Personalized Context-aware Multimodal Robot Feedback

Roshni Kaushik

CMU-RI-TR-25-03

February 24, 2025

The Robotics Institute
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA

Thesis Committee:
Reid Simmons, *Co-chair, Carnegie Mellon University*
Henny Admoni, *Carnegie Mellon University*
Jean Oh, *Carnegie Mellon University*
Sonia Chernova, *Georgia Institute of Technology*

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Robotics.

Abstract

In the field of human-robot interaction (HRI), integration of robots into social settings, such as healthcare and education, is gaining traction. Robots that provide individualized support to improve human performance and subjective experience will generally be more successful in these domains. Robots should personalize their interactions, be aware of the contextual nuances surrounding their behavior, and effectively understand and generate nonverbal cues (as humans' perceptions and responses are heavily influenced by nonverbal behavior). They should also consider factors such as personality traits, the physical environment, and emotional states to provide tailored, context-aware assistance and support during the interactions. This thesis explores personalized context-aware multimodal robot feedback, focusing on affective nonverbal behavior.

We first consider the problem of estimating *context*, specifically modeling key aspects of the human state. We predict engagement-related events in an educational activity before the end of that activity, which could allow the robot to provide feedback early enough to improve the human's experience. We then explore generating *nonverbal affective* robot behavior by correlating a simulated robot's movements with displayed emotion. We develop a user study to show that matching the robot's conveyed emotion with a matching affective movement has a positive impact on the human's performance in a sorting game. Next we design a physical robot exercise coach as a platform where we can estimate context (exercise performance, fatigue level, etc.). With a user study, we examine the changes in human perception of and performances with different robot feedback styles. This provides us a basis on which to begin tailoring feedback styles to the individual. Finally, we develop a *personalized context-aware* robot using a contextual bandit approach to dynamically adapt the robot's feedback style to optimize the human's performance, learning over time which style to use and when. This brings together all the work presented in this thesis and aims to create a holistic framework for generating personalized context-aware multimodal feedback that positively impacts the interaction with the human.

Acknowledgments

I want to thank the many people who made this work possible. First and foremost, to my advisor Reid Simmons. I could not have asked for a better advisor and have benefited greatly from your expertise, insight, and kindness. You truly care about your students both as researchers and as people. I value our research and plant chats and feel that I have become a better researcher because of you.

I also want to thank the other members of my committee, Henny Admoni, Jean Oh, and Sonia Chernova. Henny and Jean, you have provided feedback on my researcher all the way from my qualifiers, and your suggestions and questions during our research chats have been invaluable in shaping my research. Sonia, thank you for providing mentoring all the way back in 2016 when I did a summer program at Georgia Tech with you, and that summer was my inspiration to go to pursue robotics research in graduate school. I am very grateful that you were able to be part of my committee and that I could benefit from your expertise during this work. I also want to thank my masters advisor at University of Illinois at Urbana-Champaign, Amy LaViers, who really helped me find my footing in graduate school and my research interest in human-robot interaction.

I owe a great deal to the collaborators I am lucky to have had during the last several years. I learned a lot from the students who came before me, Pallavi, Mike, Reuben, Sam, and Zhi. Thank you also to Rayna, I am glad that became friends. Running a user study together is truly a bonding experience. I was also very lucky to have a large number of undergraduate students working with me and I truly value their contributions. Devika Venugopalan, Vera Zhong, Sara Pavlinek, Audrey May, Tarun Trivedi, Mahati Manda, Kevin Ni, Eshaan Joshi, Jaehee Kim, Peter Vanegas, Allison Moore, Mark Chen, Priyanshi Garg, Bharath Sreenivas, Steven Qu, Adrian Thinnyun.

I also need to thank my friends and colleagues at CMU. Terri, Michelle, Suresh, Pragathi, Pat, Gavin, and others, thank you for all the good times and for all your research feedback throughout the years. And finally, thanks to the admins who made all the logistics a breeze: Suzanne Muth, Peggy Martin, and Lynnetta Miller.

And lastly, to my family and Bill. I would not be here without you. You are my greatest support, and I am thankful for you every day.

Funding

This work was supported by NSF Grant (Teachable Robots) #IIS-1939047 and NSF Grant (AI-CARING) #2112633

Contents

1 Introduction **1**

2 Related Work **5**
2.1 Human Feedback . . . 5
2.2 Generating Robot Feedback . . . 6
2.3 Robot Feedback Improving Performance . . . 7
2.4 Robot Feedback Improving Experience . . . 8
2.5 Personalization and Context . . . 8
2.6 Summary . . . 9

3 Approach **11**
3.1 Estimating Context . . . 12
3.2 Generating Robot Feedback . . . 13
3.3 Developing Feedback Styles . . . 15
3.4 Personalizing Feedback . . . 15

4 Human State Estimation **17**
4.1 Problem Formulation . . . 17
4.2 Robotutor Dataset . . . 18
4.2.1 Automatic Feature Extraction . . . 19
4.2.2 Description of Dataset . . . 20
4.3 Methodology . . . 21
4.3.1 Training Gaussian Mixture Models . . . 21
4.3.2 Probability Propagation . . . 22
4.3.3 Classification . . . 22
4.3.4 Optimizing Performance . . . 24
4.4 Results . . . 24
4.4.1 Earliness and Guessing . . . 25
4.4.2 Facial and Task Features . . . 26
4.5 Discussion . . . 27

5 Affective Nonverbal Feedback **29**
5.1 Generating Affective Movements . . . 30
5.1.1 Robot Movements . . . 30

5.1.2 Movement Generation User Study 32
5.1.3 Movement Generation Discussion 35
5.2 Sorting Game Task Design 37
5.2.1 Example Rule Inference Process 38
5.2.2 Choosing a Card Order 39
5.2.3 Learner Model 41
5.3 Affective Nonverbal Feedback User Study 42
5.3.1 Affective Nonverbal Feedback Study Design 42
5.3.2 Experimental Design 45
5.3.3 Affective Nonverbal Feedback Results 46
5.3.4 Affective Nonverbal Feedback Discussion 50

6 Feedback Style Preferences 52
6.1 Exercise Evaluation 52
6.1.1 Segmentation 54
6.1.2 Comparison to References 55
6.2 Feedback Generation 58
6.2.1 Verbal 59
6.2.2 Face 60
6.2.3 Body 61
6.3 Feedback Cadence 62
6.3.1 Feedback Cadence User Study 63
6.3.2 Feedback Cadence Study Discussion 67
6.4 Feedback Style Study Design 69
6.4.1 Study Procedures 69
6.5 Feedback Style Results 71
6.5.1 Performance 72
6.5.2 Perception of the Robot 73
6.6 Feedback Style Discussion 74

7 Personalized Feedback Style 76
7.1 Adaptive Feedback Model 76
7.1.1 Contextual Bandit Simulation Results 80
7.2 Analysis of Dataset 85
7.2.1 Performance and Preference Groups 87
7.2.2 Model Results 89
7.3 Adaptive Study Design 94
7.3.1 Study Procedures 96
7.4 Adaptive Study Results 97
7.4.1 Performance vs. Ratings vs. Stated Preference 97
7.4.2 Stated Style vs. Adaptive Performance 99

7.4.3 Performance per Performance Group 100
7.4.4 Performance by Context 102
7.5 Adaptive Study Discussion 103

8 Conclusion and Future Directions 108
8.1 Key Attributes of Applicable Problems and Design Considerations . . 108
8.2 Verbal and nonverbal feedback 110
8.3 Balancing Competing Objectives 111
8.4 Limitations and extensions of the contextual bandit formulation for an exercise coach 112
8.5 Other Future Directions 113
8.6 Conclusion 114

A Appendix 115
A.1 Study Form 115
A.1.1 Consent Form for Participation in Research 115
A.1.2 Demographics Form 118
A.1.3 Survey completed after each round of exercise 119
A.2 LLM Prompt for Generating Robot Verbal Feedback 120

Bibliography 125

When this dissertation is viewed as a PDF, the page header is a link to this Table of Contents.

List of Figures

3.1 **Left**: The human performs the task and responds to the robot's feedback based on their own intrinsics. **Middle**: The robot observes the human's response and estimates the context. **Right**: The robot uses its context estimation and its own intrinsics to generate a multimodal response. **Goal**: The overall goal is to personalize the robot's intrinsics to improve specific outcomes. 12
3.2 The humanoid robot Quori with a leaning torso, gesturing arms, and face that can project any image. 14

4.1 **Left**: Single frame from recorded videos of the RoboTutor application with front-facing camera feed (face obscured for anonymity). **Right**: Feedback screen appearing at the end of each activity. 19
4.2 Example predicting *feedback* choice using the described approach with two possible thresholds λ_1 and λ_2 shown. 23
4.3 Comparison of $\alpha = 1$, $\alpha = 0.8$, and guessing the most common label, for *feedback* (left) and *backbutton* (right). The average value of metrics F1 score, earliness, frequency, and S are illustrated with a standard deviation bar. Significant differences at the $p = 0.05$ level are indicated with asterisks. Significance was tested only for F1 score and earliness. 25
4.4 Comparison of using only facial features, only task features, and all features, for *feedback* (left) and *backbutton* (right), with the optimal hyperparameters and $\alpha = 0.8$. The average value of metrics F1 score, earliness, frequency, and S are illustrated with a standard deviation bar. Significant differences at the $p = 0.05$ level are indicated with asterisks. Significance was tested only for F1 score and earliness. . . . 26

5.1 Torso and left arm angles labeled on the Quori robot (image from http://www.quori.org/) 31
5.2 Each participant saw 10 randomly chosen videos out 224 and answered a set of Likert questions about their perception of emotion in the robot movement. 34
5.3 All 81 cards (left) in the sorting game (right) belong in one of two bins. 37
5.4 Example of how a human could infer the rule that states "all green cards belong in the left bin, and all others in the right bin" 38

5.5 Example screenshots of the *neutral* and *matching affective* conditions when the correct bin is the left bin. 44
5.6 In the trial phase, the participant dragged the card from the gray area to one of the two bins. After clicking the “Submit Choice” button, the robot provided feedback on their choice and moved the card to the correct bin, if it was placed incorrectly. 45
5.7 Participant performance compared to simulated learners for the *easy* and *difficult* rule. Demonstration cards and trial cards are shown, and the title of each plot indicates the rule being learned. 46
5.8 Two-way ANOVA tests with significant results indicated. The asterisks indicate a significance of $p < 0.05$. 48

6.1 Bicep curls (left) and lateral raises (right) 53
6.2 Angles computed during bicep curls 54
6.3 One of the bicep curl angles segmented into repetitions with the alternating solid blue and dashed green colors indicating where the segmentation method found the end of one rep and the beginning of another 55
6.4 An illustration of happy (top row) and sad (bottom row) at two different levels of intensity: 50% (left column) and 100% (right column). 61
6.5 Screenshot for a video from the study with the human exerciser on the left and the robot providing verbal and nonverbal feedback on the right. 64
6.6 Results of all Likert-style measures with perceptions of verbal feedback in the first row and nonverbal feedback in the second row. Significant differences are marked with one asterisk ($p < 0.05$) or two asterisks ($p < 0.01$). Each box plot shows the interquartile range of the data, as well as the min and max of the data. The mean is shown with a green triangle. 66
6.7 Study setup with participant doing bicep curls 70
6.8 Performance measures compared between the three different styles. Statistical differences are shown with one asterisk ($p < 0.05$) and two asterisks ($p < 0.01$). 72
6.9 Subjective preferences based on Likert-style questions, scored 1-5. Values above 3 are a preference for **encouraging**), and values below 3 are a preference for **firm**. Statistical differences from the center value are shown with one asterisk ($p < 0.05$). 74

7.1 The robot observes the context c from the human and chooses an action a. The human responds to the robot’s action, and their performance r forms the reward for the robot to train policy Π. 77

7.2 The average reward received shown in blue for each experiment (over 30 trials) with one standard deviation shaded in light blue. The average reward per set increases as the bandit sees more data, with the 20th set closer to the expected optimal reward (shown in red) compared to the 1st set. 81
7.3 Experiment 1: Simulation results with no human fatigue dependence in performance. The human's performance is the same across all fatigue levels, and the optimal action for the bandit is **very firm** across all reps. 82
7.4 Experiment 2: Simulation results with basic human fatigue dependence in performance. The human's performance decreases as fatigue increases, but the optimal action for the bandit is **very firm** across all reps. 83
7.5 Experiment 3: Simulation results with complex fatigue dependence in performance. The bandit should choose **very firm** for low fatigue, is relatively indifferent between all styles for moderate fatigue, and should choose **encouraging** or **very encouraging** for high fatigue. . 84
7.6 Comparison of average reward for all three fatigue situations with a simple bandit and a contextual bandit. Significant differences are indicated with $***$ ($p < 0.001$), and the mean of each distribution is also marked. 85
7.7 Example of model results for one participant. This participant performed much better with the **encouraging** style (37.5% with **firm**, 85.7% with **encouraging**), which the model quickly learns by choosing the **encouraging** style 96.7% of the time. 91
7.8 Expected bandit reward - actual reward for all 29 participants (x-axis is participant sorted by difference). The bandit outperforms the true data for most participants. 92
7.9 Example of model results for one participant, who performed approximately equally with the two styles (52.6% with **firm** and 68.4% with **encouraging**). The model does not learn their preferences as quickly since their performance with the two styles is quite close. 93
7.10 Expected contextual bandit reward - expected simple bandit reward for all 29 participants (x-axis is participant sorted by difference). . . . 94
7.11 The style the participant performs best with, rates the highest, and states as their preference before exercising, separated out by younger and older adults. **Firm** is shown in maroon, **Encouraging** in green, and neither/no preference in white. 97
7.12 Performance comparison with stated style preference and with the **adaptive** style for the 14 participants who chose either **firm** or **encouraging** for their stated preference at the beginning of the session. 100

