in such a way that indirectly prompts the students to complete three reaching tasks. When a SB

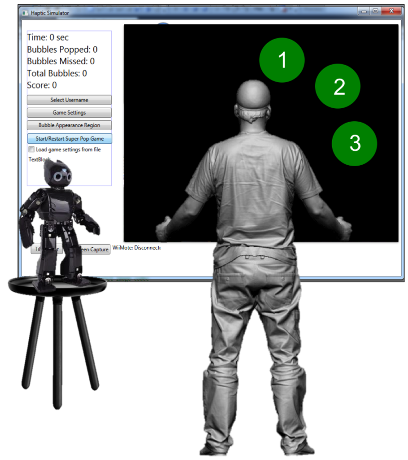


Figure 38: Darwin is strategically placed between the student and the game, which enables him to interact with both.

set is displayed, Darwin will direct his gaze towards the student and instruct the student to perform a task. The tasks are given in the same order to all of the students. For both Groups 1 & 2, Darwin states the task verbally and also provides the students with an iconic gesture to further demonstrate the desired performance (discussed further in the next chapter). However, for Group 2, Darwin also interjects a motivational phrase upon task completion. Lastly, the interaction with the system is followed with an exit survey to evaluate whether or not the students enjoyed their experience.

5.2.4 Results

The survey following the interaction with the system covered topics about the student's motivation, the game's perceived difficulty, and the student's overall physical effort exerted. We asked the students to rate their agreement with a series of statements on a 5-level Likert scale that ranged from 1 (Strongly Disagree) to 5 (Strongly Agree). The overall means and standard deviations of the students' responses are displayed in Table 17. In general, there seems to be a relatively positive perception of the entire system with the REA.

Table 17: Statistical analysis of survey responses

Statement	Group 1		Group 2		p-value
	m	SD	m	SD	
I enjoyed playing the game overall	4.00	1.05	4.10	0.56	0.39
I think I performed well in the game	3.90	0.87	4.20	0.91	0.27
I would like to play this game more often	2.70	1.05	3.30	1.15	0.08
I would be willing to play the game every day for a few minutes	2.80	1.31	3.30	1.05	0.10
I would be willing to play the game twice a week for at least 30 minutes	2.50	1.26	3.00	0.94	0.10
It would be nice if I could play the game with other children	3.40	1.57	3.60	1.26	0.38
The game was so engaging that I lost track of the time	2.90	1.19	2.60	0.96	0.24
Training with this game is less fun than with regular physiotherapy	1.60	0.89	2.16	0.75	0.24
If repeatedly played, my speed and accuracy would improve	4.00	1.33	4.00	0.66	0.50
The game was too fast - I would have liked to play a slower version	1.80	0.42	1.90	0.56	0.29
The game was too difficult. I would have liked to play an easier version	1.60	0.51	1.60	0.51	0.50
I could predict what was going to happen after I had made a movement	4.00	0.94	3.80	0.91	0.17
I found it hard to play the game by moving my arms	1.30	0.67	1.4	0.51	0.37
I become more tired from playing this than regular physiotherapy	2.20	0.44	1.66	0.81	0.19
I have learned new movements by playing this game	2.11	0.78	2.00	0.70	0.38
I think I could learn new movements by playing the game more often	3.33	1.11	3.11	1.36	0.42

In addition to the questions discussed in the previous subsection, we also left room on the survey for students to provide freeform comments that reflected their experience as a whole. Out of the 20 people who participated in the study, 4 students decided to provide comments. In particular, two comments were about the system with the REA not providing verbal feedback (Group 1), and two comments were about the system with the REA providing verbal motivational feedback (Group 2). These responses are outlined in Table 18.

5.2.5 Discussion

In regards to motivation, students in both groups stated that they enjoyed playing the game with a score of 4.05 (Agree = 4; SD = 0.82). This results correlates with results from previous studies on student motivation with VR systems and/or robot-assisted

Table 18: Freeform responses

Group	Response
1	“Sound of the robot was not clear. Missed a few instructions.”
	“Consider changing the tone of Darwin’s voice to have a higher pitch or place him closer to the user so that his instructions can be heard clearly.”
2	“Sometimes it was too slow for me to realize what was going on. The visual feedback of where you were (where your arms were) was really useful. I see this extending past cerebral palsy and being really useful for rehab of all kinds (e.g. shoulder).”
	“Could not make out what Darwin instructed on the first set of bubble pops. After that I inferred what he wanted from his movements.”

rehabilitation [32, 31, 58, 25]. However, if we compare Group 1 to Group 2, the students enjoyed playing the game slightly more in Group 2 with a score of 4.1 and with a lower standard deviation of 0.56. The addition of motivational phrases in the learning scenario also increased the students’ confidence in how well they perceived their performance. Group 2 stated that they thought they performed well with a score of 4.2 (Agree = 4; SD = 0.91), while Group 1 had a score of 3.9 (Agree = 4; SD = 0.87).

Furthermore, we are interested in developing a learning system that students will be able to use long term to ensure optimal results. Group 2 exhibits a lot of potential towards reaching this goal when compared to Group 1. More specifically, Group 2 was inclined to play the game more often with a score of 3.3 (Neutral = 3; SD = 1.15) whereas Group 1 had a score of 2.7 (Neutral = 3; SD = 1.05). We also probed both groups to see if they would prefer to use the system every day for a few minutes or twice a week for 30 minutes. As expected, Group 2 was more accepting of both ideas with an average score of 3.15 (Neutral = 3; SD = 0.98), while Group 1 had a score of 2.65 (Neutral = 3; SD = 1.26).

Lastly, the freeform responses were able to reveal subtle differences between the two groups, although only 20% of the students decided to give feedback. In general, the students in Group 1 were more critical of the entire system, and did not hesitate to state the areas that needed improvement. For instance, one student stated that the "sound of the robot was not clear" and because of this he or she "missed a few instructions." Another student of Group 1 stated that we should "consider changing the tone of Darwin's voice to have a higher pitch or place him closer to the student so that his instructions can be heard clearly."

However, when we looked at the responses that Group 2 gave about the system, the students seemed to be more sensitive in their critique and ended on a positive note about the system. For instance, one student stated that "sometimes it was too slow...to realize what was going on." Then the same student ended by stating that "the visual feedback of where you were (where your arms were) was really useful" and he or she can "see this extending past cerebral palsy and being really useful for rehab of all kinds (e.g. shoulder)." Another student in Group 2 stated that he or she "could not make out what Darwin instructed on the first set of bubble pops." The same student ended by saying that after that he or she "inferred what he wanted from his movements."

The freeform responses suggests similar trends show in the prior engagement study when learning computational tasks (Section 5.1). In particular, one can argue that in Group 2, a personal relationship was built between Darwin and the human through use of these socially-supportive motivational phrases. Because of this, the students were more sensitive to Darwin's 'feelings' when giving feedback. As previously noted, the relationship between the student and the instructor is a very critical component to achieving optimal learning.

Although we received beneficial information by conducting this study, there is still obvious room for improvement with the proposed system. It was apparent in both groups that at times it was difficult to hear Darwin's voice at the start of each session. Therefore, in future studies we plan to make the necessary adjustments to the protocol.

5.2.6 Conclusion

Across both groups, everyone enjoyed their interaction with the system. Unlike the control group, it was evident that the students were able to build a relationship with Darwin when he provided motivational cues during the learning scenario. In particular, the students did not wish to disappoint him with their responses and made sure to mention something positive about the system after their critique. In addition, the group that was provided motivational cues during the learning scenario had a richer experience. Based on their responses, they are more likely to interact with the system long-term, which is ideal for optimal learning and retention. Therefore, this works suggests that verbal engagement is ideal for enhancing performance with motor tasks.

CHAPTER VI

LEARNING WITH ROBOTIC EDUCATIONAL AGENTS

In addition to increasing student engagement with Robotic Educational Agents (REA), we aim to increase student learning/performance. Our studies have provided many insights on effective methods to communicate tasks and corrective feedback to students during a learning scenario [21, 37]. We discuss the overall system approach, which builds on the current REA system that employs verbal cues embedded on the robotic agent to increase student motivation (Chapter 5). The result of this study fulfills our fourth contribution:

> *4. Develop a system that uses verbal and nonverbal cues embedded on a robotic platform to increase student performance through guided instruction and corrective feedback.*

6.1 Guided Instruction

This work builds on the previous study, in that we have chosen to focus on learning motor tasks similar to those performed in physical therapy. We utilize the same SDP framework. However, instead of evaluating the effects of verbal motivational cues on student engagement, we will focus on improving student performance by providing guided instruction during interaction. The robotic agent continuously tracks the student's kinematic performance and autonomously provides objective verbal and nonverbal cues in order to increase the efficacy of the intervention. To determine how various cues affect an individual's kinematic performance, we have tested the complete system with 20 able-bodied adults. Namely, we computed the total amount of time it took the students to successfully complete a reaching task as a function of the verbal or nonverbal cues received. The results show that movement times improve at a faster rate for the group provided with both verbal and nonverbal feedback versus verbal feedback alone. Exit surveys also suggest that the targeted population deemed the system enjoyable.

6.1.1 Description of System

6.1.1.1 Performance Metrics

There are several kinematic parameters associated with reaching movements, including movement smoothness [76], movement speed [28], movement time [63], and elbow range of motion (ROM) [36]. This combination of kinematic parameters best describes the reaching movements performed by the student while interacting with the SDP game. In regards to this study, we only focus on computing and correcting the student's *movement time (MT)* [63].

The MT parameter describes the amount of time the student takes to complete a reaching task (i.e. to move from one bubble to another). The initial and final positions of the student's hand are captured when they 'pop' the start and target bubbles respectively. The MT for a given task is determined by the system clock, which starts when the student 'pops' the initial bubble and stops when they 'pop' the target bubble.

6.1.1.2 The Robotic Platform

To enable interaction with the virtual environment and feedback to the human, Darwin was pre-programmed with a library of verbal and nonverbal behaviors. These behaviors enabled Darwin to work together with the student in the virtual environment as a teammate, yet also provide corrective feedback when the task was performed incorrectly. The 'teamwork' aspect that is conveyed through these socially supportive verbal and nonverbal behaviors is ideal for optimal learning [78]. The behaviors used in this investigation to provide corrective feedback to the student are described in

Table 19, and the basic arm movement gesture is displayed in Figure 39.

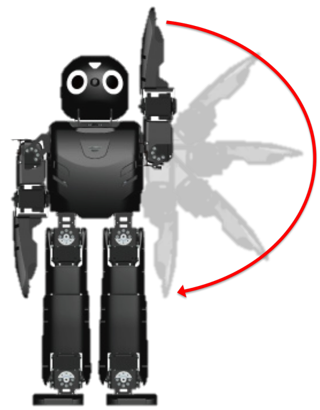


Figure 39: The robotic platform Darwin performing the basic reaching movement gesture. There are a total of three gestures, and the key difference is movement time.

Table 19: Instructional feedback

Task	**Verbal**	**Nonverbal**	
		Movement Time	**Description**
1	"Please pop bubbles one through three at a speed that feels normal."	1.60s	Darwin extends his left arm above his head. He then moves his shoulder joint 180ntil his arm is downward and parallel to his torso (Figure 39).
2	"Please pop bubbles one through three again, but move as slow as possible."	16.32s	
3	"Please pop bubbles one through three again, but move at a speed that is a little slower than normal."	5.28s	

6.1.2 Hypothesis

Feedback is provided to individuals who are learning motor tasks to help improve the individual's movement performance. Because feedback can come from a variety of sources, we have chosen to utilize an embodied robotic agent to provide corrective feedback to the students. Although studies have been conducted that show verbal cues alone are best for *motivational* feedback [23], we hypothesize that receiving a combination of verbal and nonverbal cues will be better than receiving verbal cues alone for *instructional* feedback.

6.1.3 Experimental Design

To evaluate the effectiveness of the robot platform providing corrective feedback during a motor task, we employed a between-groups design for this study. A total of 20 participants took part in this experiment and all were recruited from the Georgia Institute of Technology in Atlanta, GA. The population consisted of both males and females in the age range of 18-45 years old (mean = 28.4, SD = 5.7 years; Male: 15, Female: 5). The participants signed the IRB (Institutional Review Board) approved consent form before engaging in the testing sessions.