7.13 Performance comparison by performance group for each style. Statistical differences are marked at the $p < 0.05$ and $p < 0.01$ level with $*$ and $**$ respectively, and groups that are similar at a 95% confidence level are marked as "same" 101
7.14 The style the participant performs best with across all fatigues as well as by fatigue level, separated out by younger and older adults. **Firm** is shown in maroon, **encouraging** in green, and neither/no performance difference in white. 102
7.15 Pie chart indicating the percentage of people that matched across 3 groups: style they performed best with, style they rated the highest, and style they stated as their preference. This chart aggregates the information shown in Figure 7.2.1 104
7.16 Pie chart indicating the percentage of people that match each of the three contextual situations explored in simulation: no fatigue dependence, basic fatigue dependence, and complex fatigue dependence. Participants that did not reach multiple fatigue levels in all rounds are marked as insufficient data. 106

List of Tables

4.1 Description of Feature Set 20

5.1 Torso and arm movements from literature associated with 7 emotions 32
5.2 Movement Properties 33
5.3 The value in each cell indicates the value of the property (Table 5.2) that was either described in the literature (Table 5.1) to be associated with a particular emotion (blue, italicized text) and/or had the highest Likert score compared to movements with a different value of that property (black, unitalicized text). 36
5.4 Tukey Post-hoc test results on accuracy. Significant differences when holding one variable constant and varying the other variable are shown at the $p < 0.05$ (*) and $p < 0.01$ (**) levels. 49

6.1 Correspondence between the corrective feedback, numeric evaluation, and the distance to the closest reference 57
6.2 Examples of verbal and nonverbal feedback to different situations for the three feedback styles. The **neutral** style does not have verbal feedback while the human is exercising. 58
6.3 Conditions tested in our feedback cadence study design 63

7.1 Performance and preference groups for the 29 participants. The diagonal indicates agreement between the style preference and the style the participant performs best with. 88
7.2 Performance and stated style preference for 10 participants in the older adult study 89
7.3 Feedback style that participants performed best with, split by age group. 89
7.4 The style the participant performs best with compared to the style they rated the highest. The data are split by older adults, younger adults, and all participants. 98
7.5 The style the participant performs best with compared to the style they stated as their preference. The data are split by older adults, younger adults, and all participants. 99

1 Introduction

The field of human-robot interaction (HRI) aims to create robots that integrate into human social settings. Robots are becoming prevalent in a variety of domains, moving from controlled factories to more complex environments such as healthcare, education, and our homes. In these domains, it is increasingly crucial for robots to personalize their interactions, be aware of the contextual nuances surrounding their behavior, and effectively utilize nonverbal communication cues. Failure to personalize interactions and instead rely on a one-size-fits-all strategy can lead to missed opportunities to improve individual performance and experiences based on preferences and context.

Personalized context-aware multimodal feedback focuses on tailoring the robot's responses and expressions to individual users' preferences while considering the dynamic context that shapes the way humans interact. We specifically focus on *affective*, or emotional nonverbal feedback, as we see potential for robots to convey emotions nonverbally, in a context-aware fashion, to influence humans.

When humans interact, they adapt to each other and consider the preferences, requirements, and capabilities of each other, and when robots interact with humans, they should endeavor to do the same. Robots that personalize their behavior have great potential in many applications, including educational robots that provide individualized support and exercise coaches that motivate people to exercise better. Different people also respond to feedback and guidance in different ways. In the case of an exercise coach, some people might prefer a firmer approach and others a more encouraging one. Additionally, some might in general prefer a firm approach based on their personality, but as they become more fatigued, they might like more encouragement to keep going. Considering factors such as personality traits and user-specific requirements, personalized robots can provide tailored assistance and support, fostering an engaging and productive interaction.

Context awareness goes hand in hand with personalization. Robots need to comprehend and respond to the situational factors that influence human communication. Context includes a broad spectrum of information, including the physical environment, temporal aspects, and even the emotional states of the individuals involved. A context-aware robot understands and interprets this contextual information, so it can make informed decisions, adjust its behavior, and provide appropriate responses. By being context-aware, robots can personalize to the individual even better and enhance the human's satisfaction and performance.

Integrating nonverbal behavior into the framework of personalized context-aware robots holds great potential for robots to interact in more nuanced ways. Nonverbal cues, such as facial expressions, gestures, body language, and tone, are a significant portion of how humans communicate [45]. For example, in education, teachers can change their nonverbal immediacy, defined as the degree of perceived physical or psychological closeness between people, using different nonverbal behaviors [66]. A teacher smiling and leaning in to explain a concept to a student will have a very different impact compared to a teacher crossing their arms and frowning while communicating the same information. The student may better internalize the feedback from the smiling teacher and may view the frowning teacher as more strict. The way a human communicates emotion through nonverbal behavior can have a significant effect on the impact of their feedback and how they are perceived. Similarly, robot nonverbal behavior affects humans in different ways, such as the perception of the robot, emotion recognition and response, behavioral response, and human task performance [79].

In this thesis, we explore the creation of personalized context-aware robot feedback, with a specific focus on affective nonverbal behavior. Unlike other types of nonverbal behavior, such as deictic gestures (e.g. pointing), affective nonverbal behavior focuses on conveying emotion through facial expressions, body gestures, and other modalities.

We explore two domains: education and exercise. Personalized feedback in education can be vital; teachers often encourage their students after mistakes and celebrate their successes [16]. A robot tutor can provide that one-on-one support, responding with feedback to the student to help them learn better and enjoy the learning process. Exercise is another domain where personalization can be useful. Personal trainers are a great resource for individuals to improve at their own pace and can provide feedback that a group setting or exercising alone cannot provide. A robot exercise coach could fill that role, providing personalized corrections and motivation by understanding the human's current state as well as the way they prefer to be coached.

Thesis contributions

Research Question: How can we design personalized context-aware multimodal robot feedback to improve human performance and subjective experience?

Chapter 4 primarily explores how to estimate the human state and presents an example predicting measures of engagement during an educational activity, early enough for a robot to intervene and correct any undesirable behavior.

- **Hypothesis**: We can reliably estimate engagement early during an educational activity using a combination of facial and task-related features.
- **Findings:** Our results show that we can predict with high accuracy with more than 80% of the activity remaining, and we also see that a combination of facial and task-related features outperforms using only facial features.

Chapter 5 explores how to generate affective nonverbal robot behavior, which we show improves learning during an educational activity, even though the affective behavior does not provide additional information related to the task.

- **Hypothesis:** Affective nonverbal robot feedback can significantly improve performance and subjective experience in an educational activity.
- **Findings:** A robot using affective nonverbal behavior to react to the human's correct and incorrect responses resulted in a significantly better human performance in a sorting game task, compared to a robot that provided the same content of feedback without the emotional nonverbal response.

Chapter 6 presents a physical robot exercise coach that evaluates the form and speed in real time of human exercise. The robot reacts in a multimodal way to explore people's varying preferences for and performances with different feedback styles.

- **Hypothesis:** Different feedback styles of an exercise coach robot have significantly different impacts on people's performances and subjective experiences.
- **Findings:** We implemented two feedback styles (firm and encouraging) for a robot exercise coach, and found that people preferred the robot to provide context-aware feedback and that people had varying performances and perceptions of the two styles.

Chapter 7 extends the work in Chapter 6 by answering the question: how can a robot determine which feedback style to use and when, in a context-aware way? We present a contextual bandit approach that accurately chooses the feedback style to optimize performance in complex fatigue scenarios for an exercise coach robot.

- **Hypotheses:** A robot learning style preferences in real time can significantly improve human exercise performance. Additionally, a robot that takes context into account can significantly improve performance over a robot that uses non-contextual learning.
- **Findings:** Our results show that the adaptive contextual bandit approach is successful in determining the style to choose in real time and performs well for humans who perform better with either one style or equally with both. We also found that a bandit that takes context into account can outperform one that does not in some contextual situations.

2 Related Work

2.1 Human Feedback

In human interactions, different types of feedback can result in differing task performances and experiences, and people use a combination of verbal and nonverbal modalities to communicate. When speech is accompanied by gestures, people interpret may interpret the human's meaning differently, as nonverbal behavior helps disambiguate verbal feedback [45]. This motivates our use of both verbal and nonverbal feedback for the robot exercise coach presented in Chapter 6.

Nonverbal communication is defined as the process of stimulating meaning in the minds of others through messages that are not language-based [73]. This form of communication is composed of several categories including kinesics (body motions, posture, facial expressions, etc.), paralanguage (voice volume, pitch, tempo, etc.) and proxemics (use of space) [85]. Nonverbal communication can also convey emotions, with some research suggesting that up to 90% of emotional meaning is conveyed through nonverbal behaviors [62]. Humans rely on a combination of many characteristics to display emotion, including facial expressions [23], which have been decomposed into Facial Action Units in [24]. In our work, we utilize estimates of these Facial Action Units to model the human's emotional state (Chapter 4).

Nonverbal communication can serve a variety of purposes, including complementing verbal exchanges, revealing emotional states, and influencing the performance of others [92]. Nonverbal behavior, specifically movement, gesture, posture, facial expressions, and gaze, was shown to be vital in communication between a coach and athlete [22].

The robot exercise coach presented in Chapter 6 utilizes these modalities to convey affective feedback to the human during an exercise session. In students, nonverbal behaviors, such as slouching, could indicate a lack of interest, and leaning forward could show attentiveness [66]. Nonverbal behavior of students has even been shown to have an impact on teachers' emotions [32]. Emotion communication researchers found that body movements of professional actors showed patterns in portrayals of specific emotions [21], and we find similar patterns with affective robot behavior in Chapter 5.

Humans can also react differently to feedback given in different ways, motivating our use of multiple feedback styles in Chapter 6. Researchers found that students had different preferences for how teachers should give different types of feedback [16]. Researchers also found that people generally prefer process-based criticism over personal criticism [36], which motivated our design of robot verbal feedback in Chapter 6. Choosing the correct feedback style can even have an impact on task performance. For example, in one study, groups that received error feedback outperformed groups that received other types of feedback in learning how to jump vertically [67]. Human teacher nonverbal behavior has been shown to affect both task performance and perception of the teacher [6, 7, 93]. For example, students thought their teacher meant what she said when her voice became louder and when she stood and looked down at them.

We attempt to leverage this effect in our work by determining how to choose the correct style of feedback for each person (Chapter 7). In the exercise domain specifically, feedback preferences can be impacted by a variety of factors, as shown in [34], including physical health, status, and educational level.

2.2 Generating Robot Feedback

Robots providing feedback, especially by expressing emotions, is a highly studied area for both humanoid and non-humanoid robots, and robot body language has been used to convey emotion [74]. Researchers investigated how changing the poses of the NAO humanoid robot affected the perception of the robot's affect [9]. This work looked at creating an *Affect Space* where the nonverbal behaviors of the robot map to perceived emotion, resulting in a mapping different from that for humans, and we perform a similar procedure for the Quori robot in Chapter 5. Our prior work looked at how even a low-dimensional robot that mimics one aspect of human behavior with its nonverbal movements could improve the perception of the robot [37, 38].

Using designed body language for a humanoid robot, researchers tested emotion recognition in [60]. Emotional behavior by a humanoid robot (NAO or Mini Darwin Platform) was developed to assist autistic children with emotion recognition [25]. Researchers also developed a framework for displaying robot traits, moods, and emotions using nonverbal behaviors [68].

Laban movement analysis [51] is another approach to generate robot movements. Flight paths were used to convey affect in [82]. Using Laban efforts, head movements were used to convey a robot state using Keepon and NAO robots [49], and the path of mobile robots was used to create expressive motion [48]. Researchers have also parameterized hand movements (e.g., waving) using pose and motion parameters that participants used to create movements corresponding to different moods [95].

Robots influence humans through their nonverbal communication in human-robot social interactions and [79] separates the impact on humans into cognitive framing (perception of the robot), emotion recognition and response, behavioral response (directly measured human responses), and task performance. In education, social robots have been shown to be effective in increasing cognitive and affective outcomes [10], and our work also seeks to improve both performance and subjective experience through robot feedback.

2.3 Robot Feedback Improving Performance

Robots have used their behavior to improve human performance in a variety of tasks. Higher nonverbal immediacy led to greater learning gains in child-robot interactions [46]. Nonverbal behavior of the robot (deictic rather than affective in this case) improved task performance for difficult collaborative tasks [1]. Robot gestures made difficult tasks feel easier, specifically looking at perceived workload and task performance [54]. Eye contact and iconic robot gestures improved message retention in an interaction between robots and humans [90]. In an evacuation scenario, a robot's nonverbal expressions improved participants' compliance, causing them to respond earlier and faster [65]. The gaze cues of a robot improved the performance of the human agent in a cooperative task, showing the importance of nonverbal behavior in communication [12]. A robot exercise coach was shown to reduce mistakes in combination with increased fitness [76]. Many of these works do not explore how affective nonverbal behavior affects performance (as we show in Chapter 5), and we also utilize multimodal behavior to impact performance, where the robot learns the feedback style with which the human performs best in an online fashion (Chapter 7).

2.4 Robot Feedback Improving Experience

Robots have also used their behavior to improve the human experience during an interaction. Participants found that a robot programmed to have a positive mood increased the valence and arousal of its audience compared to using a negative mood [96]. Researchers have also developed a model for a NAO robot teacher to express different levels of warmth and competence with its body postures and hand gestures [71]. In a longitudinal study, robots improved social engagement by tutoring children [81]. A study on the effects of affective human-like and robot-specific behavior showed that these behaviors impact the perception of the robot and the human's affect [74]. Robot motion and mimicry had a significant impact on similarity and closeness in robot-mediated communication [17]. Combining nonverbal gestures with verbal information can further improve the human experience, as found in [75]. One main difference between these works and the research presented in this thesis is the focus on multiple feedback styles and understanding how these different styles affect the human's experience (Chapter 6).

2.5 Personalization and Context

Personalization refers to the ability of a robot to tailor its actions and behaviors to individual users. Researchers developed an approach to adapt the verbal and nonverbal behavior of a robot based on human extroversion [4], and a study of post-stroke rehabilitation found that people preferred a robot matched to their extroversion/introversion [89]. The personalization of the tutors to the learning differences of the students resulted in an improvement between the results of the preand post-test in an educational activity [53]. Additionally, participants in a study varied their preferred distance from the robot based on various personality traits, such as proactivity [91]. Another study found that participants preferred to interact with a robot that synchronized its movements with that of the human [84]. A robot personalizing to the individual elicited and maintained engagement and motivation during an exercise program [35]. A social robot can even personalize its facial expressions to respond to the user during an interaction [69]. Students interacting with a robot that personalized its feedback showed a significant increase in emotional response [30]. Personalizing language to the individual was more effective than formal language in multimedia instruction [72] and the customization of multimodal robot behavior based on attention score history increased engagement [3]. These works motivated our use of multiple feedback styles so that the robot could personalize its behavior to improve its interaction with the human. However, many of these works do not learn the personalization in real time, which differentiates our approach.

Context is an important aspect of human behavior [47], and contextual information can improve emotion recognition [14]. Research showed that data analyses that do not take context into account can be difficult to interpret [92], and we see a similar result in Chapter 4 where the use of task-related features in addition to facial expressions improved the prediction of the human state.

Researchers found that context can have a large impact on both quantitative and qualitative results; gaze models had varied results based on the location of the experiment [70], and children perceived a robot differently based on their presence at school or at camp [71]. The environmental context had an effect on the recognition rate of stylized walking sequences [33]. Even cultural differences have an impact; facial expressions are perceived differently in different cultures [57, 58]. Having the robot take both task information and a human model into account to inform its feedback improves the human experience and the relevance of the robot's feedback.

2.6 Summary

This thesis incorporates these different aspects of related work. We see in Section 2.1 that humans use both verbal and nonverbal feedback. In our work, we generate multimodal robot behavior using verbal feedback, facial expressions, and body movements to leverage the expressive capabilities of all of these modalities. We also found that humans react differently to feedback given in different ways, and we explore this aspect in our work by creating multiple feedback styles to allow the robot to choose the style to which each person responds best (Chapters 6 and 7).

We also see in Sections 2.3 and 2.4 that robots providing feedback during an interaction can improve both performance and subjective experience. However, most of the previous work in this area only focuses on one of these outcomes, and we show in Chapter 5 that our multimodal robot feedback can improve both performance and experience. We also explore trade-offs between these outcomes in Chapter 7, trying to understand the impact of optimizing for one outcome over the other.

Furthermore, Section 2.5 explored how humans and robots personalize their behavior to improve interactions, as well as to take contextual information into account to improve the relevance of feedback. Although these other works may not incorporate personalization and contextually-aware feedback simultaneously, our work does combine these aspects in addition to adapting the robot's behavior in real time in response to the human's changing performance (Chapter 7).

3 Approach

We seek to develop personalized, context-aware robot feedback, specifically focusing on nonverbal affective robot behavior. We consider settings where a person interacts with a robot repeatedly, such as performing multiple sets of exercises or multiple learning activities. At a high level, our approach (Figure 3.1) to address this can be summarized as follows:

- The human has certain intrinsics including their personality, motivation, and preferences that affect the way they respond to feedback.
- They perform the task while observing the multimodal feedback of the robot.
- The robot observes the human and estimates the context which could include task information (e.g. task type, interaction time) and a human model (e.g. performance, fatigue, attention).
- The robot has a specific feedback style that governs how it responds to the estimated context with a multimodal response (verbal, facial expressions, body language, etc.).

We aim to close the loop by *personalizing* the robot's feedback. Using the human's response, the robot can modify its feedback style to improve both the human's performance and their subjective experience of the interaction. For example, the robot could observe the human performing better after a correction and could note that the style it used for the correction was effective for that contextual situation. Context is a key component of this process because estimating the human's performance and state can allow the robot to see what effect its feedback is having and make any modifications that could be beneficial. For example, the robot can have an estimation

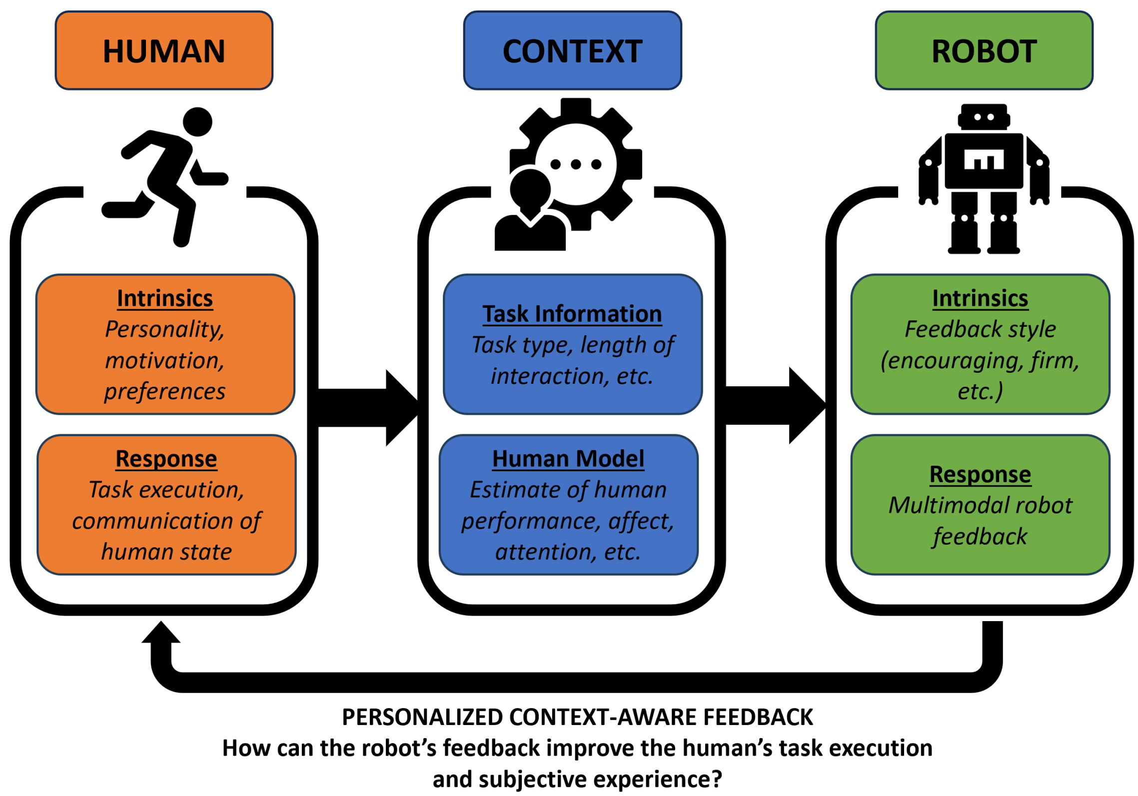

Figure 3.1: **Left**: The human performs the task and responds to the robot's feedback based on their own intrinsics. **Middle**: The robot observes the human's response and estimates the context. **Right**: The robot uses its context estimation and its own intrinsics to generate a multimodal response. **Goal**: The overall goal is to personalize the robot's intrinsics to improve specific outcomes.

of fatigue based on its observations of the human and how long an exercise session has progressed, and when that fatigue is high, the robot could modify its feedback to be more encouraging and not push the human too much, as they are already fatigued.

In the remaining sections of this chapter, we demonstrate how the work in this thesis fits into our overall approach.

3.1 Estimating Context

The first step of our approach is estimating the context: both the task information and a human model. Task information refers to task type, interaction length, and other information about the interaction that is not directly related to human sensing. For example, if the robot is presenting the human with an educational activity, the robot will already know the correct answer to any questions presented, as well as the types of questions and the duration of the interaction. To construct a human model, the robot must observe certain features from the human (e.g., facial expressions, body movements, heart rate) and use them (in conjunction with the task information) to estimate useful information such as task performance, fatigue, and attention. Knowing information such as how long a task has been going on, as well as measured facial expressions, can improve the robot's estimate of the human state (e.g., fatigue, attention). Incorporating task-related features can be vital in interpreting the observations of the human. A smile in frustration and a smile of joy are both smiles, but knowing whether the individual has just made a mistake or a correct response can distinguish between them.

Chapter 4 presents an approach to estimate one aspect of the human state: engagement. In this work, we use task features and facial features to predict whether students would exit an activity early, as well as what feedback they would choose to provide at the end of an activity. Being able to predict these events early would allow a robot to intervene with appropriate feedback.

Chapter 6 includes a prediction of performance, specifically, how well an individual performs an exercise. This work uses real-time pose estimation to analyze the human's movement and evaluate both their form and speed while exercising. This information is vital when the robot determines what feedback to respond with.

Chapter 7 explores the estimation of the human state (fatigue) as a vital part of context to accompany the work in Chapter 6. We then use this more complete estimation of the context to inform the robot's feedback and personalize its behavior.

3.2 Generating Robot Feedback

Once the robot has estimated the context, it must react in a multimodal way. In general, the types of feedback that the robot could provide fall into two categories: reinforcement and corrections. If the human is performing well in the task (e.g. correct responses in an educational activity or good form for an exercise), then the robot should provide encouragement and reinforcement of the good performance. If the human is not performing well in the task (e.g. incorrect responses in an educational activity or improper exercise form), then the robot should provide a (gentle) correction to guide the human to improve their performance. From the literature, we can see how humans provide both encouragement and corrections and the way that people could prefer that feedback during an interaction. However, we do not know a priori how a human would respond to a robot giving feedback, and in this thesis, we explore the design of verbal and nonverbal robot feedback using the robot Quori (Figure 3.2).

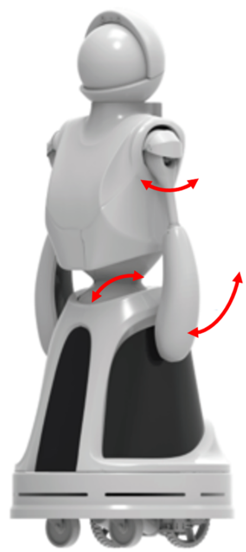

Figure 3.2: The humanoid robot Quori with a leaning torso, gesturing arms, and face that can project any image.

Chapter 5 presents an approach for generating affective body movements for Quori, where specific body patterns are correlated with conveying certain emotions. We then show that the robot's affective nonverbal feedback has a positive impact on participants' learning during a user study compared to a baseline of only verbal and no nonverbal reactions. We also show that mismatching or inappropriate nonverbal feedback does not have the same positive impact on learning; the robot's verbal and nonverbal feedback must be aligned and related to the contextual situation.

Chapter 6 extends feedback generation by adding additional robot capabilities: more nuanced verbal reactions and facial expressions. These additional modalities work together with the robot's body movements to provide more complex multimodal feedback. This greater variety of possible feedback allows us to differentiate our feedback styles, and in turn, move toward personalizing the feedback styles to the individual.

3.3 Developing Feedback Styles

The next step of our approach is developing the way the robot should respond, the feedback model it uses to turn the estimated context into the robot's feedback. For example, given the same stimulus, a human teacher will respond in differing ways based on their own personality as well as what they think their student will respond best to. If the teacher knows that the student appreciates very direct corrections, they will try to respond in the way the student prefers. Our robot, therefore, must be able to respond in a variety of ways to the same estimated context.

In Chapter 5, we show that the robot's feedback centered around contextually appropriate affective nonverbal behavior was effective in improving the human's performance compared to a neutral feedback style that did not react nonverbally to the human's actions.

We also explore this idea in Chapter 6, where we develop two exercise coaching styles (based on what we learned from domain experts): a firm approach and an encouraging approach, both of which use aligned verbal and nonverbal reactions. Through a user study, we compare these approaches and see how participants' preferences for these styles vary, as well as their performance differences.

3.4 Personalizing Feedback

The final stage in our approach addresses how the robot should choose which feedback style to use for an individual in a contextually aware way, with the goal of optimizing performance. For example, a person may perform better with less encouragement when they are not tired but with more encouragement when they become fatigued. Other individuals may perform better with less encouragement regardless of their level of fatigue. As we show in Chapter 7, what people say they prefer does not necessarily match up with what they truly perform best with, so for a robot to determine the appropriate feedback style to use in each situation, it must adapt in real-time.

Chapter 7 uses a contextual bandit approach [13] to achieve this goal. With a contextual bandit algorithm, the robot observes the context (fatigue), chooses an action (feedback style), and observes the reward as a result of its decision (human response). Then it trains on the combination of context, action, and reward to learn which styles to use, and when, for an individual. We explore this personalization with a human model in complex contextual situations as well as in a user study that compares the effectiveness of this adaptive approach over simply choosing a static feedback style. This final work showcases the overall effectiveness of personalized context-aware robot feedback.