For this study, all students received instructional feedback selected randomly with one variable between the two groups:

- **Group 1:** The robotic platform gives *verbal* instructional feedback to assist the student when completing a reaching task in the virtual environment.
- **Group 2:** The robotic platform gives a combination of *verbal and nonverbal* instructional feedback to assist the student when completing a reaching task in the virtual environment.

Before the students played the SDP game, they were assigned to one of the groups. All students completed a total of two games with their dominant hand. The first game is the original Super Pop VR^{TM} game played without the robotic platform and, therefore, no feedback is given to the student. This initial game is used to familiarize the student with the motor-task learning game and eliminate any novelty effects. The second game is played with the robot platform. The group that the student is assigned to determines the level of corrective feedback administered from the robot platform.

For the games where the students play with the robot platform, Darwin is positioned to the left of the projection screen, yet between the student and the screen. During game play, a total of three sets of bubbles are displayed during this game, and each set is correlated to one of the three tasks described in Table 19. When a bubble set is displayed, Darwin will direct his gaze towards the student and instruct the student to perform a task. For Task 1, the student is instructed to move at a 'normal' speed; for Task 2, the student is instructed to move 'as slow as possible'; for Task 3, the student is instructed to move at a speed 'a little slower than normal.' The tasks are given in the same order to all of the students. For both Groups 1 & 2, Darwin states the task verbally; however, for Group 2 Darwin also provides the students with an iconic motor gesture to further demonstrate the speed that is desired of them. Lastly, the interaction with the system is followed with an exit survey to ensure that all of the students understood the instruction given from Darwin.

6.1.4 Results

6.1.4.1 Performance

To evaluate the performance of the proposed system, we collected data from the students of both groups and averaged the resulting outcome metrics with respect to movement time for each task. For the context of this study, the slower the student moves, the higher is their movement time. Figure 40 shows the average movement times of each group for each task along with the movement times of the robot platform Darwin. By observing the linear trend-lines moving from Task 1 to Task 2, it is evident that Group 1, Group 2, and the robotic platform Darwin improve at a positive rate with a slope of 0.25, 0.56 and 0.90, respectively. In particular, Group 2 improves at a faster rate than Group 1 and is able to perform closer to the robot's guided direction. Both groups also reach their highest movement time during Task 2, as expected.

Because we hypothesized that receiving a combination of verbal and nonverbal cues will be better than receiving verbal cues alone for corrective feedback, we also conducted a one-tailed t-test on the collected data as shown in Table 20. Here we see that there are little differences between the performance of the groups for Task 1 and Task 3. However, for Task 2 there is a statistically significant difference between

Group 1 (verbal cues only) and Group 2 (combination of both verbal and nonverbal) with a p-value of 0.01.

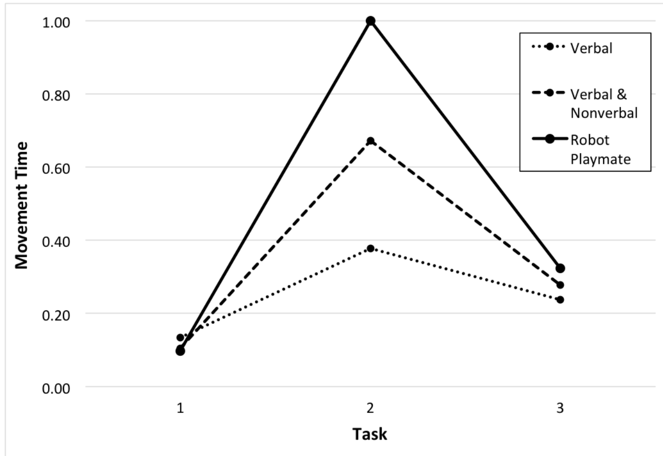


Figure 40: The normalized average movement times with respect to each task for Group 1, Group 2, and the Robot platform.

Table 20: Statistical Analysis of Movement Time

Task	**Group 1**	**Group 2**	**p-value**
1	0.13	0.11	0.14
2	0.38	0.67	0.01*
3	0.24	0.28	0.26

6.1.4.2 Exit Survey

The survey following the interaction with the system covered topics about the student's motivation, the game's perceived difficulty, and the student's overall physical effort exerted. We asked the students to rate their agreement with a series of statements on a 5-level Likert scale that ranged from 1 (Strongly Disagree) to 5 (Strongly Agree). The overall means and standard deviations of the students' responses are displayed in Table 21. In general, there seems to be a relatively positive perception of the entire system with the robotic playmate.

Table 21: Statistical Analysis of Survey Responses

Statement	Group 1		Group 2		p-value
	m	SD	m	SD	
I enjoyed playing the game overall	4.20	0.78	4.00	1.05	0.29
I think I performed well in the game	3.60	0.96	3.90	0.87	0.13
I would like to play this game more often	3.10	1.28	2.70	1.05	0.23
I would be willing to play the game every day for a few minutes	3.30	1.15	2.80	1.31	0.13
I would be willing to play the game twice a week for at least 30 minutes	3.10	1.19	2.50	1.26	0.10
It would be nice if I could play the game with other children	4.00	1.05	3.40	1.57	0.21
The game was so engaging that I lost track of the time	2.90	0.99	2.90	1.19	0.50
Training with this game is less fun than with regular physiotherapy	3.00	1.15	1.60	0.89	0.07
If repeatedly played, my speed and accuracy would improve	4.20	0.42	4.00	1.33	0.35
The game was too fast - I would have liked to play a slower version	2.00	0.66	1.80	0.42	0.25
The game was too difficult. I would have liked to play an easier version	2.00	0.66	1.60	0.51	0.05
I could predict what was going to happen after I made a movement	3.50	1.17	4.00	0.94	0.12
I found it hard to play the game by moving my arms	1.70	0.48	1.30	0.67	0.05
I become more tired from playing this than regular physiotherapy	1.85	0.69	2.20	0.44	0.03*
I have learned new movements by playing this game	1.77	0.66	2.11	0.78	0.11
I think I could learn new movements by playing the game more often	3.62	0.91	3.33	1.11	0.50

6.1.4.3 Freeform Responses

The survey following the interaction with the system allowed students to provide freeform comments that reflected their experience as a whole. Out of the 20 people who participated in the study, 7 students decided to provide comments. In particular, five comments were about the system with the robotic platform providing verbal feedback (Group 1), and two comments were about the system with the platform providing both verbal and nonverbal feedback (Group 2). These responses are outlined in Table 22.

6.1.5 Discussion

Although the survey results (Table 21) show that across all groups there is an overall positive response to the system, the performance results show that there is still potential in improving the efficacy of motor-task learning interventions by coupling

Table 22: Freeform responses

Group	Response
1	"Right now, this is more of a test run than a game. If (or when) the patterns' locations or order change, then I could be more compelled to play. Consider color-coding the patterns (the 1-2-3 bubbles)."
	"I didn't know which bubbles were 1-3 so I delayed the first time. (Bubbles were not labeled.) My hands weren't found at first. Maybe tell me in the game to move them or put them in a certain place."
	"The sound from the speaker of the robot was not clear. Better quality of speaker would have made this trial better to follow the instruction given by the robot."
	"The question about 'game interface' was vague as is saw a video screen of myself overlaid with this activity. So I never thought of there being an interface as presented. Maybe Darwin's instruction could include "keeping your arm fully extended." I couldn't accurately comment on length of time to commit to the game w/o knowing the goal. Maybe the survey should explicitly ask about previous physiotherapy experience."
	"Robot partner sometimes difficult to understand."
2	"Sound of the robot was not clear. Missed a few instructions."
	"Consider changing the tone of Darwin's voice to have a higher pitch or place him closer to the user so that his instructions/encouragement can be heard clearly."

serious games with a robot platform that provides real-time corrective feedback. The students enjoyed their experience, as well as the robot's presence. Not only was the system accepted, the students verbalized that they wanted to continue playing the game with Darwin even after testing was complete.

In regards to performance, Figure 40 suggests that the corrective feedback cues that Darwin provided had a positive effect on the outcome of the *Movement Time* performance metric, further indicating that there is benefit in providing concurrent corrective feedback during game play. Overall, both groups began at a certain baseline for Task 1, were able to slow down for Task 2, and then speed up for Task 3. However, we observe key differences between Group 1 and Group 2, which provide feedback to the student in unique ways. More specifically, if we look at Task 2 when Darwin instructs the student to move 'as slow as possible,' Group 1 has a normalized movement time of 0.38 (SD = 0.26) and Group 2 has a normalized movement time of 0.67 (SD = 0.30). This result suggests that the addition of nonverbal cues for feedback on the robotic platform allows the student to perform better, as hypothesized.

Lastly, the freeform responses received from the students gave a lot of insight to take into consideration for future studies. A large number of students commented that the robot platform's voice was hard to understand; however, the results in Figure 40 show that the were able to accurately complete the tasks as instructed. For Group 2, this could be attributed to the use of gestures to further explain the task.

6.1.6 Conclusion

The proposed system for providing objective instructional feedback to individuals who have some form of motor skills disorder while they perform recommended motor tasks has the potential of improving motor learning associated with such protocols. The comparison between the groups' average movement times showed that the students in Group 2 (verbal and nonverbal) improved faster than the students in the Group 1 (verbal only), while the exit surveys showed that the target population accepts a system such as ours regardless of the feedback type they received. As such, these results suggest that the proposed system can increase learning best by using both verbal and nonverbal cues for instructional feedback. In the future, we plan to improve our methodology by 1) increasing the number of students, 2) redefining the threshold values for the performance metrics, and 3) testing with the motivational cues embedded on the robotic platform.

6.2 Corrective Feedback

Previous studies have shown that external corrective feedback provided to individuals performing movement tasks increases the motor learning associated with their intervention protocols, thus allowing for the individuals to rapidly improve on their performance. After showing in [21] that a combination of verbal and nonverbal cues is the most efficient method for providing guided instruction, we now show that motor learning associated with intervention protocols can be further increased by embedding human-like behaviors on the robotic platform such that it provides continuous low-resolution feedback. This would allow student performance to improve at each iteration of feedback and, ultimately, converge to an individualized learning goal. To show that students can converge to a reference point or learning goal for a given kinematic parameter, we recruited 14 children to interact with the system. Results show that students converged to their learning goal in an average of four trials. Further analysis shows that the students' response to the system imitates that of an overdamped second-degree order system, suggesting that all individuals that interact with our proposed system will reach their learning goals.

6.2.1 Description of System

We base this work on our previous findings in [21] that showed how an individual's performance with respect to movement time (MT) can be affected by verbal and nonverbal corrective cues provided by a robotic platform (Section 6.1). We adhere to the Super Darwin Pops (SDP) system previously developed to aid in motor task learning for individuals who have some form of motor skill disorder [39, 38].

6.2.1.1 Performance Metric

As mentioned previously, movement time is a kinematic parameter of interest in motor skill learning because it directly correlates with the speed of an individual's movements. In general, for a given learning goal, the idea is to compare the individual's MT with a ground truth value to determine their kinematic performance and then influence their learning such that their MT approaches an ideal MT via corrective feedback provided by a robotic agent.

We define an ideal MT using the model of human movement, **Fitt's law** [34]. This model predicts the amount of time a student should take move from one point to another in a virtual environment. We compute a MT reference for each student as a function of the distance between the start and target virtual points and the width of the target as in [40].

6.2.1.2 Robotic Platform

Darwin was pre-programmed with a library of verbal and nonverbal behaviors, which he uses to provide appropriate corrective feedback to the students depending on their performance. Table 23 organizes these behaviors based on the student's reference MT or learning goal, where *target* is the reference MT window. Low-resolution means that minimal detail is provided to the student on how to adjust their performance. After each task the student completes, the system assesses the student's performance relative to the kinematic parameters and feeds the information to Darwin. Depending on the analysis, Darwin speaks the corresponding verbal feedback and then performs the nonverbal gesture as described in Table 23 and shown in Figure 41. This combination of low-resolution verbal and nonverbal cues is used to modify the student's kinematic performance towards reaching the learning goal.