4 Human State Estimation

Summary: This chapter explores an approach for estimating key features of the human state. Specifically, we utilized an educational game [61] where students provide feedback at the end of an activity and have the option to exit the activity early. Predicting these user engagement related events that are a key component of context could help an agent provide appropriate feedback. We present an approach that uses a combination of human and task features, fits Gaussian Mixture Models, and propagates the probability of the engagement-related events over the course of an activity. Our results show that we can predict these events early enough to intervene during the activity and additionally show the advantage of using both task and human features for improved accuracy. Based on these results, we used a similar approach (combining task and human features) for the exercise coach presented in Chapter 6.

The contents of this chapter were published in [39].

4.1 Problem Formulation

Intelligent tutoring systems, whether in the form of a physical robot or tablet/computer interface, have great potential to personalize the educational experience and meet the needs of many different learners. These systems generally take some feedback from the user to adapt the educational tasks, attempting to optimize learning, engagement, or other performance measures. For a specific application, a tutoring system designer has to determine how to adapt the tutor's behavior based on the changing tutor/user/task interactions.

In this work, we used data from an intelligent tutoring system [61] to develop an approach for estimating user engagement through the prediction of key events. Ideally, the system should detect early when the tutor should change its behavior (e.g., when the student is disengaged), so that the tutor could intervene with feedback. To achieve this, we used task and human features derived from an existing educational application dataset to develop a model that predicts two key events: (1) whether the student exits the activity early and (2) the feedback the student provides to the system at the end of the activity.

To predict these engagement-related events, we utilized a time-series classification approach. Time-series classification methods take a set of observations labeled by a class and predict the class based on the observations. Deep learning is a popular tool for time series classification [28], and LSTMs have been used successfully on time series of the same length. However, these approaches generally make a prediction at the end of the time series, while we wish to predict before the end of the activity. A measure of earliness of prediction has been used in time series classification work in addition to an accuracy measure to evaluate model performance [26, 64, 94], and we used a similar approach to predict both accurately and early.

4.2 Robotutor Dataset

RoboTutor [61] is an educational application running on an Android tablet that contains many activities, including reading illustrated stories, practicing writing words, and solving math problems. Previous work on data from this app included affect detection based on expert labels [2] and correlation between some user behaviors and facial features automatically detected by OpenFace [80]. This app was a finalist in the Global Learning XPRIZE[1], and during that time was used by children ages 6-12 in Tanzania. The data used for this project come from beta sites and are screen recordings during the student sessions, which include a video from the front-facing camera ([61] has more details, including ethical considerations).

Figure 4.1 (left) is a screenshot of a story reading activity, which contains an image and text from the story. The text is highlighted in green as it is read to the

[1] https://learning.xprize.org/

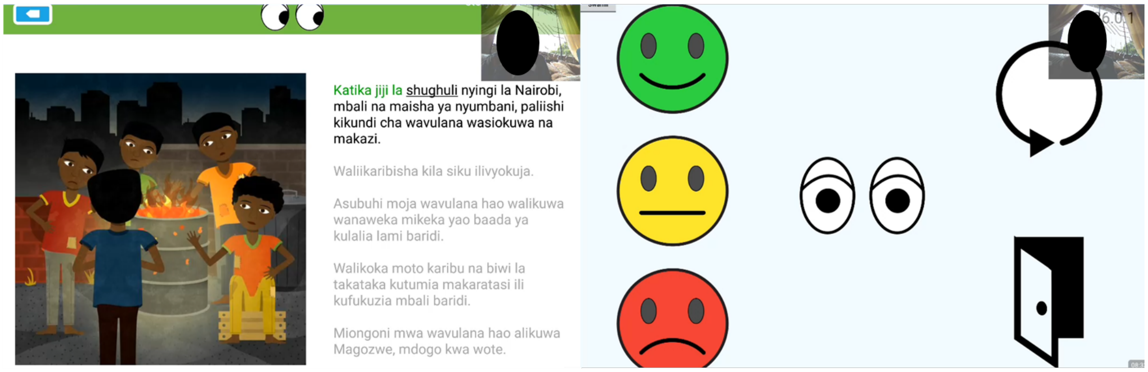

Figure 4.1: **Left**: Single frame from recorded videos of the RoboTutor application with front-facing camera feed (face obscured for anonymity). **Right**: Feedback screen appearing at the end of each activity.

student. There is also a *backbutton* in the top left corner, which the student can use to end the activity early. At the end of an activity (whether it has ended early or fully completed), a feedback screen (Figure 4.1, right) appears where the student can choose a red, yellow, or green circle to indicate how they felt about the activity, and the remaining icons on this screen help navigate between activities.

4.2.1 Automatic Feature Extraction

To create prediction models, we first extracted a set of features from the screen-recordings. We only used activities that contain stories to limit the variety of activities to analyze. For each activity, we collected two different labels: *Feedback* (the student's choice on the feedback screen: red, yellow, or green) and *Backbutton* (whether the student exited the activity early). Screen taps, which appeared on the videos as white circles, were detected to determine the feedback chosen or whether the backbutton was pressed.

We extracted a set of *facial features* (Table 4.1, A-K) over the course of the activity using OpenFace [5]. These are the same features computed by previous work with this dataset [2] and include features that have been used frequently in affect recognition [15]. The six facial action units (F-K) are coded by their regression values corresponding to the intensity of presence [24]. We also extracted a set of *task features* (Table 4.1, L-Q) that relate to the educational activity itself. These features were chosen because they were easily extracted from a frame on the tablet screen and could represent information correlated with student engagement. As we did not collect this dataset ourselves, the features we could use were limited by those that were recorded at the time of data collection.

Table 4.1: Description of Feature Set

	Facial Features
A	**Head Proximity**: the scalar distance of the head from the camera
B	**Head Orientation**: the magnitude of rotation of the head
C	**Gaze Direction**: the averaged angle of gaze between the two eyes
D	**Eye Aspect Ratio**: related to blinking of the eye
E	**Pupil Ratio**: the ratio of the area of the pupil to the area of the eye
F	**AU04**: Brow Lowerer
G	**AU07**: Lid Tightener
H	**AU12**: Lip Corner Puller
I	**AU25**: Lips Part
J	**AU26**: Jaw Drop
K	**AU45**: Blink
	Task Features
L	**Position of Activity in Video**: sequential order of activity in video
M	**Picture Side**: left or right side of the screen
N	**Activity Type**: story read or story echo
O	**Progress**: non-decreasing scalar indicating how far along in the story, computed by green vs. black text on a page (see Figure 4.1)
P	**Time from Activity Beginning**: in seconds
Q	**Time from Educational Session Beginning**: in seconds

4.2.2 Description of Dataset

Our dataset consisted of 105 video recordings of student sessions, each 20-30 minutes long. We first extracted individual story activities from the videos. Since the final activity of each video often corresponded to the instructor telling the student to stop the session, this activity was not included. Our activity data set was then composed of 423 activities of length 5-950 seconds, with most activities shorter than 200 seconds. The distribution of the *feedback* labels was 13.2% red, 77.1% yellow, and 9.7% green, and the distribution of *backbutton* labels was 87.9% no backbutton and 12.1% backbutton.

Data are represented as $\{T^{(i)}, X^{(i)}, Y_1^{(i)}, Y_2^{(i)} \mid i = 1, ..., 423\}$ where $T^{(i)}$ is a vector such that $T_j^{(i)}$ is the time of the jth frame of activity i; $X^{(i)}$ is a matrix such that $X_j^{(i)}$ is a vector corresponding to the features computed for the jth frame of activity i; $Y_1^{(i)} \in \{0, 1, 2\}$ corresponds to a *feedback* choice of red, yellow, or green; and $Y_2^{(i)} \in \{0, 1\}$ corresponds to an activity ending naturally or the *backbutton* being pressed. The goal of our approach is to accurately predict Y_1 and Y_2, given only a few frames of T and X.

4.3 Methodology

We used facial and task features to predict which feedback was chosen at the end of an activity or whether the backbutton was pressed to end the activity early, with the goal of predicting as accurately and early as possible. During the course of the activity, we combined individual observations using a Bayesian framework and used Gaussian Mixture Models (GMMs) to provide the probabilities needed for Bayesian updating. The approach chose hyperparameters to optimize the desired balance between the F1 score (weighted by α) and earliness (weighted by $1 - \alpha$), where α is an input to the learner. We describe the methodology for predicting the *feedback* labels here with $K = 3$ labels; the *backbutton* case is analogous.

4.3.1 Training Gaussian Mixture Models

Given a set of training data of the form $\{T^{(i)}, X^{(i)}, Y_1^{(i)}\}$, we trained a Gaussian Mixture Model for each label, i.e. red, yellow, and green. GMMs were chosen as they could represent multi-modal data and output a probability that an observation belongs to a class, which we used in the probability propagation step. However, we noticed that the distribution tended to change over time (e.g., at the beginning of an activity, facial features tend to be less informative predictors), so to improve the prediction we trained multiple GMMs for each label by first creating M intervals from the distribution of activity lengths.

Each of the M intervals had a starting and ending time (e.g. the first interval may include 0-30 seconds, the second 30-120 seconds, etc.) such that the number of activities ending in each interval is approximately the same. For each of the K labels, we first found all the activities within the training data with that label, took only the time steps of the activities corresponding to time steps within a particular interval, and then trained a GMM with N components on those data. This resulted in a total of MK GMMs trained. Given the training data, our approach learned the models and optimized for M and N.

4.3.2 Probability Propagation

We then used the GMM models to predict $P(C_k)$, the probability of the k^{th} class, for each time step of an activity. We initialize $P(C_k) = \frac{1}{K}$, which corresponds to a random guess. Then we let $\mathbb{X}_j^{(i)} = \{X_1^{(i)}, X_2^{(i)}, ..., X_j^{(i)}\}$ be the observations known at the jth time step. We calculated the probability at the next step using a modified Bayes rule from [52]:

$$P(C_k|\mathbb{X}_j^{(i)}) = \frac{P(X_j^{(i)}|C_k)P(C_k|S)P(C_k|\mathbb{X}_{j-1}^{(i)})}{P(X_j^{(i)}|\mathbb{X}_{j-1}^{i})} \tag{4.1}$$

where $P(C_k|\mathbb{X}_{j-1}^{(i)})$ is the computed probability from the previous time step; $P(X_j^{(i)}|C_k)$ is the output of the GMM that was trained on a time interval including $T_j^{(i)}$ corresponding to class C_k; and $P(C_k|S)$ is the static prior from the training distribution to avoid model drift. For example, if C_k = yellow and yellow labeled 80% of the training data, then $P(C_k|S) = 0.8$.

The denominator is a constant over all k, so is normalized out by ensuring that the $P(C_k|\mathbb{X}_j^{(i)})$ sum to 1. To try to ensure conditional independence between observations and reduce the effect of noise in the features, the features $X^{(i)}$ are averaged over a one second interval, and the probability is updated once a second. If no features are present during one second, due to errors in face detection, the probability from the previous time step is used unchanged.

4.3.3 Classification

If the goal was to classify at the end of each activity, we would use the highest $P(C_k)$ in the final time step. However, classifying earlier is beneficial, since that information could be used to modify a tutoring system's behavior. To achieve this, we set a threshold $\lambda \in [0, 1]$ such that if any $P(C_k)$ exceeds λ, we classify[2] the activity as belonging to class k. If no $P(C_k)$ exceeds λ for the entire activity, we counted it as an inconclusive result, as predicting a class at the end of an activity does not have utility, since we know what the students did and no intervention is possible.

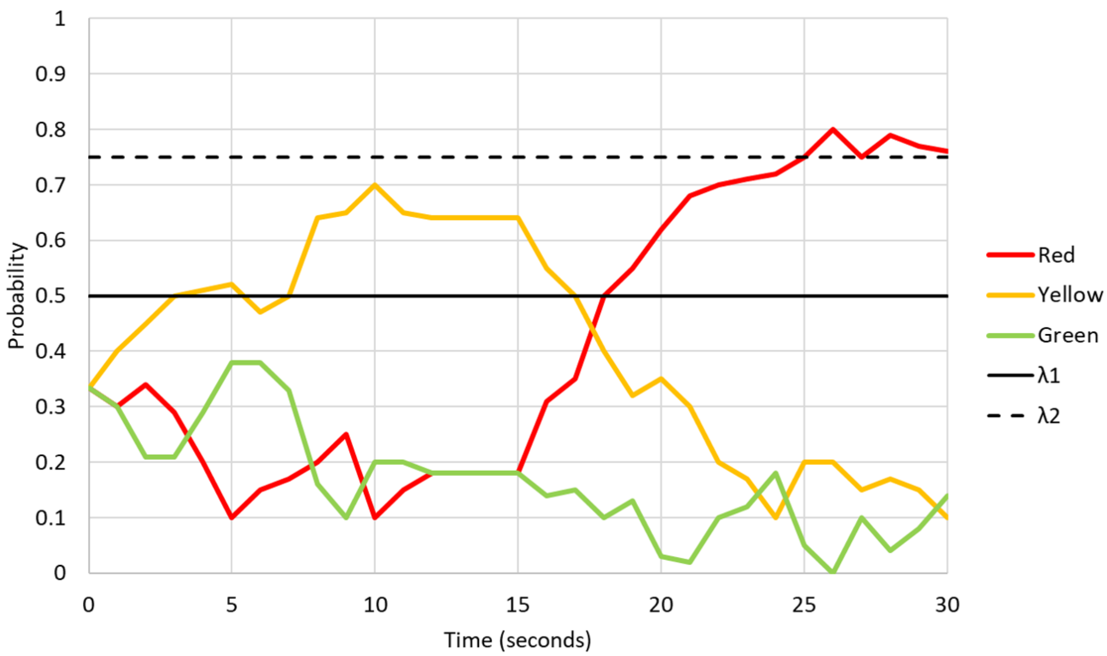

Figure 4.2: Example predicting *feedback* choice using the described approach with two possible thresholds λ_1 and λ_2 shown.

Figure 4.2 illustrates the prediction process applied to an activity of length 30 seconds for the *feedback* case. The probabilities are initialized to 1/3, and the observations are combined using Equation 4.1. Note that between 12-15 seconds, OpenFace failed to find the face, resulting in a flat probability curve for all labels. The figure plots two different thresholds to show how the choice of threshold impacts both the time of classification and the predicted label. λ_1 predicts yellow at 3 sec, while λ_2 predicts (the correct label) red at 25 sec.

[2]We tried varying λ over time, but that did not improve results.

4.3.4 Optimizing Performance

To predict both accurately and early, we optimized using an objective function S, a function of λ (threshold), M (time intervals), and N (GMM components), as well as a weight $\alpha \in [0, 1]$. α defines how much we prefer an accurate prediction over an early one. As our dataset was quite unbalanced, we used a weighted F1 score in place of accuracy. The F1 score is calculated for each label, and we report the average weighted score by the number of true instances for each label.

The form of the objective function S is shown below.

$$S(\lambda, M, N, \alpha) = (\alpha)\text{F1 score} + (1 - \alpha)\text{Earliness} \tag{4.2}$$

Earliness, or the average fraction of an activity's time that was **not** needed for classification, is defined as $\frac{1}{n}\sum_{i=1}^{n}\frac{T_{-1}^{(i)} - \hat{t}^{(i)}}{T_{-1}^{(i)}}$, where n is the number of activities where the threshold is met; and $T_{-1}^{(i)}$ and $\hat{t}^{(i)}$ are the activity length and prediction time for activity i.

4.4 Results

Our goal was to optimize the performance metric S by changing the three parameters: $\lambda \in \{0.55, 0.60, ..., 0.95\}$, $M \in \{1, 2, ..., 6\}$, and $N = \{1, 2, ..., 6\}$. For each combination of these parameters (324 total), we performed 10-fold cross-validation and recorded the average S over all folds. We then chose the hyperparameter combination with the highest average value of S. Additionally, when comparing the performance of two different models, we used a Welch two-sample, two-tailed t-test, which does not assume that the two variances are equal.

The optimization was dependent on the choice of α, which trades off the F1 score for earliness. $\alpha = 0$ means we prioritize only earliness and $\alpha = 1$ means we prioritize only the F1 score. We performed the optimization of S for values of $\alpha \in [0, 1]$. We found that when $\alpha > 0.8$, performance drops significantly, so we chose $\alpha = 0.8$ for further analysis. Additionally, low levels of α, such as $\alpha < 0.3$, have a lower performance metric S due to lower accuracy.

4.4.1 Earliness and Guessing

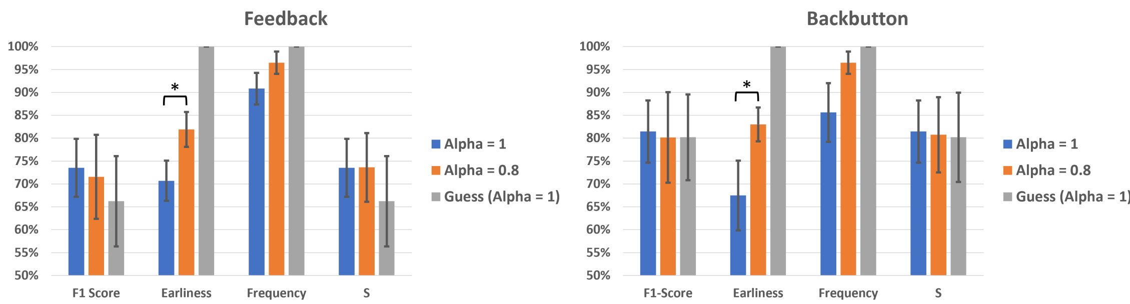

Figure 4.3: Comparison of $\alpha = 1$, $\alpha = 0.8$, and guessing the most common label, for *feedback* (left) and *backbutton* (right). The average value of metrics F1 score, earliness, frequency, and S are illustrated with a standard deviation bar. Significant differences at the $p = 0.05$ level are indicated with asterisks. Significance was tested only for F1 score and earliness.

Often, accuracy is the only metric used to evaluate prediction models. Intuitively, the more of the activity seen by the model, the more accurate the prediction will be; however, we wanted to make predictions before the activity had finished to allow time for any intervention. To understand how this trade-off manifested in our model, we compared the results of considering only the F1 score ($\alpha = 1.0$) and including earliness ($\alpha = 0.8$). We also validated our approach by comparing the performance with guessing the most common label in the training data for each activity (e.g., choosing yellow or no backbutton) at the first time step.

As shown in Figure 4.3, the F1 score was lower for a lower α, which is intuitive since $\alpha = 1$ only weights the F1 score. However, that difference was not statistically significant, while adding a weight of 20% on earliness *does* significantly change earliness for *feedback* ($t = -5.782, p < 0.001, df = 17.64$) and *backbutton* ($t = -5.507, p < 0.001, df = 13.00$). This increase in the earliness with no significant change in the F1 score indicates that including earliness improves overall performance, with respect to our goal of predicting accurately and early. Figure 4.3 also plots the frequency of prediction, where a value of less than 100% indicates that for some activities, the threshold was not reached by the end of the activity and no classification was performed. As shown in the results, $\alpha = 0.8$ classified a greater percentage of activities compared to $\alpha = 1$.

The optimal hyperparameters vary for each value of α. For *feedback*, we found $(\lambda = 0.95, M = 2, N = 3)$ for $\alpha = 1.0$ and $(\lambda = 0.7, M = 2, N = 3)$ for $\alpha = 0.8$ to be optimal. For *backbutton*, we found that $(\lambda = 0.95, M = 4, N = 1)$ for $\alpha = 1.0$ and $(\lambda = 0.55, M = 2, N = 3)$ for $\alpha = 0.8$ are optimal. Note that the optimal threshold λ when $\alpha = 1$ is much higher than for $\alpha = 0.8$, which makes sense since with $\alpha = 1$ there is no penalty for waiting longer in exchange for greater prediction confidence.

Another interesting result is that the optimal number of time intervals M was greater than 1 for all cases. This means that using multiple intervals to segment the time series data increased overall performance.

We also compared the results to guessing the most common label (shown in gray in Figure 4.3). Although the F1 score resulting from guessing was lower than our model at $\alpha = 0.8$ and 1.0 for both cases, this difference was not statistically significant (note that the earliness scores are always 1, since guessing is done at the start of an activity). We anticipate that with a larger dataset and more balanced label distribution, guessing will perform worse.

4.4.2 Facial and Task Features

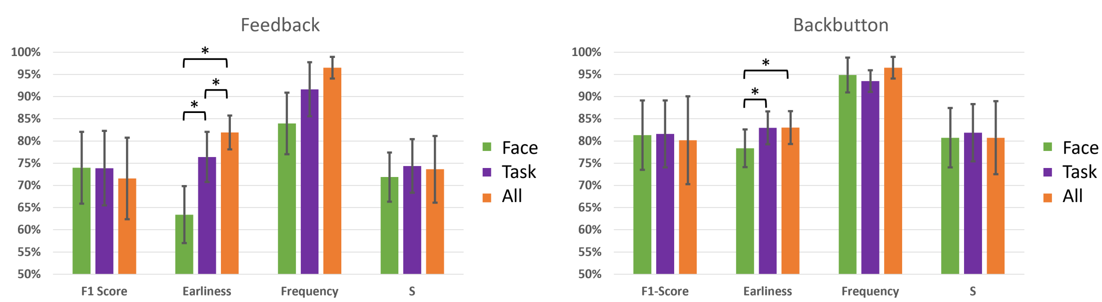

Figure 4.4: Comparison of using only facial features, only task features, and all features, for *feedback* (left) and *backbutton* (right), with the optimal hyperparameters and $\alpha = 0.8$. The average value of metrics F1 score, earliness, frequency, and S are illustrated with a standard deviation bar. Significant differences at the $p = 0.05$ level are indicated with asterisks. Significance was tested only for F1 score and earliness.

We hypothesized that the context of the task can help interpret the student's internal state. Facial features alone have been used extensively to predict affect, as in

[78, 86]. To evaluate this, we compared the performance of using only facial features (A - K in Table 4.1), only task features (L - Q in Table 4.1), and all features, shown in Figure 4.4 with significant differences indicated with asterisks.

We found significant differences in earliness when comparing a model using only facial features and a model using all the features. Specifically, the t-test resulted in $(t = -7.422, p < 0.001, df = 14.62)$ for *feedback* and $(t = -2.50, p = 0.02, df = 17.67)$ for *backbutton*. Additionally, for the *feedback* case, there was a significant difference between earliness using only task features and using all features $(t = -2.417, p = 0.03, df = 15.76)$.

An interesting result is that the task features alone predicted earlier in both cases compared to facial features alone; with the F1 score not significantly different. This does seem non-intuitive, since facial features have been used extensively for affect recognition. Task features do not encode any information directly from the student and instead record progress in the activity, so it seems unlikely that they would outperform facial features. A potential explanation is the noisy data output by OpenFace. The students moved rapidly in the camera frame, and occasionally another student appeared in the frame during an activity. In contrast, task features were less noisy, as they were computed from relatively static and predictable items on the tablet screen during an activity.

Another explanation could be that engagement is tied closely to the time a student has spent using the tablet (one of the task features), perhaps due to fatigue. We conducted an ablation study, removing one feature at a time, to determine which features had the greatest impact on the objective function S. We found that feature P (time from activity beginning) and feature Q (time from educational session beginning) resulted in statistically significantly better performances when included compared to when removed. This lends validity to this explanation that engagement is tied closely to these time-related features.

4.5 Discussion

Our framework for predicting whether an event occurs at the end of an activity before the activity actually ends, using a combination of facial and task features, is easily generalizable and can predict any event occurring at the end of a time series given a set of features computed over the course of that time series. We can handle time series of varying lengths without trimming or warping the data, as is often necessary for other time series approaches, such as LSTMs [11, 28].

Since "ground-truth" engagement does not exist, our approach used the *feedback* and *backbutton* events as proxies, which means that the interpretation of the results can be ambiguous. An issue brought up by those who collected this data [61] is that not all students understood the semantics of the three feedback buttons. The meaning of each was not clearly explained to them and therefore each student interpreted the buttons slightly differently. This could explain why the F1 score is significantly lower for *feedback* compared to *backbutton* (an unambiguous indicator of engagement or fatigue). We mitigate this issue in our work in Chapter 6 with an exercise coach that can analyze human performance in real time along with periodic surveys to collect subjective measures.

5 Affective Nonverbal Feedback

Summary: This chapter explores affective nonverbal feedback for the robot Quori. We first present a user-driven approach to generate body and arm robot movements that correlate with various emotions and compare those results to literature. We then developed a sorting game task in which players guess which of two bins cards belong to based on their properties. We also present a methodology for choosing the order of cards based on a model of how players learn the rule over time. We used this game to develop a user study with a simulated robot giving feedback between each card, comparing a *neutral* robot with minimal nonverbal movements and a *matching affective* robot performing a nonverbal movement corresponding to the affect it is trying to convey (happy for correct, sad for incorrect). Our results showed that participants learned the rule better (higher sorting accuracy) with the *matching affective* robot compared to the *neutral* one, and we also examined subjective measures and interactions with the difficulty of the sorting rule. We additionally tested nonverbal behavior that did not match the contextual situation (e.g. sad movement coupled with a correct answer), and this did not improve the sorting accuracy, indicating the importance of verbal-nonverbal congruence. This work showcases the utility of including affective nonverbal feedback, and we utilized these movement patterns Chapter 6 when developing a physical robot feedback system.

The contents of this chapter were published in [40, 41].

5.1 Generating Affective Movements

This section describes how we generated a variety of movements using an approach similar to [83] and asked users to identify the emotions observed in those movements to determine which movement patterns were correlated with which perceived emotions. Since robot morphology is different from human morphology, we chose this user-driven approach so we can later be more confident that the movements we generate for the robot will have the correct perceived emotion.

Quori (Figure 5.1) is a robot designed for and by the HRI community [87]. This robot has gesturing arms, a face that can project any image, and a waist joint to lean forward and backward, movement that has been shown to correlate with various emotional behaviors [14, 20, 21]. Our goal in this work was to develop nonverbal robot movements for the Quori robot to display specific emotions. Although emotional behaviors are composed of many modalities (speech, movement, facial expressions, etc.), we focused on body movements exclusively in this work to explore how even low degree-of-freedom (DOF) motion can convey emotions.

We first focused our attention on the types of torso and arm movements in humans in the literature that are correlated with the displaying of different emotions. Table 5.1 summarizes our review of this literature with specific description of movements for each of the seven emotions. These are the six basic expressions of happiness, sadness, fear, disgust, anger, and surprise from [24], with the addition of interest (useful in an educational context).