Table 23: Low-resolution feedback given from the robotic platform

<table>
<tr><th>Movement Time, MT</th><th>Verbal</th><th>Nonverbal</th></tr>
<tr><td>$MT > target$</td><td>"Great job. Move a little faster like this..."</td><td rowspan="2">Darwin performs the gesture at the correct movement time. He extends his left arm above his head. He then moves his shoulder joint 180 degree until his arm is down and parallel to his torso (Figure 41).</td></tr>
<tr><td>$MT < target$</td><td>"Great job. Move a little slower like this..."</td></tr>
<tr><td>$MT = target$</td><td>"Fantastic."</td><td></td></tr>
</table>

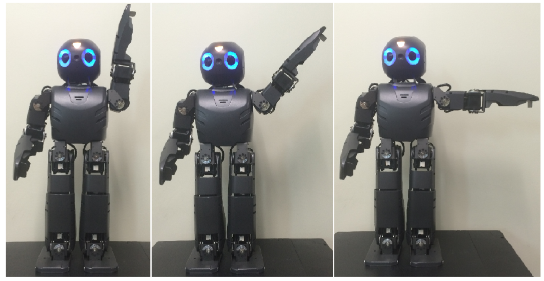


Figure 41: DARwIn-OP is used as the humanoid robotic platform in this study. Image shows snapshots of Darwin performing the nonverbal gestures.

6.2.2 Hypothesis

An embodied robotic agent providing guided instruction to individuals during motor-task learning has been shown to improve individual kinematic performance [21] and increase overall engagement levels [91]. Given that the main objective of performing movement tasks with a human instructor is to reach a learning goal over an extended period of time, we aim to reach these long-term learning goals through implementation of a robotic platform providing continuous low-resolution feedback. We hypothesize that, by interacting with this proposed system, students will improve their kinematic performance at each instance of completing a reaching task and, ultimately, converge to a targeted learning goal.

6.2.3 Experimental Design

To evaluate how well a student's kinematic performance relative to the *Super Pop* VR^{TM} game is affected by the feedback provided by a robotic platform, we recruited 14 children to interact with the system. Five females and nine males, ranging in age between 15 and 16 years (mean age = 15.5 years, standard deviation = 0.5 years), participated in this study. The parents of the participants signed the IRB (Institutional Review Board) approved consent form allowing their children to engage in the testing sessions.

At the start of each session, Darwin instructs the students to raise their arms as high as possible. The system is then calibrated to position the Super Bubbles as previously described. In addition, the system computes the student's reference movement time (MT) and a $\pm$ 150 ms margin of error is added to the final value. Each student interacts with the system for one round. Each round consists of three main steps: 1) student performs the reaching task, 2) the system compares the student's MT to their reference, and 3) Darwin provides the corresponding corrective feedback until they reach the goal or until the time runs out (6 minutes). A flowchart with the logic of the testing sessions is shown in Figure 42.

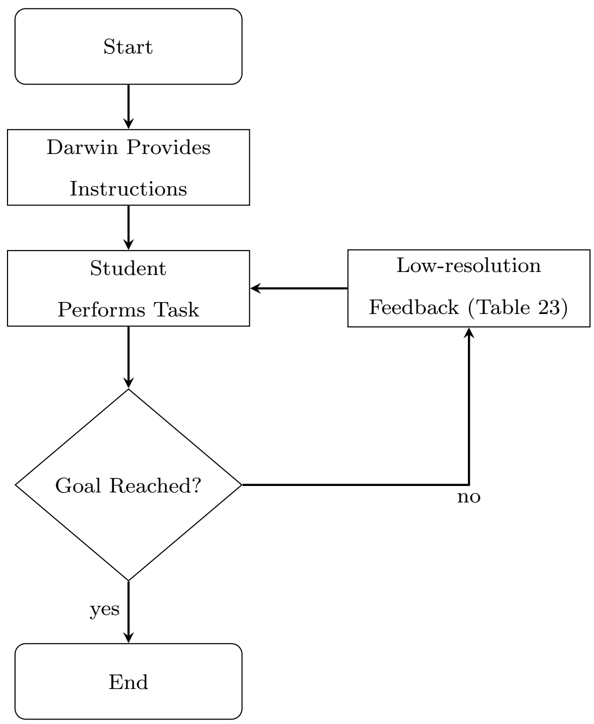


Figure 42: Flowchart describing the interaction between the student and the system.

At the start of each round we make sure that the students' starting MT is much greater than the reference MT by having Darwin speak the following:

> *"Please pop bubbles one through three and move as slow as you can. Like this..."*

Darwin then performs the nonverbal gesture that is 10 times slower than the one described in Table 23 to show the student how the reaching task should be completed. These instructions prompt the students to move as slow as they can thus defining a common starting point between all students.

6.2.4 Results

Thirteen out of the 14 students converged to their corresponding movement time (MT) references. The students that converged reached their reference windows at an average of **4 ± 1.8 trials**. The MTs for all students for their first and last trials, and their corresponding reference MTs are shown in Figure 43 and Table 24. By comparing the students' MTs for their last trials to their respective reference MTs, we can see that all but Participant 10 converged to their goals. As an example, the response of Participant 4 who converged to the reference in 3 trials is shown in Figure 44. On the other hand, the response of Participant 10 who did not converge is shown in Figure 45. We discuss some of the potential reasons for this student not converging in the following section.

We performed a two-tailed paired t-test analysis to determine if the amount of individuals that converged out of all who participated is statistically significant. Let $\vec{a}$ and $\vec{b}$ be two vectors containing the absolute difference between the students' MTs and their corresponding reference MTs for their first and last trials respectively, where each element belongs to a different student. Let $\vec{d}$ be the vector containing the difference between these $\vec{a}$ and $\vec{b}$ (15).

$$\vec{d} = \begin{bmatrix} d_1 \\ d_2 \\ ... \\ d_n \end{bmatrix} = \begin{bmatrix} |Ft_1 - Rt1| \\ |Ft_2 - Rt2| \\ ... \\ |Ft_n - Rtn| \end{bmatrix} - \begin{bmatrix} |Lt_1 - Rt1| \\ |Lt_2 - Rt2| \\ ... \\ |Lt_n - Rtn| \end{bmatrix} \tag{15}$$

where n is the number of the students being analyzed, Ft_i and Lt_i are the i^{th} student's

MT in the first and last trials respectively, and Rt_i is the i^{th} student's reference MT. Thus, we define our null-hypothesis as: the sample mean of $\vec{d}$ is equal to zero (i.e. there is no statistical difference between $\vec{a}$ and $\vec{b}$). Our t-test analysis on $\vec{d}$ results in a p-value $\ll 0.01$. As such, we reject the null-hypothesis and conclude that there is a statistical difference between the students' performance in their first and last trials. This suggests that the student that did not converge to their reference MT does not affect our claim that individuals will reach their performance goals by interacting with our system.

To further support this claim, we plotted two boxplots that summarize the absolute differences between the student' MTs and their corresponding reference MTs for the first and last trials of all students (Figure 46). We use this information to visually compare how far away the students were from their reference values in the first trials with how far away they were in the last trials. As expected, the students' MTs were further away from their references in the first trials than they were in the last. The median difference for the first trials is much greater than zero (**4,051.3 ms**), and the median difference for the last trials is approximately equal to zero (**67.6 ms**).

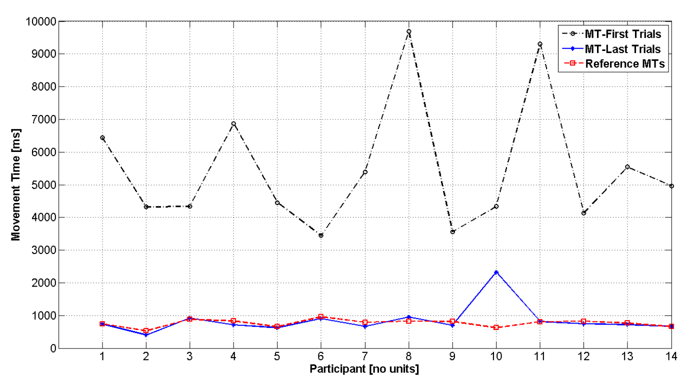


Figure 43: Participants' MTs at their first and last trials, and their respective reference MTs.

Table 24: Quantitative results for all participants

Participant	Reference MT [ms]	Start MT [ms]	End MT [ms]	Trials to Convergence
1	741.50	6445.29	735.73	5
2	526.99	4320.93	402.59	3
3	883.51	4335.50	919.65	4
4	830.11	6878.28	714.98	3
5	659.69	4453.29	624.07	3
6	958.16	3445.10	902.77	4
7	789.77	5391.94	662.53	7
8	828.02	9687.42	951.08	3
9	814.50	3558.70	698.05	5
10	628.62	4331.25	2322.97	N/A
11	804.38	9306.65	814.15	4
12	823.45	4138.47	743.69	2
13	768.07	5544.83	722.22	1
14	653.31	4962.00	667.03	8

6.2.5 Discussion

The fact that all students converged at an average of four trials suggests that it does not take much robot guidance before they reach the reference window. As expected, the students that converged to their reference movement times (MTs) modified their behavior according to Darwin's commands after completing each reaching task. Participants moved faster (decreased their MT) in the trials where their previous MT was greater than the reference window, and moved slower (increased their MT) in the trials where their previous MT was less than the reference window. Regarding Participant 10 who did not converge, Figure 45 shows that the student did not correct his/her behavior according to Darwin's commands for 5 out of 10 trials. Given that Participant 10 did not follow Darwin's feedback cues in half of the attempted trials, the student's behavior can be considered to be random. This analysis suggests that, in order for the students to improve their kinematic performance and reach their goals, they have to understand Darwin's commands and do their best to follow them.

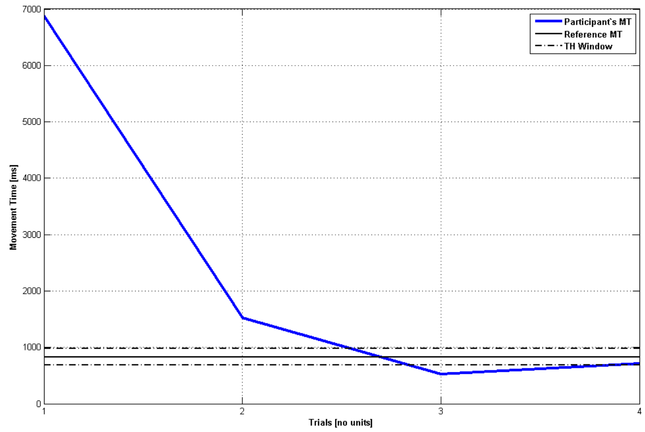


Figure 44: MT response curve of Participant 4: convergence at Trial 3.

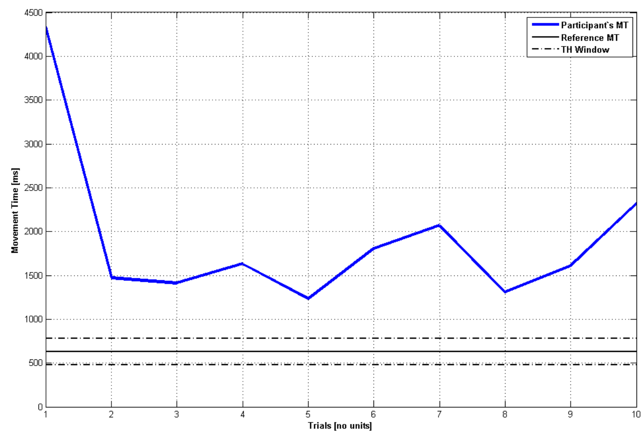


Figure 45: MT response curve of Participant 10: didn't converge.

Otherwise, their random behavior would be comparable to not receiving targeted feedback at all.

The students' behavior through their interaction with the system is an important factor that determines if the student will converge to the reference or not. As such, to determine how well all students followed Darwin's feedback cues, we analyzed their performances throughout their interactions with the system. All students finished their rounds with different number of trials. This is to say that the MT curves as a function of the number of trials are of different lengths for all students. Thus we took

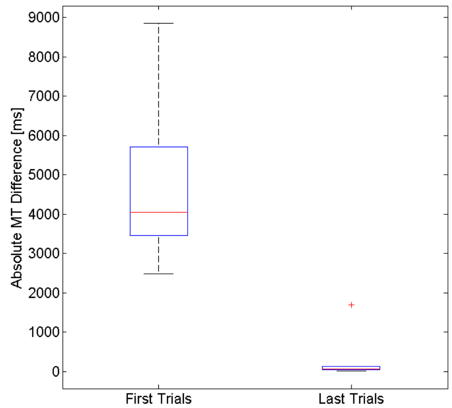


Figure 46: Boxplots showing the comparison of the absolute differences between the participants' MTs and their reference MTs for the first and last trials.

the following steps to aggregate all the response curves: 1) normalize each curve with respect to the MT relative to the maximum MT value out of all students and trials, 2) parameterize each curve with respective to their arc lengths, 3) normalize each curve with respect to their corresponding maximum arc length, and 4) interpolate each curve. After applying this methodology, the x-axis of all curves range between 0 and 1, and all curves now have the same number of points equally distributed in the x-axis. This allows us to average all the curves and obtain the average response curve as a function of the curve's arc length (Figure 47).

In general, (16) describes the solution of an overdamped second-order system as a function of time. We fitted the resulting average MT response from all students to this solution and obtained (17). It is important to keep in mind that the final MT response is a function of its arc length and not of time given that the original MT responses were a function of the number of trials. Even so, the behaviors of the average response and the fitted curve are maintained.

$$y(t) = c_1 * e^{r_1 t} + c_2 * e^{r_2 t} \tag{16}$$

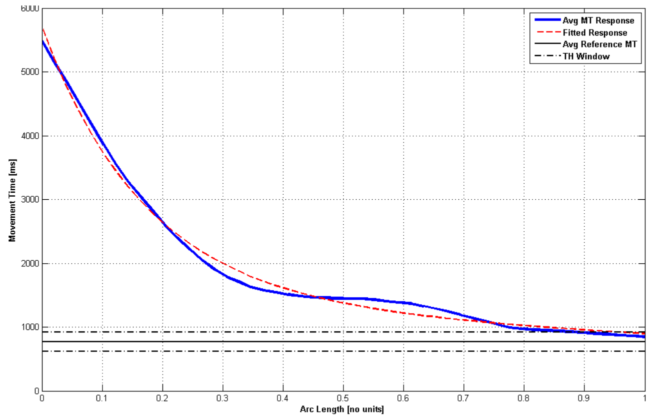


Figure 47: Average reference MT, average MT response from all students, and the fitted response to the solution of an overdamped second-order system.

$$y(t) = 4172 * e^{-5.996t} + 1541 * e^{-0.5535t} \tag{17}$$

where r_1 and r_2 are the equation's characteristic roots, and c_1 and c_2 are the coefficients of the exponentials. The comparison between the average MT response from all students and the fitted solution is shown in Figure 47. Not only does it show that the average MT response converges to the average reference MT, but it also shows that the average response mimics the solution of an overdamped second-order system. This analysis suggests that each new student that interacts with our system will have a similar response and reach their learning goal, thus supporting our claim that our system can be used to aid individuals in reaching their performance goals by providing targeted feedback from a robotic educational agent.

6.2.6 Conclusion

The developed system was able to provide the students with continuous corrective feedback that enabled 92% of the students to reach their targeted learning goals. In addition, the results mimic that of an overdamped second-order system, which suggests that each new student that interacts with our system will have a similar response and reach their learning goal.

At first glance, the results of this investigation (Figure 47) appear to resemble the learning curve discussed previously in Chapter 2. However, before we can strongly make this claim, we need to show that the learning is retained long after the initial learning goal has been reached (i.e. the curve needs to plateau). Once the reference MT or learning goal is reached in this investigation, the learning scenario terminates; therefore, the response is not able to plateau. As such, moving forward we plan to evaluate if and when retention is achieved in the proposed REA system by extending the testing sessions.

CHAPTER VII

RETENTION WITH ROBOTIC EDUCATIONAL AGENTS

Previous studies have shown that various methods of repetition integrated into the learning process is able to enhance student retention. In this chapter, we discuss the overall system approach, which integrates the engagement model, motivational feedback, guided instruction, and corrective feedback within a robotic educational agent (REA). The result of this study fulfills our fifth contribution:

> *5. Develop a learning model, which validates retention is achieved in the system after learning a new task.*

7.1 Description of System

We base this work on our previous findings, which show that a student's kinematic performance can be improved by a robotic platform providing motivational feedback, guided instruction, and corrective feedback [19, 21, 37]. We adhere to the same system used in these studies, Super Darwin Pops (SDP), which combined the embodied robotic agent Darwin [45] with the serious game Super Pop VR$^{\text{TM}}$ [39]. More details and images of how the system is coupled together can be found in Chapter 5 and 6.

7.1.1 Learning Curve

In this investigation we aim to evaluate a system that achieves individualized learning. As such, we will evaluate the performance of the students through use of the learning curve described in (18), where the student performance, P, is a function of the number of practice trials, N. The initial performance of the student is described as B, and the learning rate of the student is described as β.

$$P(N) = Be^{-\beta N} \tag{18}$$

It is noted that (18) is typically used to evaluate a single student's progress instead of aggregating the performance of a group (as we have done in prior studies).

7.1.2 Performance Model

For the context of this study, the learning goal is defined as completing a reaching task within a specified movement time window (300ms). Furthermore, performance, P, is defined as the absolute difference between the student's actual movement time, MT, and their learning goal, LG, as shown in (19).

$$P(MT) = |MT - LG|. \tag{19}$$

Because LG is a 300 ms window, it has a high and low bound. Therefore, depending on what side of the window the student's MT falls (i.e. if the student moves too fast or too slow), LG will vary between the high and low bound. More specifically,

$$P(MT) = |MT - LG_{\text{high}}| \quad \forall MT > LG_{\text{high}}, \tag{20}$$

$$P(MT) = |MT - LG_{\text{low}}| \quad \forall MT < LG_{\text{low}}. \tag{21}$$

Lastly, if the student's MT falls within the LG window, he or she has successfully reached the goal as shown in (22).

$$P(MT) = 1 \quad \textit{when} \;\; LG_{\text{high}} \geq MT \geq LG_{\text{low}}. \tag{22}$$

We have chosen to equate a successful attempt to 1 to imitate the idea that the student's performance is approaching 0, similarly to an exponential decay. This will enable us to compute a solution to the student's exponential learning curve.

7.2 Hypothesis

An embodied robotic agent providing guided instruction and corrective feedback to individuals during motor-task learning scenarios has been shown to improve individual kinematic performance [21, 37]. Given that a major component of learning is retaining the information long after instruction has been completed, we aim to reach these retention goals through implementation of a robotic platform providing repetition tactics throughout the learning process. We hypothesize that, by interacting with this proposed system, students will retain the information learned when asked to repeat the task and without being provided additional feedback. We will validate this hypothesis through use of the learning curve described in (18).

7.3 Experimental Design

To evaluate how well a student's learning relative to the *Super Pop VR*^{TM} game is retained by the methods employed by the robotic platform, we recruited 7 participants to interact with the system. A total of 3 females and 4 males, ranging in age between 19 and 32 years (mean age = 27 years, standard deviation = 4.6 years), participated in this study. The participants signed the IRB (Institutional Review Board) approved consent form before engaging in the testing sessions.

As described in the previous chapters, after the student raises their arms, the system is calibrated to position the Super Bubbles. The system then computes the student's reference movement time and a $\pm$ 150 ms margin of error is added to the final value. Each cycle consists of four main components: 1) student performs the reaching task, 2) the system compares the student's MT to his or her reference, 3) Darwin provides the corresponding corrective feedback until the goal is reached, and 4) the student attempts to repeat the correct movement $n = 3$ times without additional feedback. A flowchart with the logic of the testing sessions is shown in Figure 48.

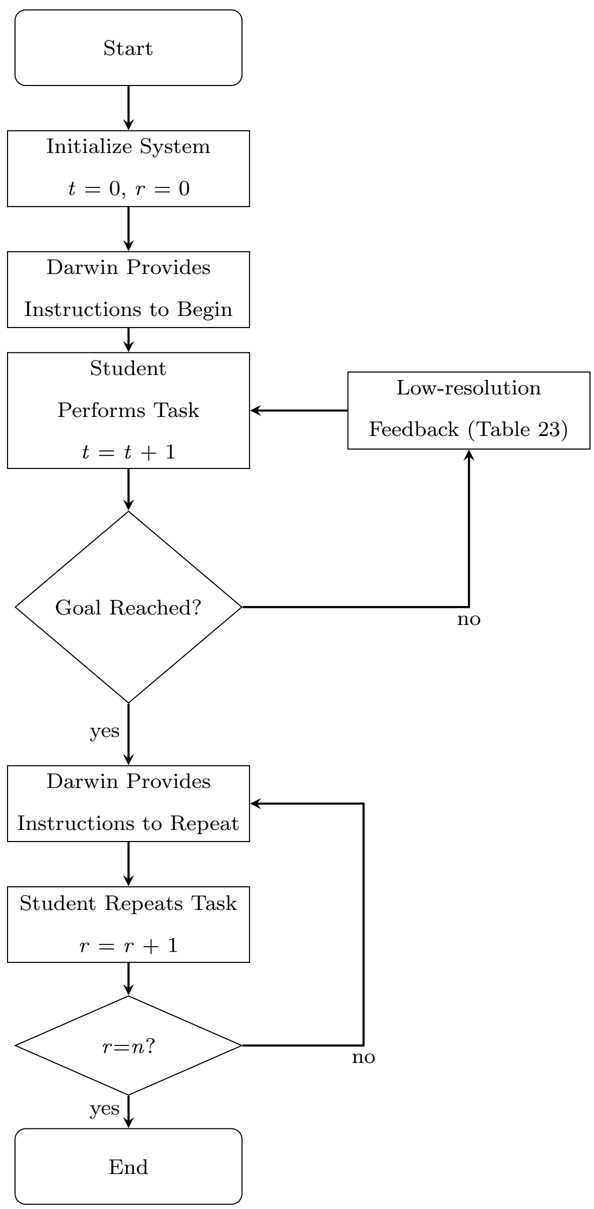


Figure 48: This flowchart describes the interaction between the student and the system. The number trials the student takes to reach his or her learning goal is described as t, and the number of repetitions the system provides once the goal is met is described as r. The minimum r required to prove the hypothesis is n (determined empirically).

7.4 Results and Analysis

All 7 students were able to reach their respective learning goals. However, the number of attempts needed to reach the goal ranged from 1 to 4 trials (m = 2.2; SD = 1.2). In addition, all 7 students were able to keep their movement times of the last 3 trials (when no additional feedback is given from Darwin) within a 300 ms window – the windows ranged from 19 ms - 236 ms. We first calculated each student's learning goal based on the reference movement time predetermined by the system (Table 25). From there, we calculated the student's performance using the model (19) - (22) as shown in Table 26. Lastly, we fitted the exponential learning curve (18) to each student's performance, and his or her solution is shown in Table 27.