5.1.1 Robot Movements

We developed a simulation environment using ROS/Gazebo that allowed us to command the robot's arms and torso to a desired position at a specified speed (maximum of 1 rad/s for all joints). As shown in Figure 5.1, the torso has one DOF (θ_1) and can lean forward to 0.47 rad and backward to -0.21 rad. Each arm has two co-located joints: a rotational shoulder joint with no joint limits ($\theta_{2,\text{left/right}}$) and a lifting shoulder joint with limits ± 1.1 rad ($\theta_{3,\text{left/right}}$).

To constrain the large number of possible movements created by combining these DOFs, we developed a set of design considerations. For torso motion, we only created

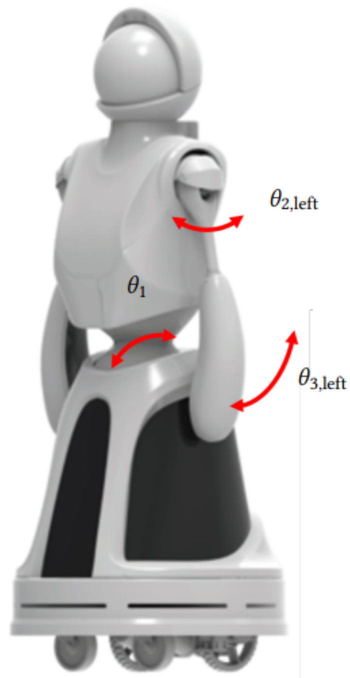

Figure 5.1: Torso and left arm angles labeled on the Quori robot (image from http://www.quori.org/)

movements starting at a neutral position ($\theta_1 = 0$) and ending forward or backward at a small or large angle. The torso can move between its start and end position at a slow or a fast speed. The discretization of the possible end positions to 4 possibilities and the speeds to 2 levels allowed us to reduce the number of possibilities but still include a variety of torso movements.

When choosing possible arm movements, we first determined whether the movement would be symmetric or asymmetric. For symmetric movements, both arms would start forward or at the robot's sides and could end either forward, at the robot's sides, or above the robot's head. This motion could be performed at a slow or a fast speed. For asymmetric movements, the left arm was stationary either forward, at the robot's side, or above its head. The right arm could be stationary in either the forward or side position. It could also move from a forward or side position to a forward, side, or high position, at a slow or fast speed. By only including asymmetric movements with the left arm stationary, we eliminated mirror image movement with the right arm stationary and the left arm moving. We believed that the emotion

Table 5.1: Torso and arm movements from literature associated with 7 emotions

Emotion	From Literature
Happiness	Symmetrical up-down motion of arms [21]; Hands kept high, hands made into fists and kept high [88]; Slight lean backwards, arms raised high [20]
Sadness	Leaning forward, Hands at sides [14]; Hands over head [88]; Leaning forward, Hands at sides [20]
Fear	Leaning backward [21]; Hands out to sides [14]; Body backing, Hands over head, trying to cover body [88]; Leaning backward, Arms slightly forward [20]
Disgust	Leaning backward, Arms forward [20]
Anger	Leaning forward [21]; Leaning forward, Arms crossed, on hips [14]; Hands on waist, hands into fist or low, fast hand lift [88]; Leaning forward, Arms forward [20]
Surprise	Hands over head [88]; Leaning backward, Hands over head [20]
Interest	Leaning forward, Arms resting at side [21]

perceived in the movement is agnostic to which arm is moving. We also chose not to include movements with both arms moving asymmetrically as this would exponentially increase the number of movements to evaluate.

After applying these constraints, we categorized the 224 resulting movements using the six properties shown in Table 5.2, each movement taking one value for each property. We decided not to analyze differences due to arm start position because we want to implement robot behavior that will go from a neutral position to an ending position and back as a nonverbal reaction, so that property is not listed in this table.

5.1.2 Movement Generation User Study

The goal of this study was two-fold. Firstly, we wanted to determine which movement properties were correlated with which emotions. Secondly, we wanted to compare the emotions perceived in Quori's movements with those seen in the corresponding human movements (Table 5.1).

Table 5.2: Movement Properties

Property	**Value 1**	**Value 2**	**Value 3**
Torso End	forward	backward	
Torso Degree	small	large	
Torso Speed	slow	fast	
(Arm) Symmetry	true	false	
(Right) Arm End	forward	sides	high
(Right) Arm Speed	slow	fast	

Movement Generation Study Design

We developed a study to test whether these 224 movements were perceived as displaying emotion. This study, created with Qualtrics and administered using Amazon Mechanical Turk, presented 10 randomly selected videos to 145 participants. Due to randomization, each video was not seen the same number of times but on average was seen 6-7 times. The participants were mostly in the 25-44 age range.

For each video, we asked the participants "How much is the robot expressing each emotion?" for the 7 emotions shown in Table 5.1 plus *Neutral*. For each of the 8 emotions, participants chose from a 5-point Likert scale with labels *Not, Slightly, Somewhat, Moderately*, and *Intensely* to indicate their perceived intensity of that emotion in the robot movement (Figure 5.2). We chose to collect the participants' perceptions using a set of Likert scales to allow for multiple emotions to be seen at different intensities in the same robot movement. This also allowed for no or very little emotion to be seen in a movement, as opposed to other question styles that may force the participant to choose a single emotion seen in a movement.

Before this set of 10 questions, each participant confirmed consent to participate in the study and saw a training video (not from the 224 videos) to practice answering the Likert questions. At the end of the survey, we asked a series of demographic questions, including age, gender, ethnicity, and familiarity with robots.

To determine whether a movement on Quori is perceived as conveying the same emotion as a similar human movement (Table 5.1), we needed to determine the movement properties with which a particular emotion was associated from our survey results. We first grouped the movements by value for a specific property. For the first property, we compared whether emotions were perceived differently between

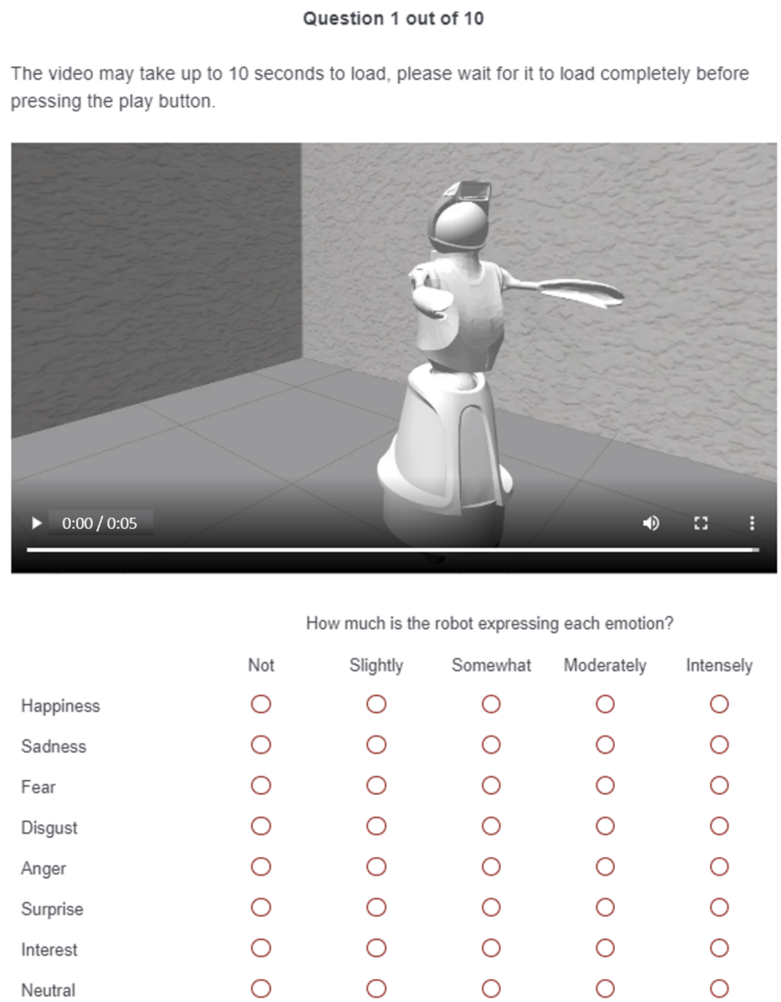

Figure 5.2: Each participant saw 10 randomly chosen videos out 224 and answered a set of Likert questions about their perception of emotion in the robot movement.

movements with the torso forward and those with the torso backward. To allow for the same participant to see multiple movements from the same group, we performed a Friedman test with a Neymeni post hoc that included a blocking variable (the participant). For each emotion, we compared the Likert values (as ordinal categorical variables) between groups. The result of this statistical test was a p-value, with a value below 0.05 indicating that the two groups were significantly different. We repeated these comparisons for each property.

Movement Generation Results

The results of comparing the distribution between values in each property are shown in Table 5.3. This table combines two types of information: (1) the results of the significant survey with p-values below 0.05 and (2) the information from the literature review (Table 5.1).

We see that certain emotions have many more significant results from the survey than others (black, unitalicized text). *Surprise* and *Sadness* each have many properties with significant results; *Happiness* and *Fear* have few properties; and *Disgust*, *Anger*, *Interest*, and *Neutral* have either one or no discriminating properties. This may point to the need for additional DOF or variety in the robot movements to display these emotions perceptibly. It may also indicate that emotions that had few significant properties are generally more difficult to recognize due to the variety of their expression in humans due to cultural or contextual differences.

We were interested to see where the results of the survey and the literature agreed and disagreed. As seen by the large number of green squares in the table, there were many cases where these two agreed, especially in the end position of the arms and the torso end position. Only in one case did the two disagree (red background), which is the end position of the arms to display *Interest*. This disagreement could be due to the large variety of behaviors associated with the human display of *Interest*. For example, raising one's hand (arm position high) and reaching out to shake someone's hand (arm position forward) could be considered displaying *Interest* in specific situations.

5.1.3 Movement Generation Discussion

We presented a user-driven approach to evaluate perceived emotion in a large set of nonverbal behaviors on the humanoid robot Quori. We compared the results of our survey with human torso and arm movements in the literature that are shown to be associated with emotions and illustrated many parallels. We found that certain movement properties, such as the torso leaning backward or the arms ending high, are associated with particular emotions. We describe our implementation of these movements to generate affective movements with the robot in Section 5.3.1.

A possible limitation of this work is that the simulated robot used here may have a different effect compared to the physically present robot. Despite this limitation,

Table 5.3: The value in each cell indicates the value of the property (Table 5.2) that was either described in the literature (Table 5.1) to be associated with a particular emotion (blue, italicized text) and/or had the highest Likert score compared to movements with a different value of that property (black, unitalicized text).

Key:	Literature and Survey agree	Literature and Survey disagree	
	Survey	*Literature*	

	Torso End	**Torso Degree**	**Torso Speed**
Happiness	*Backward*		
Sadness	Forward (p = 0.001) *Forward*	Large (p = 0.001)	
Fear	*Backward*		Fast (p = 0.073)
Disgust	*Backward*		Fast (p = 0.007)
Anger	*Forward*		
Surprise	Backward (p = 0.001) *Backward*	Large (p = 0.01)	
Interest			
Neutral			

	(Arm) Symmetry	**(Right) Arm End**	**(Right) Arm Speed**
Happiness	Symmetric (p = 0.037) *Symmetric*	High (p = 0.001) *High*	
Sadness	Forward, Sides (p = 0.001) *Forward, Sides, High*		Slow (p = 0.004)
Fear	Forward (p = 0.001) *Forward, Sides, High*		
Disgust		*Forward*	
Anger		*Forward*	*Fast*
Surprise	Symmetric (p = 0.006)	High (p = 0.001) *High*	Fast (p = 0.005)
Interest	High, Forward (p = 0.002) *Sides*		
Neutral			

our work showed strong correlations between what is found in the literature and how people perceive emotion in our humanoid robot.

5.2 Sorting Game Task Design

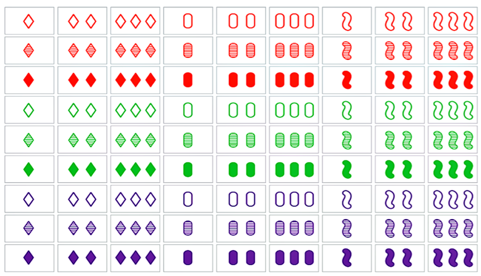

Figure 5.3: All 81 cards (left) in the sorting game (right) belong in one of two bins.

To test the effect of affective behavior on learning, we first created an educational task, specifically a sorting game. In this game (Figure 5.3), there is a rule that defines which cards belong in which of two bins (left or right). The purpose of the game is for players to sort a series of cards and infer the rule based on discovering which cards belong in which bin. The cards (taken from the game Set[1]) have 4 properties (color, number, shading, and shape), and each property has 3 possible values, resulting in a total of 81 cards.

The set of all possible rules is very large, but only some of those rules are reasonable for a human to understand. For example, a rule that randomly sorts cards into the two bins would be technically possible, but there is no pattern for a human to infer. Humans tend to see patterns even in random events (apophenia) and tend to make generalizations based on a small sample group [29].

We therefore define rules of two different forms, which we will call *easy* and *difficult*, which have clear patterns for the human to recognize. An *easy* rule sorts cards based on a single property; for example, "all diamonds go in the left bin and all squiggles/ovals go in the right bin" uses the shape property. A *difficult* rule sorts cards based on two properties; for example, "all red diamonds and green/purple

[1]https://www.setgame.com/welcome

ovals/squiggles go in the left bin and green/purple diamonds go in the right bin" uses both the shape and color properties. We did not choose rules with more complex patterns than *difficult* rules because we believed it would be difficult for a human to infer a greater rule complexity with the limited number of cards they would be presented with.

5.2.1 Example Rule Inference Process

For example, consider the rule "all green cards belong in the left bin, and all others in the right bin" and see how a human could infer this rule. Throughout this example, for simplicity we assume that the human considers only *easy* rules. If the first card the player sees is green-one-diamond-solid (top row, Figure 5.4), they have no prior information about the rule and will simply guess a bin randomly. From the rule, we know that the card actually belongs in the left bin, and the game will indicate that as the correct placement.

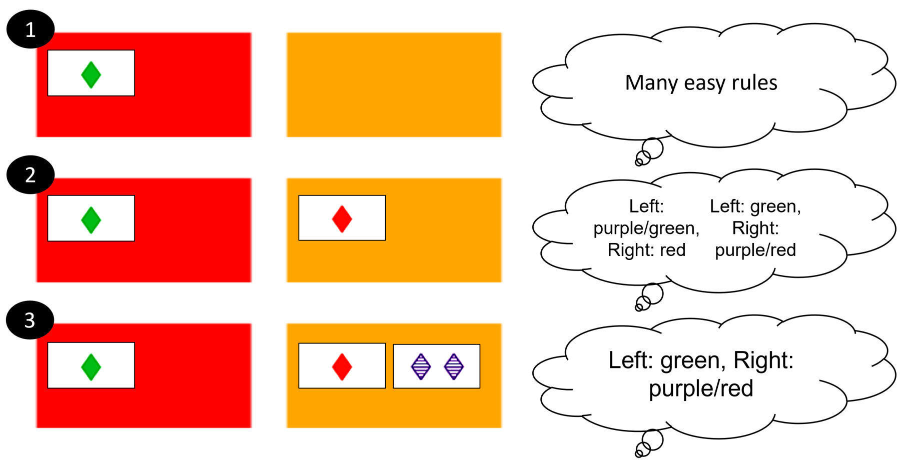

Figure 5.4: Example of how a human could infer the rule that states "all green cards belong in the left bin, and all others in the right bin"

When the next card, for example, red-one-diamond-solid (middle row, Figure 5.4), is shown to the player, they still do not have a clear idea as to the rule, but let us say that they guess that the rule is "all diamonds in the left bin, all others in the right bin." Using their guess of the rule, they will sort this card into the left bin. However, the true placement is in the right bin and the game will indicate as such. Now the player has a clearer model of the rule. They will see from the first two cards that the only property that can define the rule is color. In fact there are only two possibilities as to the *easy* rule, shown in the thought bubble in middle row of Figure 5.4.

If the next card presented is 'purple-two-diamond-striped' (bottom row, Figure 5.4), the player still has an equal chance of getting this question correct, since the two possible rules sort the card into two different bins. However, after the game indicates the correct placement of the card (right bin), the player can now infer the rule, having eliminated the first possible rule. They should now be able to sort every following card correctly and only needed those three cards to correctly infer the rule.

There are three important takeaways from this example. First, when the list of possible rules is only of the *easy* variety, it is much simpler to keep track of and eliminate possible rules, when given example cards. Looking at *difficult* rules as well, we anticipated that this process would be more difficult for humans and that it would take longer for them to learn the rule. The second is the importance of the order in which the cards are presented to the players. Certain cards will be more informative than others when it comes to eliminating possible rules, and uninformative cards can lead to it taking much longer to learn the rule. In the next section, we discuss how to generate an 'optimal' card order that will try to help students as much as possible in learning the rule quickly. And finally, the player is essentially guessing a bin randomly for the first few cards, so it may be advantageous to give them a head start by sorting the first few cards so that the accuracy of their sorting is more correlated with their learning of the rule.

5.2.2 Choosing a Card Order

This game was constructed so that players could see a series of cards, which helped them to learn the rule. We designed an algorithm to choose an informative card order to present to players that maximizes information gain (Algorithm 1). We first define a set of hypotheses that contains all possible rules that are of the *easy* (24) or *difficult* (432) form, for a total of 456 possible rules.

Given a rule to be learned, the algorithm begins by looking at all the cards that have not yet been played and finds the number of hypotheses eliminated by each card. A card *eliminates* a hypothesis if the true rule and the hypothesis sort the card

Algorithm 1 Generating a card order to maximize the number eliminated hypotheses to learn a rule

1: r, N $\triangleright$ True rule, Number of cards to generate
2: $\mathcal{H} = \mathcal{H}_e, \mathcal{C}$ $\triangleright$ All easy hypotheses, all cards
3: $\mathcal{Q}$ $\triangleright$ Card order
4: **if** $|\mathcal{H}| = 0$ **then**
5: $\quad \mathcal{H} = \mathcal{H}_d$ $\triangleright$ Add in difficult hypotheses only if all easy have been eliminated
6: **end if**
7: **for** $i \in \{1, 2, 3, ..., N\}$ **do**
8: $\quad$ **for** $c \in \mathcal{C}$ **do**
9: $\quad\quad \mathcal{H}_c \leftarrow f(r, \mathcal{H})$ $\triangleright$ Hypotheses eliminated by c
10: $\quad\quad q_c \leftarrow |\mathcal{H}_c|$
11: $\quad$ **end for**
12: $\quad m \leftarrow \max_c q_c$ $\triangleright$ Max hypotheses eliminated
13: $\quad \mathcal{C}_m \leftarrow \{c \in \mathcal{C} s.t. q_c = m\}$ $\triangleright$ Cards eliminating max hypotheses
14: $\quad$ **if** $|\mathcal{C}_m| = 1$ **then**
15: $\quad\quad c' \in \mathcal{C}_m$ $\triangleright$ Pick best card
16: $\quad$ **else**
17: $\quad\quad c' \in g(\mathcal{C}_m, \mathcal{Q}, r)$ $\triangleright$ Card sorted into bin with least cards
18: $\quad$ **end if**
19: $\quad \mathcal{Q} \leftarrow [Q, c']$ $\triangleright$ Add to card order
20: $\quad \mathcal{C} \leftarrow \mathcal{C} \backslash c'$ $\triangleright$ Remove chosen card from possible cards
21: $\quad \mathcal{H} \leftarrow \mathcal{H} \backslash \mathcal{H}_{c'}$ $\triangleright$ Remove hypotheses eliminated by card
22: **end for**

into different bins (Lines 8-10). Next, we find the maximum number of hypotheses eliminated by a card (Line 12).

If only one card eliminates this maximum number, it is chosen to be presented to the player (Lines 14-15). If multiple cards eliminate the maximum number of hypotheses (Lines 16-17), we then choose a card randomly from the choices, where the probability of choosing a card is based on the number of cards we have already seen that belong in each bin. This ensures that the number of cards in each bin is roughly balanced for the entire card order, and the player is less likely to receive several cards in a row that belong in one bin.

After choosing the 'most informative' card, it is added to the card order for the rule, it is removed from the set of remaining cards, and all hypotheses that it eliminates are removed from the set of hypotheses (Lines 19-21). The process is repeated by evaluating all remaining cards against the remaining hypotheses to choose the next most informative card. Note that we initialize the set of hypotheses to be considered as all the *easy* hypotheses, and we only add the *difficult* hypotheses if all the *easy* hypotheses have been eliminated (Lines 4-5).

In both the rule inference process and the card order algorithm in Lines 4-5, we make the assumption that the human will consider only *difficult* hypotheses when all *easy* hypotheses have been eliminated. We initially did not include this assumption when generating the card order for the following study. When we saw the results in Figure 5.7, the way humans performed did not match how we assumed they would. We then added this assumption of only considering *easy* hypotheses first into the learner model, resulting in a prediction of results much closer to how people performed. This means that the card order generated for the user study was not truly an 'optimal' order, but was 'optimal' considering our incorrect assumption of the way humans would eliminate rules. In the remainder of this chapter, we will call **optimal** the learning model/card order that takes the *easy* hypotheses into account first, and the **suboptimal** the learning model/card order that assumes that all *easy* and *difficult* hypotheses are taken into account at the beginning of the learning process.

This method of choosing cards is specifically designed to mimic how we believe people would play this game. We hypothesized that they would have certain beliefs about the possible rules and that as cards are presented, they would remove rules from that set as those rules are violated. Choosing a card order based on removing the most hypotheses the fastest should allow the players to quickly arrive at the true rule.

5.2.3 Learner Model

Once we developed the card order algorithm, we could use the underlying assumed human inference process to define a learner model for how a player would learn the rule. This allowed us to simulate how we thought people would perform given a rule and compare these simulated results to the actual human results.

When a simulated learner sees a card, they look at all the remaining hypotheses and choose the bin to which the card belongs by choosing a hypothesis randomly from that set. After seeing the true sorting of a card, the simulated learner will eliminate all hypotheses that incorrectly sort that card from their set of remaining hypotheses. With a smaller set of hypotheses, the simulated learner should improve over time, eliminating hypotheses after each card is revealed in its correct place, until only the correct rule remains. Being able to simulate how learners will play this game allowed us to compare this modeled performance to how humans actually learned the same rule, and validate the assumptions made when generating an 'optimal' card order. Note that we are ignoring possible human biases here, such as having a bias towards rules based on color rather than shape. We compared how simulated learners and actual human learners performed with a fixed card order and sorting rule in the following user study as a way to validate our learning model. Note again that we implemented the **suboptimal** learning model for the study, but we used the **optimal** learning model for our post hoc analysis, so our simulated results matched the actual results more closely.

5.3 Affective Nonverbal Feedback User Study

The goal of this user study was to determine whether the robot's affective feedback had a positive impact on human learning during this card sorting task. We compared feedback between *neutral* behavior (no nonverbal reaction to correct/incorrect responses) and *matching affective* behavior (happy reaction to correct responses and sad reaction to incorrect responses). Our hypotheses were as follows:

- **H1:** Participants will learn the rule better with the *matching affective* robot.
- **H2:** Participants will have a more positive subjective experience of the game with the *matching affective* robot (more engaged, lower perceived game difficulty, higher perceived learning).
- **H3:** Participants will have a more positive perception of the *matching affective* robot (higher intelligence and animacy).

5.3.1 Affective Nonverbal Feedback Study Design

Robot Feedback

We designed two types of nonverbal robot behaviors for this study – *neutral* and *matching affective*. The behaviors were implemented using a Gazebo [50] simulation of

Quori [87]. For the *neutral* behavior, the robot performed a slight random movement of its joints. This was also used while the participant was deciding what to do during the game. We chose to have the *neutral* robot move and not remain stationary to still impart some animacy to the robot, so the differences we saw would be based on the *matching affective* behaviors rather than whether the robot was moving at all. For the *matching affective* behavior, we wanted the robot to display an emotion that was correlated with the correctness of the participant's guess. We used the correlations found in Section 5.1 to determine how the robot should move. Specifically, happiness was correlated with backward torso movement and symmetric arm raising, and sadness was correlated with forward torso movement and slow movement of one arm forward. We used these results to design nonverbal movements for Quori for correct (happiness) and incorrect (sadness) answers.

After the participant chose a bin to place the card, the robot turned toward the correct bin (turning only slightly in the *neutral* behavior), performed the nonverbal movement (neutral, happiness, or sadness), displayed text feedback, and returned to a neutral position. We generated a few slightly different movements for the correct and incorrect cases to have more variety in the movement options. We also generated a few slightly different phrases for the displayed text feedback. Some examples of feedback for correct answers and incorrect answers are:

- Correct: Nice work; Good thinking; Great work
- Incorrect: Hmm, not quite; Maybe think about the pattern in a different way

Figure 5.5 shows example screenshots from feedback videos when the correct bin is the left bin. The text feedback appears and the robot turns towards the correct bin (left) in all cases. In the *neutral* condition, the robot moves slightly and it performs larger movements in the *matching affective* condition. Text feedback appears in all cases, so participants would have no confusion as to whether their choice was correct, and any differences in performance would be solely due to the movement of the robot.

Demonstration and Trial Phases

Each round in the study consisted of two phases: the demonstration phase and the trial phase. In the demonstration phase, participants first saw two demonstration cards, determined by the **suboptimal** card order generation method (Section 5.2.2).

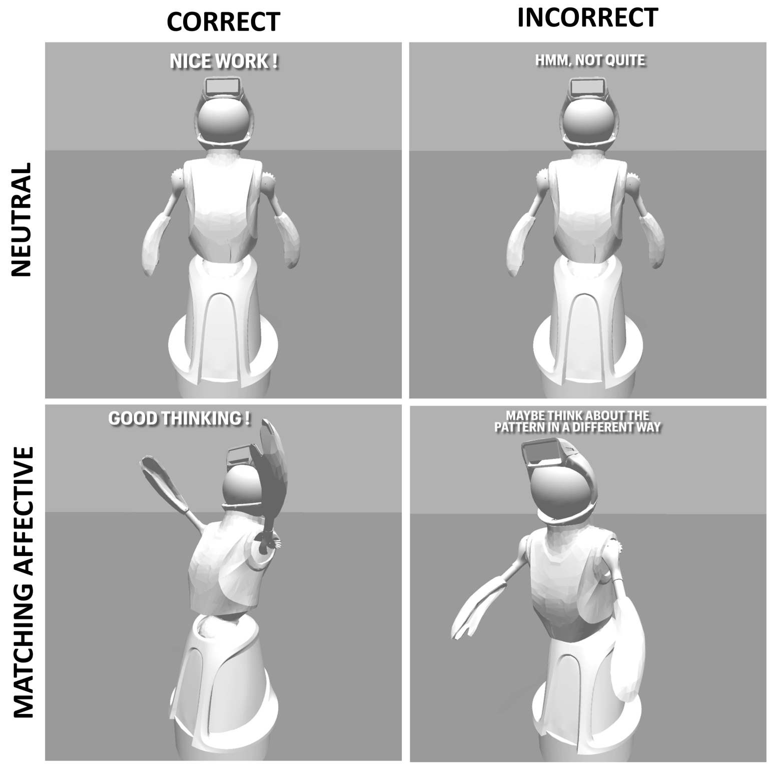

Figure 5.5: Example screenshots of the *neutral* and *matching affective* conditions when the correct bin is the left bin.