Table 25: The learning goal, LG, of the participants (P1-P7) based on their reference MT (ms)

Parameter	**P1**	**P2**	**P3**	**P4**	**P5**	**P6**	**P7**
Ref. MT	608	613	746	648	788	748	646
LG	[458, 758]	[463, 763]	[596, 896]	[498, 798]	[638, 938]	[598, 898]	[496, 796]
LG_{low}	458	463	596	498	638	598	496
LG_{high}	758	763	896	798	938	898	796

If we take a closer look at the performance of Participant 1, the individual was able to reach the learning goal after 1 trial (Figure 49). In addition, after reaching the learning goal, he or she was able to replicate this task during the last three attempts when no additional feedback was given from the robotic educational agent. The student was able to keep the last three trials within an 85 ms window, which is

Table 26: Performance of participants based on model, *P(MT)*=$|MT - LG|$ (ms)

Trial	**P1**		**P2**		**P3**		**P4**		**P5**		**P6**		**P7**	
	MT	**P**	**MT**	**P**	**MT**	**P**	**MT**	**P**	**MT**	**P**	**MT**	**P**	**MT**	**P**
0	1290	523	1714	950	2214	1317	5408	4609	3262	2324	2939	2041	4982	4186
1	583	1	412	50	759	1	371	126	1103	164	1212	314	461	35
2	641	1	830	66	433	162	763	1	1689	751	732	1	739	1
3	556	1	385	77	540	55	664	1	1308	370	742	1	584	1
4	626	1	597	1	523	72	725	1	828	1	602	1	463	33
5			446	17			747	1	732	1	738	1	505	1
6			477	1					599	38				
7			465	1					835	1				

Table 27: Solution to exponential learning curve, *P(N)*=$Be^{\beta N}$

Parameter	**P1**	**P2**	**P3**	**P4**	**P5**	**P6**	**P7**
B	532	950	1317	4609	2324	2041	4186
β	2.092	1.088	1.081	2.213	1.136	1.974	1.955
R^2	0.166	0.709	-0.437	0.566	0.614	0.649	0.186

not only accurate, but extremely precise. Participant 1 was able to achieve a learning rate β of 2.092.

Participant 2 was able to reach the learning goal after 4 trials (Figure 50). In addition, after reaching the learning goal, he or she was able to replicate this task 2 of the last 3 attempts when no additional feedback was given from the robotic educational agent. The student was able to keep the last three trials within a 19 ms window, which is extremely precise. Participant 2 was able to achieve a learning rate β of 1.088.

Participant 3 was able to reach the learning goal after 1 trial (Figure 51). However, after reaching the learning goal, he or she was not able to replicate this task during the last three attempts when no additional feedback was given from the robotic educational agent. Although the student was not able to reach the learning goal again, he or she was still able to keep the last three trials within a 107 ms window,

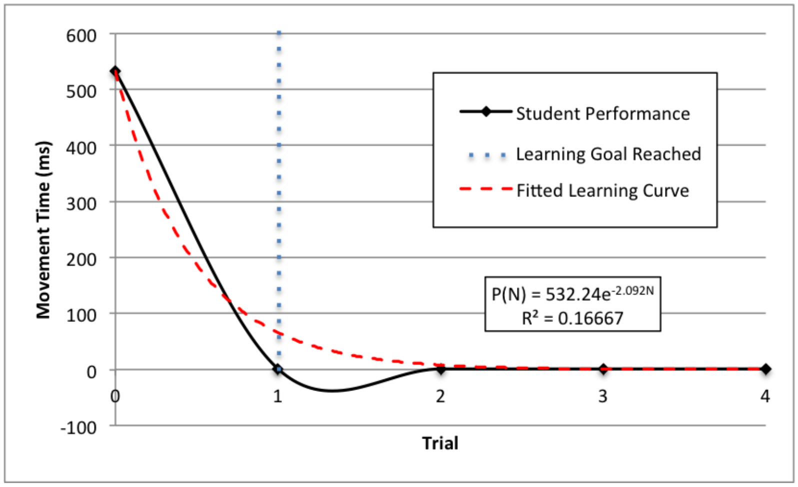


Figure 49: Participant 1's actual performance *P(MT)* plotted with his or her fitted learning curve *P(N)*.

which is very precise. Participant 3 was able to achieve a learning rate β of 1.081.

Because Participant 3 was unable to replicate the task during the learning session, we had the student perform another learning cycle. The individual was still able to reach the learning goal again after 1 trial (Figure 52). However, there was an improvement on his or her performance. Namely, after reaching the learning goal, he or she was able to replicate this task 2 of the last 3 attempts when no additional feedback was given from the robotic educational agent. In addition, the student was able to keep the last three trials within a 137 ms window, which is very precise. Participant 3 was able to achieve a learning rate β of 1.845, which is an improvement.

Participant 4 was able to reach the learning goal after 2 trials (Figure 53). In addition, after reaching the learning goal, he or she was able to replicate this task during the last three attempts when no additional feedback was given from the robotic educational agent. The student was able to keep the last three trials within an 83 ms window, which is not only accurate, but extremely precise. Participant 4 was able to achieve a learning rate β of 2.213.

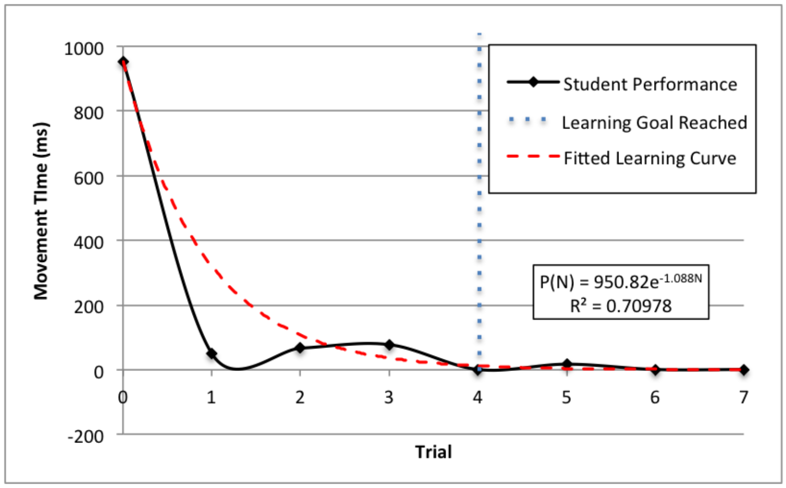


Figure 50: Participant 2's actual performance *P(MT)* plotted with his or her fitted learning curve *P(N)*.

Participant 5 was able to reach the learning goal after 4 trials (Figure 54). In addition, after reaching the learning goal, he or she was able to replicate this task 2 of the last 3 attempts when no additional feedback was given from the robotic educational agent. The student was able to keep the last three trials within a 236 ms window, which is precise. Participant 5 was able to achieve a learning rate β of 1.136.

Participant 6 was able to reach the learning goal after 2 trials (Figure 55). In addition, after reaching the learning goal, he or she was able to replicate this task during the last three attempts when no additional feedback was given from the robotic educational agent. The student was able to keep the last three trials within a 140 ms window, which is not only accurate, but very precise. Participant 6 was able to achieve a learning rate β of 1.974.

Participant 7 was able to reach the learning goal after 2 trials (Figure 56). In addition, after reaching the learning goal, he or she was able to replicate this task 2 of the last 3 attempts when no additional feedback was given from the robotic educational agent. The student was able to keep the last three trials within a 121 ms window, which is very precise. Participant 7 was able to achieve a learning rate β of

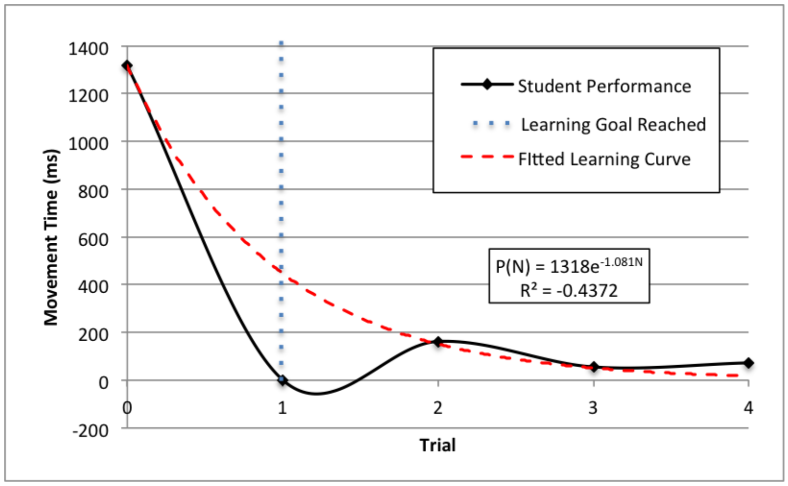


Figure 51: Participant 3's actual performance *P(MT)* plotted with his or her fitted learning curve *P(N)*.

1.955.

7.5 Discussion

All students were able to reach their learning goal; however, each individual learned at different rates as denoted by their number of trials and/or learning rate, β. In particular, 2 students were able to reach their learning goal in 1 trial, 3 students were able to reach their learning goal in 2 trials, and 2 students were able to reach their learning goals in 4 trials. This can be attributed to the fact that all students learn at different speeds and also have different learning types. Nevertheless, this investigation proves that all students are able to learn when provided the appropriate feedback.

At first glance, it appears that an ideal learning rate, β, for this motor-task is approximately 2. It is evident through Figures 49-56 that the curves with a β closer to 1 take longer to converge (plateau) to the learning goal when compared to the others. In particular, Participant 2 with a β = 1.088 and Participant 5 with a β = 1.136 both required 4 trials to reach their learning goal. However, after they reached

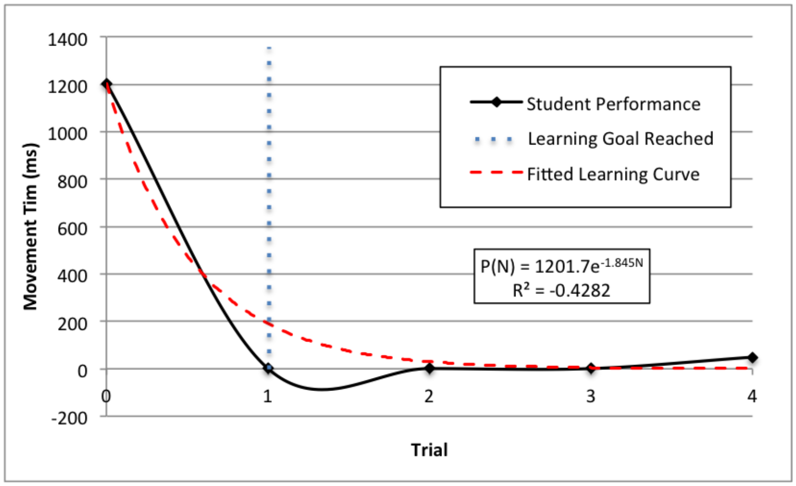


Figure 52: The results of Participant 3 after completing a second cycle of the system. Here is his or her actual performance *P(MT)* plotted with his or her fitted learning curve *P(N)*.

the goal, they both were fairly accurate in maintaining their performance. Participant 3 also has a low $\beta = 1.081$; however, he or she was able to reach the learning goal after 1 trial. The reason why the β is low is because after Participant 3 reached the goal, he or she was not successful in maintaining this level of performance. Therefore, according to this solution, learning has not been achieved.

All 7 students were able to retain their window of performance in the final three trials when no additional feedback is given from the robotic agent. Moreover, the movement time range was 19 ms - 236 ms (m=113 ms; SD = 66), where Participant 2 was able to achieve the minimum of 19 ms. The fact that all students were within a 300 ms window for the final three trials exhibits a high level of precision across all students.

In addition, 3 of the 7 students were able to achieve the learning goal 100% of the last three trials – Participant 1, 4, and 6. Here, retention is achieved because the students are able to mimic their initial movement perfectly. Another 3 of the 7 students were able to achieve the learning goal 67% of the last three trials – Participant

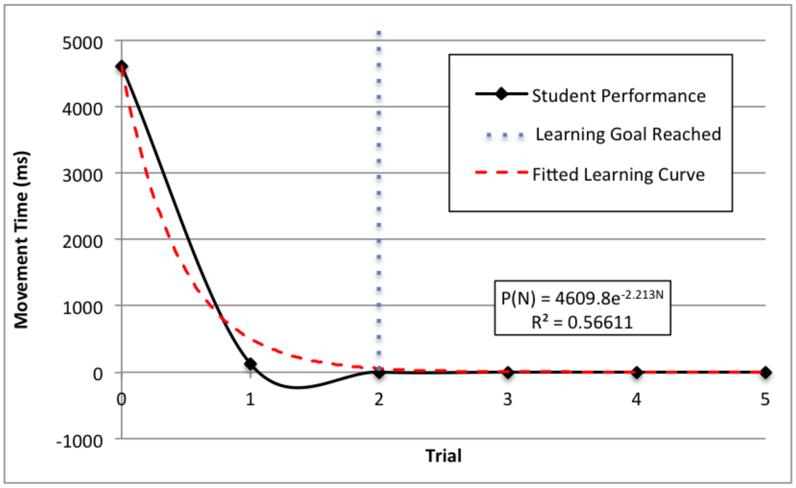


Figure 53: Participant 4's actual performance *P(MT)* plotted with his or her fitted learning curve *P(N)*.

2, 5, and 7. In the cases when the students did not achieve the goal, he or she was at most 38 ms away from the goal, which is negligible. Therefore, retention is achieved with these students as well. Lastly, 1 of the 7 students did not achieve the learning goal for the last three trials – Participant 3. This student was able to achieve the initial learning goal after his or her first attempt; however, this may have been due to "luck" and not necessarily learning/understanding. Furthermore, the individual may not have had enough practice trials to truly understand the learning goal before being asked to repeat the motor task.

As a short follow-up study, we had Participant 3 complete the learning cycle for a second time. These results showed an improvement in his overall performance. The individual was able to achieve the initial learning goal after one trial and improved his or her learning rate β from 1.081 to 1.845. The individual was also able to achieve the learning goal 67% of the last three trials. In the instance when he or she did not achieve the goal, it was by a negligible 49 ms. Therefore, retention is achieved for Participant 3, but only after two cycles of the system had been completed.

After looking at the results, a naive individual may say that Participant 1 is the

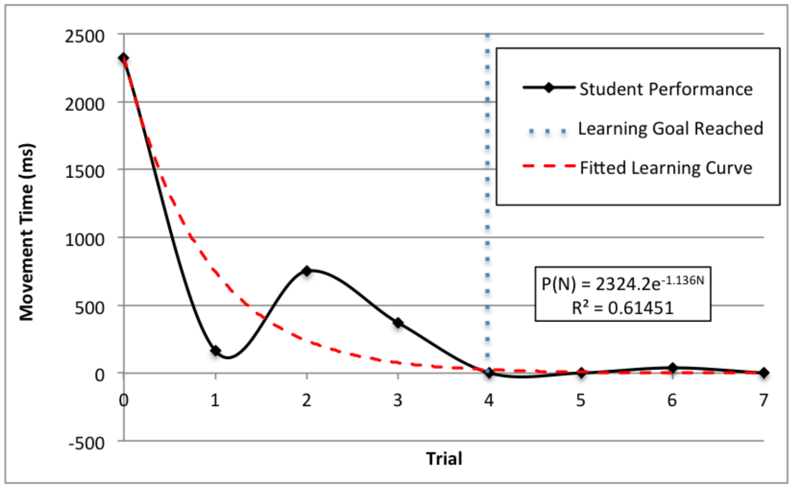


Figure 54: Participant 5's actual performance *P(MT)* plotted with his or her fitted learning curve *P(N)*.

"best" student being that he or she achieved learning after the first trial, retained this level of performance for 100% of the last 3 trials, had an ideal learning rate β of 2.092, and obtained a small window for the last three trials of 85 ms. This same naive individual may say that Participant 3 is the "worst" student being that he or she achieved learning after the first trial due to "luck," did not achieve the learning goal for the last three trials, had the lowest learning rate β of 1.081, and ultimately had to complete the learning cycle again to improve. This is by far not true. A major takeaway from this study is that students learn at different speeds and prefer variety of teaching styles. However, the REA system is still able to adapt to each student's learning preference so that he or she is able to reach the learning goal in a timeframe suitable to his or her individual needs.

7.6 Conclusion

The developed system was able provide a combination of feedback (motivation, instruction, correction) and repetition tactics throughout the learning process that enabled 100% of the students to reach their learning goals and 85.7% of the students

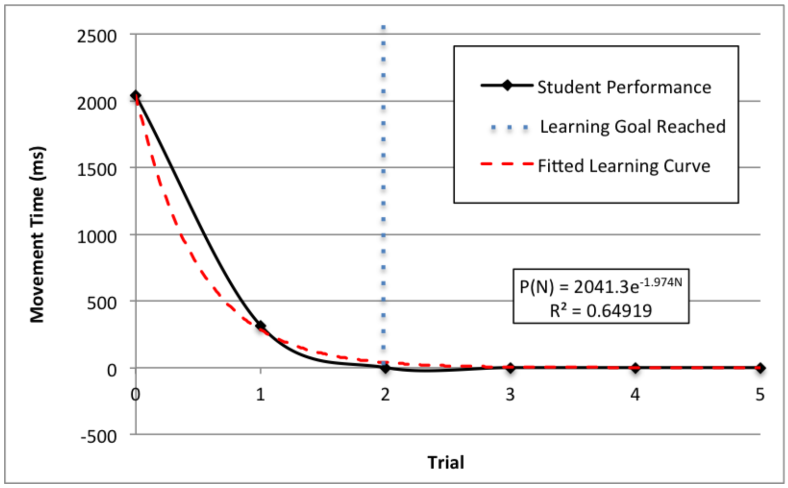


Figure 55: Participant 6's actual performance *P(MT)* plotted with his or her fitted learning curve *P(N)*.

to retain the information learned. In the one case where the student was unable to retain the information taught, the individual appeared to have achieved the initial learning goal through "luck," which unfortunately happens often in learning. This further shows the importance of evaluating learning to determine if retention is being achieved. This critical component of evaluating retention would allow an instructor to repeat the learning cycle so that all students are able to achieve optimal learning.

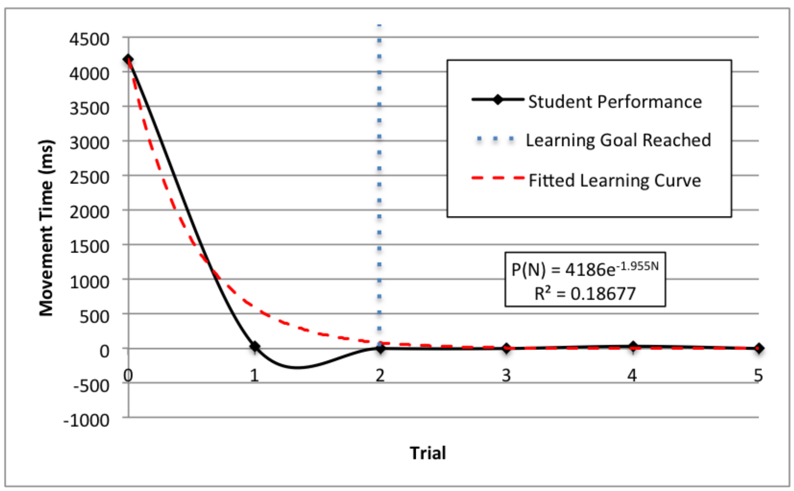


Figure 56: Participant 7's actual performance *P(MT)* plotted with his or her fitted learning curve *P(N)*.

CHAPTER VIII

CONCLUSION

In this dissertation, we aimed to develop a robotic educational agent that can interactively function in an equivalent manner as a human tutor. In result, we were able to develop an educational system that effectively monitors student engagement, applies behavioral strategies, teaches novel tasks, and improves student retention to achieve individualized learning.

8.1 Engagement Model

We begin by developing an engagement model based on the interactions between the student and the teaching device (tablet, computer, or virtual reality game), and we later expand the model by developing an eye gaze algorithm based on student fixations and saccades.

We observed that the eye gaze and head pose method of monitoring engagement as described in [9] worked well for low-level cognitive problems, but failed more often for the higher level problems. Because we want to target this system in the math domain, it is important that the engagement model is able monitor the student's attentiveness regardless of the problem difficulty. The preliminary engagement model takes time, response, and function execution into consideration, and it is able to successfully monitor engagement for all-level cognitive tasks.

The results of this study prove that we are able to predict salient points in an environment via eye gaze before an individual becomes fully cognizant of them. Furthermore, monitoring human-task inputs (measured by mouse and keyboard events and eye gaze) in a novel environment has the potential to enhance learning and understanding as well as engagement when completing a task.

8.2 Social Interaction Model

Next, we examined how emotions in social scenarios are able to enhance learning. Because empathy is a key factor used to enhance interpersonal relationships, which ultimately leads to increased enthusiasm and learning [86], we derived a gestural framework for implementing happy and sad emotions on a humanoid robotic platform [17].

This study revealed that by altering head direction, arm direction, gesture size, and gesture speed on a humanoid robotic social agent, participants are able to achieve accurate perception when the intended emotion is happy or sad. By using these key principles to categorize the gestures, the standard deviation was kept consistently at a minimum when identifying emotion. When using this framework, the participants were very confident in identifying when the intended emotion is happy, not happy, sad, not sad, and not neutral. This work suggests that engagement and motivation during social interaction can be optimized through the use of happy and sad gestures derived using the described framework.

In addition, studies have shown that the use of verbal encouragement strategies in education is able to maximize learning. This idea is derived from traditional classroom settings where teachers use a multitude of behavioral strategies to maintain the students level of engagement. Motivated by these educational practices, we developed a number of socially-supportive phrases to embed on the robotic educational agent.

8.3 Robotic Educational Agent

Next, we elaborate on the process of embedding social interaction within a humanoid-student *math*- and *motor*-learning scenario in order to re-engage students during high- and low-demand cognitive tasks, respectively. We found that across all interaction types, verbal, nonverbal, and both, the students enjoyed Darwin and were not distracted by his presence during the session. They were able to build a relationship with Darwin and did not wish to disappoint him with their performance. When compared to having no educational agent present, every interaction type that Darwin implemented was successfully able to maximize the time used in the learning environment. This was achieved by using the engagement model to monitor progression through the session. In addition, the groups that were provided motivational cues during the learning scenario had a richer experience. Based on their responses, they are more likely to interact with the system long-term, which is ideal for optimal learning and retention. Therefore, this works suggests that verbal engagement is ideal for enhancing performance with motor tasks.

After increasing student engagement with out system, we aimed to increase student learning/performance by integrating guided instruction and corrective feedback into the REA. The results of our study suggest that the proposed system can increase learning best by using both verbal and nonverbal cues for instructional feedback. In addition, the continuous corrective feedback embedded on the REA enabled 92% of the students to reach their targeted learning goals. The students' performance mimic that of an overdamped second-order system, which suggests that each new student that interacts with our system will have a similar response and reach their learning goal.

Lastly, in efforts to evaluate if learning is being retained after instruction has ended, we expanded the system to provide a combination of feedback (motivation, instruction, correction) and repetition tactics throughout the learning process. This overall system enabled 100% of the students to reach their learning goals and 85.7% of the students to retain the information learned. This further reiterates the importance of evaluating learning to ensure that retention is being achieved. In conclusion, our developed robotic educational agent effectively monitors engagement, applies behavioral strategies to increase motivation, teaches novel tasks, and improves student retention to achieve individualized learning.

APPENDIX A

EYE GAZE DATA

Table 28: Image 1 was the first image pair shown in the eye gaze study. The left image is the original image, and the right image has been manipulated. In particular, the mailbox has been moved from the right side of the road to the left side of the road. Of the 9 participants, 6 consciously noticed a change in the image, and their description of the change is listed below.

Participant	Details
3	Mailbox moved from the right side to the left side
4	The mailbox moved from one side of the road to the other
5	Mailbox moved from right to left
6	The mail box moved from the right portion of the screen to the left portion
7	I think the presence of the mailbox was the change
9	The mailbox changed sides

Table 29: Image 2 was the second image pair shown in the eye gaze study. The left image has been manipulated image, and the right is the original image. In the manipulated image, the star has been moved from the top of the left window to the top of the right window. Of the 9 participants, 6 consciously noticed a change in the image, and their description of the change is listed below.

Participant	Details
1	Yellow star moved windows
3	The yellow flower moved from the upper right wall to the upper left wall
4	The star decoration on the wall (over the left window) grew larger
6	The gold star moved from being above the right window to being above the left
7	I think one of the window doors being missing is the change
9	The golden flower moved to the left

Table 30: Image 3 was the third image pair shown in the eye gaze study. The left image is the original image, and the right image has been manipulated. In particular, the fish on the far left has been moved to the front-right side of the image. All 9 participants consciously noticed a change in the image, and their description of the change is listed below.

Participant	**Details**
1	Left fish moved
2	One of the fish moved from left to right (bottom left corner)
3	The orientation of the image changed. It shifted to the right by 45 degrees
4	One of the fish moved from the left side of the image to the foreground on the right
5	More fish were on the left and not the right
6	There was an additional fish on the left portion of the screen that disappeared
7	One of the fish being missing was the change
8	The fish changed positions from the left side to the right side
9	The fish moved from being grouped on the left to the right

Table 31: Image 4 was the fourth image pair shown in the eye gaze study. The left image has been manipulated image, and the right is the original image. In the manipulated image, the orange has been moved from the table on the left to the floor on the right next to the couch. All 9 participants consciously noticed a change in the image, and their description of the change is listed below.

Participant	Details
1	Orange moved from table to floor
2	Orange moved from the floor to on top of the table on the left of the room
3	The orange moved from the floor to the table. Also, the size of the orange was larger [at] first
4	The orange moved from the floor at the right of the image onto the table
5	The orange was on the floor
6	The orange moved from being on the floor on the right of the image to being on the table
7	There was an object on the floor that is no longer present
8	The pumpkin moved from the bottom right quadrant to the table
9	The orange moved from the floor to the table

Table 32: Image 5 was the fifth image pair shown in the eye gaze study. The left image is the original image, and the right image has been manipulated. In particular, the horse on the left of the bunch has been moved farther away from the other two horses (closer to bottom left). Of the 9 participants, 2 consciously noticed a change in the image, and their description of the change is listed below.

Participant	Details
4	One of the horses moved further away (to the left) of the other two.
5	All three horses were together

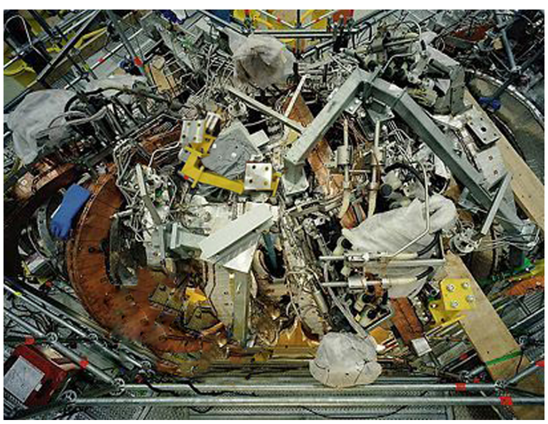
Table 33: Image 6 was the sixth image pair shown in the eye gaze study. The left image has been manipulated image, and the right is the original image. In the manipulated image, the book has been moved from the top-right of the image to the top-center of the image. Of the 9 participants, 7 consciously noticed a change in the image, and their description of the change is listed below.

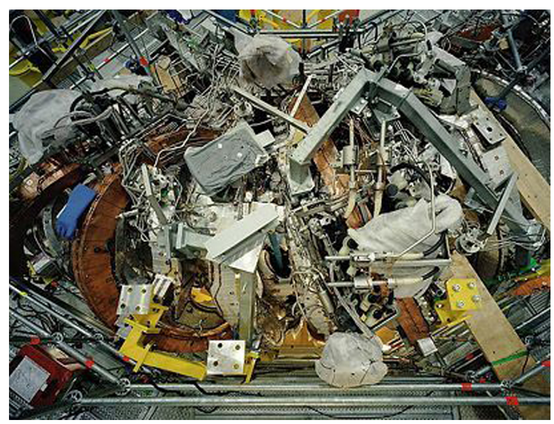
Participant	Details
1	Arm with book moved from center mountain to right mountain
3	The arm holding the book shifted from the center of the image to the right side
4	The hand coming down from the clouds moved from the cloud in the center of the sky to a cloud on the right side of the sky
5	The arm and book were in the middle of the picture
6	The arm that is coming from behind the cloud and holding the book moved from the center cloud to the cloud that is farthest to the right
7	I think the white sleeve with a book on the top right corner was a change
8	The arm on the upper right quadrant moved from the middle top quadrant to the right quadrant

Table 34: Image 7 was the seventh image pair shown in the eye gaze study. The left image is the original image, and the right image has been manipulated. In particular, the yellow flower pot on the right window's right shutter has been moved to the left window's left shutter. Of the 9 participants, 4 consciously noticed a change in the image, and their description of the change is listed below.

Participant	**Details**
2	The flower pot hanging from the rightmost shutter (with the yellow flowers) moved to the leftmost shutter (under the pink flowers)
3	The flower pot with yellow flowers shifted from the upper portion of the right window shade to the lower portion of the left window shade
4	I'm not 100% sure, but I think there were flowers (or something) on all of the window shutters, not just those on the left window
5	Flower placement on shutters

Table 35: Image 8 was the eighth image pair shown in the eye gaze study. The left image has been manipulated image, and the right is the original image. In the manipulated image, the yellow L-shaped figure in the botttom-left of the image has been moved up to the center-left of the image. Of the 9 participants, 3 consciously noticed a change in the image, and their description of the change is listed below.





Participant	**Details**
2	The yellow L-shaped part moved from the center to the bottom left corner
7	I think there was another yellow object, but I do not recall precisely
9	The yellow pieces moved from the center to the bottom center and to the right

Table 36: Image 9 was the ninth image pair shown in the eye gaze study. The left image is the original image, and the right image has been manipulated. In particular, the red person sitting on the rock (center-right) has been moved to the bridge in the center of the image. Of the 9 participants, 8 consciously noticed a change in the image, and their description of the change is listed below.

Participant	**Details**
1	Red person appeared on bridge
2	The man sitting at the edge of one of the rocks on the right is now on the bridge
3	The person moved from sitting on the river bank to sitting on the bridge
4	The man moved onto the bridge. I don't remember where he was previously - I think on the rocks off to the right of the water midway up the image
5	The man was on the bank and not the bridge
6	The person who is wearing the red shirt moved from the ledge under the tree that is located in approximately the center of the image to the bridge
8	That red kid in the middle wasn't there before
9	The boy moved from sitting beneath the tree to on the bridge

Table 37: Image 10 was the tenth image pair shown in the eye gaze study. The left image has been manipulated image, and the right is the original image. In the manipulated image, the sailboat has been moved from the left of the lake to the center of the lake. Of the 9 participants, 5 consciously noticed a change in the image, and their description of the change is listed below.

Participant	**Details**
2	Two sailboats moved from the center of the lake to the left
4	The group of boats moved from the middle of the water to the left side
5	All the ships were near each other in the middle of the picture
6	The two sail boats moved from the center on the image to the left side of the image
9	Yes, two of the boats moved from where the lone sailboat is to the front of the image

Table 38: Image 11 was the eleventh image pair shown in the eye gaze study. The left image is the original image, and the right image has been manipulated. In particular, the blue house in the center of the image has been moved to the top left of the image (on top of the big white house). None of the participants consciously noticed a change in the image.

Participant	Details
1-9	None of the participants noticed the change here.

Table 39: Image 12 was the twelfth image pair shown in the eye gaze study. The left image has been manipulated image, and the right is the original image. In the manipulated image, the house debris located on the left side of the image has been moved to the right side of the image. Of the 9 participants, 3 consciously noticed a change in the image, and their description of the change is listed below.

Participant	Details
2	The remains of one of the buildings moved from the right of the image (on top of a hill of rocks) to the left (on top of a hill of rocks)
3	The partially demolished building in the foreground of the picture shifted from the right side of the image to the left. The entire picture was inverted
4	I think there was smoke coming off of the building on the left side of the image before

REFERENCES

[1] "Boardmaker plus," *http://www.mayer-johnson.com/board maker-plus-v-6/*.

[2] "Criterion-referenced competency tests (crct)," *http://www. doe.k12.ga.us/Curriculum-Instruction-and-Assessment/Assessment/Pages/.aspx*.

[3] "The eye tribe," *https://theeyetribe.com/*.

[4] "How teachers learn to engage students in active learning." National Center for Research on Teacher Learning, East Lansing, MSU, 1993.

[5] "Exponential rise and fall to limit." Alan Fletcher, 2013.

[6] "Power law." Alan Fletcher, 2013.

[7] Aleven, A., "An effective metacognitive strategy: learning by doing and explaining with a computer-based cognitive tutor," *Cognitive Science*, pp. 147–179, April 2002.

[8] Anderson, J., Corbett, A., Koedinger, K., and Pelletier, R., "Cognitive tutors: Lessons learned," *The Journal of the Learning Sciences*, vol. 4, pp. 167–207, 1995.

[9] Asteriadis, S., Tzouveli, P., Karpouzis, K., and Kollias, S., "Estimation of behavioral user state based on eye gaze and head pose—application in an e-learning environment," in *Multimedia Tools and Applications*, vol. 41, pp. 469–493, February 2008.

[10] Bainbridge, W. A., Hart, J. W., Kim, E. S., and Scassellati, B., "The benefits of interactions with pysically present robots over video-displayed agents," *International Journal of Social Robotics*, vol. 3, no. 1, pp. 41–52, 2011.

[11] Baker, R., D'Mello, S., Mercedes, M., Rodrigo, T., and Graesser, A., "Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive–affective states during interactions with three different computer-based learning environments," *International Journal of Human-Computer Studies*, pp. 223–241, 2010.

[12] Beck, A., Canamero, L., Hiolle, A., Damiano, L., Cosi, P., Tesser, F., and Sommavilla, G., "Interpretation of emotional body language displayed by a humanoid robot: A case study with children," *International Journal of Social Robotics*, vol. 5, pp. 325–334, August 2013.

[13] Berka, C., Levendowski, D. J., Lumicao, M. N., Yau, A., Davis, G., Zivkovic, V. T., Olmstead, R. E., Tremoulet, P. D., and Craven, P. L., "Eeg correlates of task engagement and mental workload in vigilance, learning, and memory tasks," *Aviation, Space, and Environmental Medicine*, vol. 78, pp. B231–B244, May 2007.

[14] Breazeal, C., "Emotive qualities in robot speech," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, vol. 3, (Maui, HI, USA), pp. 1388–1394, October 2001.

[15] Brooks, D. and Howard, A., "Quantifying upper-arm rehabilitation metrics for children through interaction with a humanoid robot," in *Applied Bionics and Biomechanics*, vol. 9, pp. 152–172, 2012.

[16] Brown, L. and Howard, A., "Assessment of engagement for intelligent educational agents: A pilot study with middle school students," *ASEE Computers in Education (CoED)*, vol. 24, October-December 2014.

[17] Brown, L. and Howard, A., "Gestural behavioral implementation on a humanoid robotic platform for effective social interaction," in *IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, vol. 23, (Edinburgh, Scotland, UK), pp. 471–476, August 2014.

[18] Brown, L. and Howard, A., "A real-time model to assess student engagement for intelligent educational agents," in *ASEE Annual Conference*, (Indianapolis, IN, USA), pp. 24.95.1–24.95.11, June 2014.

[19] Brown, L. and Howard, A., "The effects of motivational feedback during motor-task learning using a social robot." 2015.

[20] Brown, L., Kerwin, R., and Howard, A., "Applying behavioral strategies for student engagement using a robotic educational agent," in *IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, (Manchester, England), pp. 4360–4365, October 2013.

[21] Brown, L., García-Vergara, S., and Howard, A. M., "Evaluating the effect of robot feedback on motor skill performance in therapy games," in *IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, (Kowloon Tong, Hong Kong), October 2015.

[22] Brown, L. and Howard, A., "Engaging children in math education using a socially interactive humanoid robot," in *IEEE-RAS International Conference on Humanoid Robots (Humanoids)*, (Atlanta, GA, USA), pp. 183–188, October 2013.

[23] Brown, L. N. and Howard, A. M., "The positive effects of verbal encouragement in mathematics education using a social robot," in *IEEE Integrated STEM Education Conference (ISEC)*, (Princeton, NJ, USA), March 2014.

[24] Brown, P. and Levinson, S., *Politeness.* Cambridge University Press, 1998.

[25] Brütsch, K., Schuler, T., Koenig, A., Zimmerli, L., Koeneke, S. M., Lünenburger, L., Riener, R., Jäncke, L., and Meyer-Heim, A., "Influence of virtual reality soccer game on walking performance in robotic assisted gait training for children," *Journal of NeuroEngineering and Rehabilitation*, 2010.

[26] Burger, B. and Bresin, R., "Communication of musical expression by means of mobile robot gestures," *Journal of Multimodal User Interfaces*, 2013.

[27] Cahn, J., "Generating expression in synthesized speech," Master's thesis, MIT Media Lab, 1990.

[28] Chang, J.-J., Yang, Y.-S., Wu, W.-L., Guo, L.-Y., Su, F.-C., and et al., "The constructs of kinematic measures for reaching performance in stroke patients," *Journal of Medical and Biological Engineering*, vol. 28, no. 2, pp. 65–70, 2008.

[29] Chen, Y., Kang, L., Chuang, T., Doong, J., Lee, S., Sai, M. T., Jeng, S., and Sung, W., "Use of virtual reality to improve upper-extremity control in children with cerebral palsy: a single subject design," *Journal of the American Physical Therapy Association*, 2007.

[30] Christophel, D. M., "The relationships among teacher immediacy behaviors, student motivation, and learning," *Communication Education*, vol. 39, pp. 323–340, 1990.

[31] Cioi, D., Kale, A., Burdea, G., Engsberg, J., Janes, W., and Ross, S., "Ankle control and strength training for children with cerebral palsy using the rutgers ankle cp: A case study," in *IEEE International Conference Rehabilitation Robot*, 2011.

[32] Colombo, R., Pisano, F., Mazzone, A., Delconte, C., Micera, S., Carrozza, M., Dario, P., and Minuco, G., "Design strategies to improve patient motivation during robot-aided rehabilitation," *Journal of NeuroEngineering and Rehabilitation*, 2006.

[33] Dahl, S. and Friberg, A., "Visual perception of expressiveness in musicians' body movement," 2007.

[34] Fitts, P. M., "Perceptual-motor skill learning," *Categories of human learning*, vol. 47, pp. 381–391, 1964.

[35] Fredricks, J. A., Blumenfeld, P. C., and Paris, A. H., "School engagement: Potential of the concept, state of the evidence," *Review of Educational Research*, January 2004.

[36] Gajdosik, R. L. and Bohannon, R. W., "Clinical measurement of range of motion review of goniometry emphasizing reliability and validity," *Journal of the American Physical Therapy Association*, vol. 67, no. 12, pp. 1867–1872, 1987.

[37] Garcia-Vergara, S., Brown, L., and Howard, A., "Increasing the efficacy of rehabilitation for children via a robotic playmate providing real-time low-resolution corrective feedback." [not published].

[38] Garcia-Vergara, S., Brown, L., Park, H. W., and Howard, A., *Technologies of Inclusive Well-Being: Serious Games, Alternative Realities, and Play Therapy*, ch. Engaging Children in Play Therapy: The Coupling of Virtual Reality Games with Social Robotics. Springer, 2014.

[39] García-Vergara, S., Chen, Y.-P., and Howard, A. M., "Super pop vrtm: An adaptable virtual reality game for upper-body rehabilitation," in *Virtual, Augmented and Mixed Reality. Systems and Applications*, pp. 40–49, Springer, 2013.

[40] García-Vergara, S. and Howard, A. M., "Three-dimensional fitt's law model used to predict movement time in serious games for rehabilitation," in *Virtual, Augmented and Mixed Reality. Applications of Virtual and Augmented Reality*, pp. 287–297, Springer, 2014.

[41] Gielniak, M. J. and Thomaz, A. L., "Enhancing interaction through exaggerated motion synthesis," in *ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2012.

[42] Gorham, J., "The relationship between verbal teacher immediacy behaviors and student learning," *Communication Education*, vol. 37, 1988.

[43] Granger, C. V., Hamilton, B. B., and Sherwin, F. S., "Guide for the use of the uniform data set for medical rehabilitation," *Uniform data system for medical rehabilitation project office, Buffalo General Hospital, New York*, vol. 14203, 1986.

[44] Greenwood, C. R., Horton, B. T., and Utley, C. A., "Academic engagement: Current perspectives on research and practice," *School Psychology Review*, 2002.

[45] Ha, I., Tamura, Y., Asama, H., Han, J., and Hong, D., "Development of open humanoid platform darwin-op," in *SICE Annual Conference*, pp. 2178–2181, 2011.

[46] Han, J., Jo, M., Park, S., and Kim, S., "The educational use of home robots for children," in *IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 2005.

[47] Howard, A. and Paul, W., "A 3d virtual environment for exploratory learning in mobile robot control," in *IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2005.

[48] Howard, A., "Robots learn to play: Robots emerging role in pediatric therapy.," in *International Florida Artificial Intelligence Research Society Conference*, 2013.

[49] Huang, C.-M. and Mutlu, B., "Robot behavior toolkit: Generating effective social behaviors for robots," in *ACM/IEEE Conference on Human-Robot Interaction (HRI)*, 2012.

[50] Isbister, K., Hook, K., Sharp, M., and Laaksolahti, J., "The sensual evaluation instrument: developing an affective evaluation tool," in *ACM Conference on Human Factors in Computing Systems (CHI)*, 2006.

[51] Janssen, J., Wal, C., Neerincx, M., and Looije, R., "Motivating children to learn arithmetic with an adaptive robot game," in *International Conference on Social Robotics (ICSR)*, pp. 153–162, 2011.

[52] Kidd, C. and Breazeal, C., "Robots at home: Understanding long-term human-robot interaction," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 3230–3235, 2008.

[53] Kidd, C. D. and Breazeal, C., "Sociable robot systems for real-world problems," in *IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 2005.

[54] Kory, J. and Breazeal, C., "Storytelling with robots: Learning companions for preschool children's language development," in *IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pp. 643–648, August 2014.

[55] Kulik, C.-L. C. and Kulik, J. A., "Effectiveness of computer-based instruction: An updated analysis," *Computers in Human Behavior*, vol. 7, pp. 75–94, 1991.

[56] Li, J. and Chignell, M., "Communication of emotion in social robots through simple head and arm movements," *International Journal of Social Robotics*, 2013.

[57] Looije, R., Neerincx, M. A., and de Lange, V., "Children's responses and opinion on three bots that motivate, educate and play," *Journal of Physical Agents*, June 2008.

[58] Loureiro, R., Amirabdollahian, F., Topping, M., Driessen, B., and Harwin, W., "Upper limb robot mediated stroke therapy gentle/s approach," *Autonomous Robots*, 2003.

[59] Lowe, J., "Computer-based education: Is it a panacea," *Journal of Research on Technology in Education*, vol. 34, pp. 163–171, 2002.

[60] McCrea, P. H., Eng, J. J., and Hodgson, A. J., "Biomechanics of reaching: clinical implications for individuals with acquired brain injury," *Disability and Rehabilitation*, vol. 24, no. 10, pp. 534–541, 2002.

[61] Michaud, F., Salter, T., Duquette, A., Mercier, H., Lauria, M., Larouche, H., and Larose, F., "Mobile robots engaging children in learning," 2007.

[62] Mora, R., "School is so boring: High-stakes testing and boredom at an urban middle school," *Perspectives on Urban Education*, vol. 9, no. 1, 2011.

[63] Morris, M. E., "Movement disorders in people with parkinson disease: a model for physical therapy," *Journal of the American Physical Therapy Association*, vol. 80, no. 6, pp. 578–597, 2000.

[64] Murray, L. and Arnott, L., "Toward the simulation of emotion in synthetic speech: a review of literature on human vocal emotion," *Journal Acustical Society of America*, vol. 93, no. 2, pp. 1097–1108, 1993.

[65] Mutlu, B., "Designing embodied cues for dialog with robots," *Artificial Intelligence Magazine*, vol. 32, no. 4, pp. 17–30, 2012.

[66] of Education, U. D., "The best states to teach in america," *http://online.seu.edu/the-best-states-to-teach-in-america/*.

[67] of Labor, U. B., "Supply factor: The allied health professional shortage," *http://kidmedic.weebly.com/allied-health-shortages.html*.

[68] Park, E., Kim, K., and Pobil, A., "The effects of a robot instructor's positive vs. negative feedbacks on attraction and acceptance towards the robot in classroom," in *International Conference on Social Robotics (ICSR)*, pp. 135–141, 2011.

[69] Park, H. W. and Howard, A., "Case-based reasoning for planning turn-taking strategy with a therapeutic robot playmate," in *IEEE International Conference on Biomedical Robotics and Biomechatronics*, 2010.

[70] Picard, R., *Affective Computation*. MIT Press, 1997.

[71] Posner, M. I., Snyder, C. R. R., and Davidson, B. J., "Attention and detection of signals," *Journal of Experimental Psychology: General*, 1980.

[72] Powers, A., Kiesler, S., Fussell, S., and Torre, C., "Comparing a computer agent with a humanoid robot," in *ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pp. 145–152, March 2007.

[73] Remy, S. and Howard, A., "Predicting the robot learning curve based on properties of human interaction," in *AAAI Spring Symposium*, 2009.

[74] Richmond, V. P., McCroskey, J. C., and Johnson, A. D., "Development of the nonverbal immediacy scale (nis): Measures of self-and other perceived nonverbal immediacy," *Communication Quarterly*, vol. 51, no. 4, pp. 504–517, 2003.

[75] Ritter, F. E. and Schooler, L. J., "The learning curve," *International encyclopedia of the social and behavioral sciences*, 2002.

[76] Rohrer, B., Fasoli, S., Krebs, H. I., Hughes, R., Volpe, B., Frontera, W. R., Stein, J., and Hogan, N., "Movement smoothness changes during stroke recovery," *The Journal of Neuroscience*, vol. 22, no. 18, pp. 8297–8304, 2002.

[77] Russell, J. A., "A circumplex model of affect," *Journal of Personality and Social Psychology*, vol. 39, pp. 1161–1178, 1980.

[78] Saerbeck, M., Schut, T., Bartneck, C., and Janse, M., "Expressive robots in education: Varying the degree of social supportive behavior of a robotic tutor," in *ACM Conference on Human Factors in Computing Systems (CHI)*, (Atlanta, GA), 2010.

[79] Sanders and Rivers, "Cumulative and residual effects on future student academic achievement," *Boston Public Schools - McKinsey analysis*, 1998.

[80] Schegloff, E., "Body torque," *Soc Res*, 1998.

[81] Schofield, J. W., *In Computers and Classroom Culture*. Cambridge Univ. Press, 1995.

[82] Shah, J., Wiken, J., and Williams, B., "Improved human-robot team performance using chaski, a human-inspired plan execution," in *ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2011.

[83] Smith, C. N., Hopkins, R. O., and Squire, L. R., "Experience-dependent eye movements, awareness, and hippocampus-dependent memory," *The Journal of Neuroscience*, 2006.

[84] Szafir, D. and Mutlu, B., "Pay attention! designing adaptive agents that monitor and improve user engagement," in *ACM Conference on Human Factors in Computing Systems (CHI)*, May 2012.

[85] Taggart, W., Turkle, S., and Kidd, C. D., "An interactive robot in a nursing home," in *Toward Social Mechanisms of Android Science*, 2005.

[86] Tiberius, R. and Billson, J., "The social context of teaching and learning," in *New Directions for Teaching and Learning*, 1991.

[87] Wainer, H., Dorans, N. J., Flaugher, R., Green, B. F., Mislevy, R. J., Steinberg, L., and Thissen, D., *Computerized Adaptive Testing: A Primer*. Lawrence Erlbaum, 2000.

[88] Wood, D., Bruner, J. S., and Ross, G., "The role of tutoring in problem solving," *Journal of Child Psychology and Psychiatry*, pp. 89–100, 1976.

[89] Wood, K. C., Lathan, C. E., and Kaufman, K. R., "Feasibility of gestural feedback treatment for upper extremity movement in children with cerebral palsy," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 21, no. 2, pp. 300–305, 2013.

[90] Yun, S., Shin, J., Kim, D., Kim, C. G., Kim, M., and Choi, M.-T., "Engkey: Tele-education robot," in *International Conference on Social Robotics (ICSR)*, pp. 142–152, 2011.

[91] Zheng, Z., Das, S., Young, E. M., Swanson, A., Warren, Z., and Sarkar, N., "Autonomous robot-mediated imitation learning for children with autism," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2014.