The participants were told that the robot would demonstrate where the two cards belonged according to the rule. This gave them a head start in learning the rule and established the robot as a teacher who would help them learn the rule. They were not told the difficulty options for the rule or which difficulty they were presented with.

In the trial phase (Figure 5.6), participants dragged a card from the gray staging area to the bin to which they believed it belonged (they could move it from one bin to another if they changed their mind). After clicking the "Submit Choice" button, the robot provided feedback (as discussed above) on their choice through a video, moved the card to the correct bin if it had been placed incorrectly, and then the next trial loaded. As the participant played the game, the cards previously seen (in demonstrations or previous trials) remained visible in the correct bins. This process repeated for 8 trial cards.

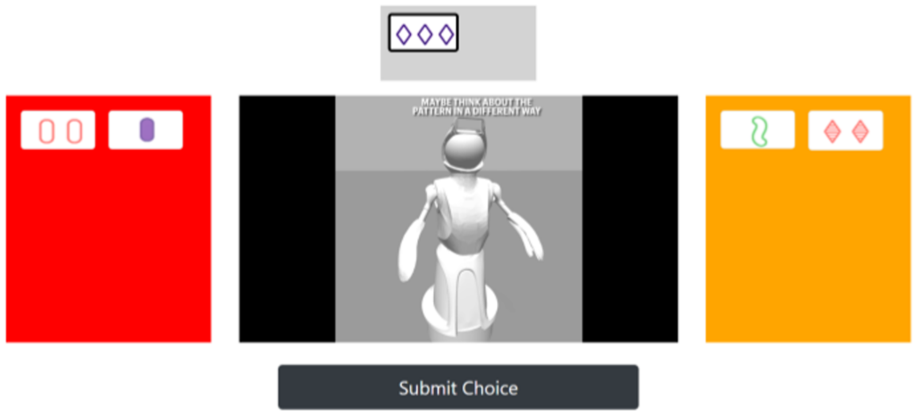

Figure 5.6: In the trial phase, the participant dragged the card from the gray area to one of the two bins. After clicking the "Submit Choice" button, the robot provided feedback on their choice and moved the card to the correct bin, if it was placed incorrectly.

5.3.2 Experimental Design

This experiment had two independent variables: **difficulty** (a choice of *easy* and *difficult*) and **feedback** (*neutral* and *matching affective*). Each participant played two rounds of the game. The participants saw either *easy* or *difficult*, and either *neutral* or *matching affective* for their first round (4 options). They then saw the other rule and either *neutral* and *matching affective* for their second round (2 options). This produced 8 conditions, which were designed to avoid any ordering effects. We treated the two rounds each participant played as independent trials. The *easy* and *difficult* rules (below) were the same for all players.

- Easy: Diamonds on the left, all others on the right
- Difficult: green-one, red/purple on the left, green-two/three on the right

The *easy* rule involved one property (shape) and the *difficult* rule involved two properties, color and number. All players saw a fixed card order for the two rules (see Figure 5.7 for the specific cards shown).

Participants completed a survey at the end of each round consisting of 5-point Likert questions to evaluate their experience. Game-related questions are listed below with scales from strongly disagree to strongly agree:

- (Engagement) I enjoyed playing the game

- (Perceived Difficulty) I thought this game was difficult
- (Perceived Learning) I feel that I learned the game well.

Questions related to the robot were taken from the Godspeed Questionnaire [8] and included 3 questions related to animacy and 2 related to intelligence. We treated these qualitative measures as numeric data with the lowest level as 0 and the highest level as 4. We averaged the 3 animacy questions to get an average animacy as well as the 2 intelligence questions to get an average intelligence. We also included an optional area for free-form feedback.

5.3.3 Affective Nonverbal Feedback Results

For this study, we recruited 160 participants on Prolific, who identified as 70% Female, 28% Male, 2% Other. Most were not at all, slightly, or moderately familiar with robots and had a bachelor's degree or less. Out of the 320 rounds, we removed the 4 outliers with an accuracy below 2/8 (2.67 standard deviations from the average accuracy).

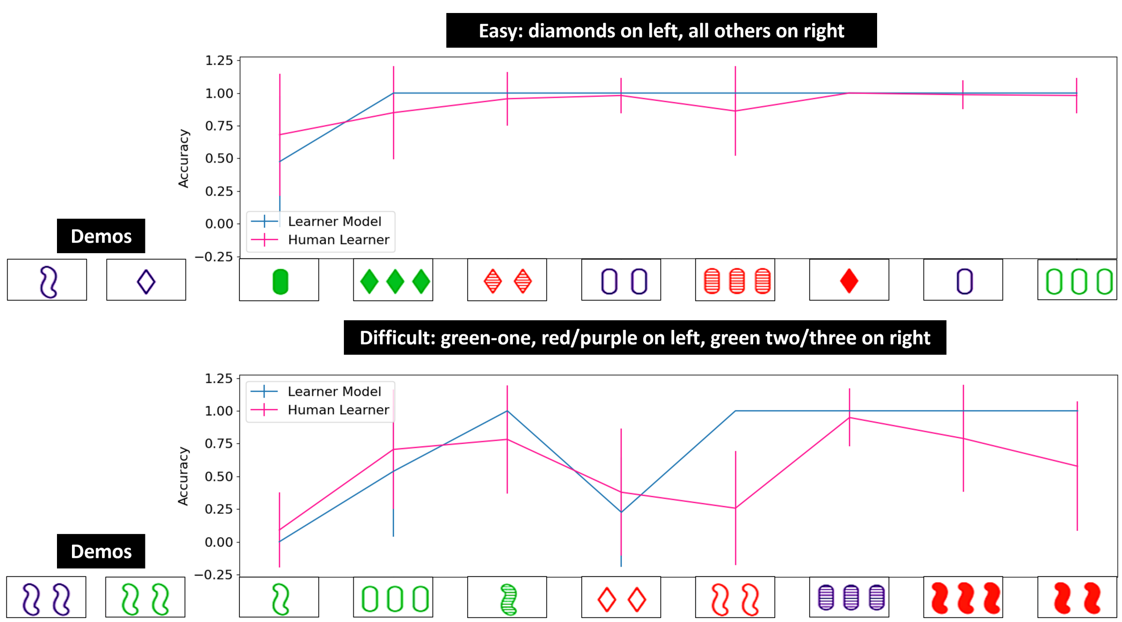

Figure 5.7: Participant performance compared to simulated learners for the *easy* and *difficult* rule. Demonstration cards and trial cards are shown, and the title of each plot indicates the rule being learned.

Figure 5.7 plots the average participants' performance for each of the 8 trial cards, separated by rule **difficulty**. One standard deviation around the mean is shown in the graph as well as the result of 80 simulated learners using the **optimal** learner model from Section 5.2.3. The variation in the simulated learners is because the learner chose a random hypothesis from its set of remaining hypotheses when deciding how to sort a previously unseen card. This random choice is what leads to simulated learners having different performances.

We can see that for the *easy* rule, participants seem to learn the rule around trial card 6, as the accuracy approaches 1 in that trial. In contrast, for the *difficult* rule, participants seemed to improve performance over time, but did not reach the high accuracy achieved with the *easy* rule. In fact, performance seemed to decline in later trials, though the standard deviation is quite large.

We also compared human performance with that of simulated learners. The simulated learners learned the *easy* rule faster than the human learners as their accuracy approached 1 after seeing 2-3 of the trial cards. For the *difficult* rule, the human learners and the simulated learners followed similar patterns of increased accuracy, followed by a dip at around trial cards 4-5. This dip could be due to the fact that the card shown was an "exception" to the *easy* rule that the participant had in their mental model, but since the true rule was *difficult*, they sorted it incorrectly and were then forced to consider the *difficult* rules. The human learners seemed to decrease in performance for the latter cards while the simulated learners learned the rule with 100% accuracy at around trial cards 5-6. This difference in performance could be due to the humans forgetting aspects of the rule (simulated learners have perfect recall of the hypotheses remaining) or fatigue/disinterest by the end of the round. It also could be that people were considering even more complex rules, perhaps considering rules that varied 3 different properties.

Improved Learning

We conducted two-way ANOVA tests with the two independent variables (**difficulty** and **feedback**). Figure 5.8 illustrates the results with significant differences marked.

An objective measure of performance in the game is the accuracy in the trials, which can be a proxy for learning. The accuracy significantly dropped when the

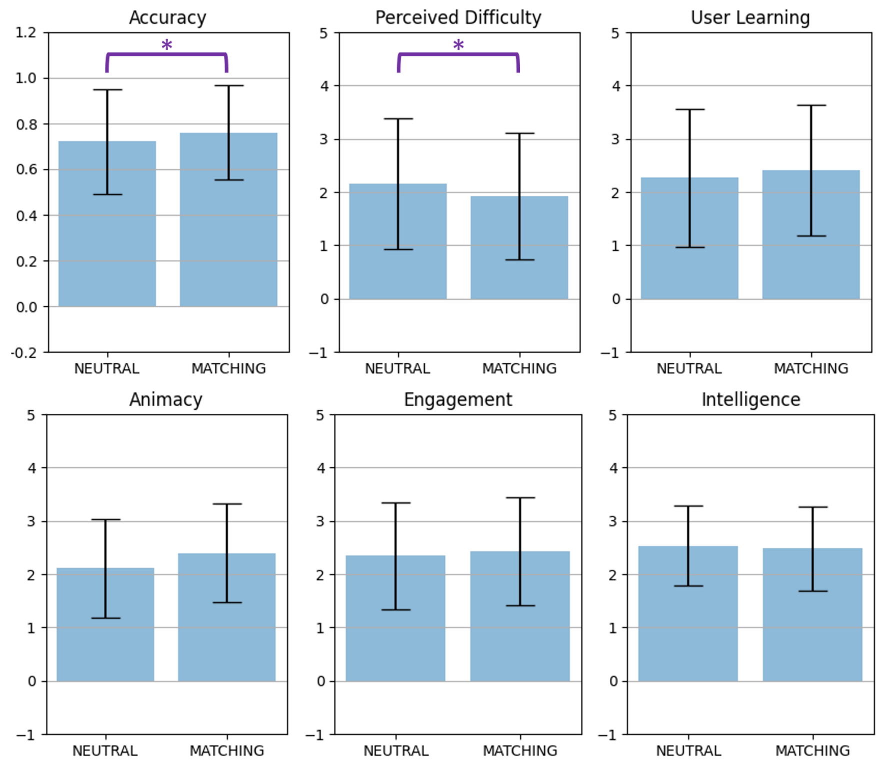

Figure 5.8: Two-way ANOVA tests with significant results indicated. The asterisks indicate a significance of $p < 0.05$.

difficulty of the rule increased $(F(1, 312) = 543.23, p < 0.001)$, which is intuitive. However, more importantly is the result that participants who viewed the *matching affective* robot feedback had a significant increase in accuracy over the *neutral* feedback $(F(1, 312) = 4.28, p < 0.05)$, shown in the top row of Figure 5.8.

The interaction between the two independent variables was also significant $(F(1, 312) = 4.04, p < 0.05)$, which supports a dependence between **difficulty** and **feedback** when determining accuracy. We performed further analysis with a Tukey test, shown in Table 5.4. An interesting result was that accuracy was significantly different for the *difficult* rule between the two robot types, but that same trend was not seen for the *easy* rule. This indicated that the *matching affective* behavior of the robot aided more when the rule was *difficult* compared to when it was *easy*.

An alternate explanation for this result is that the *matching affective* robot

Table 5.4: Tukey Post-hoc test results on accuracy. Significant differences when holding one variable constant and varying the other variable are shown at the $p < 0.05$ (*) and $p < 0.01$ (**) levels.

	Neutral vs. Matching
Easy	
Difficult	*

	Easy vs. Difficult
Neutral	**
Matching	**

engaged the participants more with more movement and caused them to pay more attention to the game and learn the rule better. To test this, we first developed *nonmatching affective* behavior, where the affect displayed by the robot was opposite to the context. If the participant answered the question correctly, the robot would say a positive message such as 'Nice work!' but the movement it performed would correspond to the sad affect.

If we found the same difference in accuracy with the *nonmatching affective* robot, then we could conclude that the improvement in accuracy was due to the increased movement rather than the matching affect. We developed a secondary study that compared *neutral* to *nonmatching affective* using the same methodology as our primary study. We found no significance in accuracy in this secondary study, so we can conclude that the *matching* aspect is vital to the improvement in accuracy, not just the increase in overall movement.

Subjective Experience of the the Game

Participants felt they learned better ($F(1, 312) = 146.76, p < 0.001$), were more engaged ($F(1, 312) = 21.04, p < 0.001$), and the rule was easier ($F(1, 312) = 161.22, p < 0.001$) with the *easy* rule compared to the *difficult* rule. With the *matching affective* robot, participants perceived the difficulty as lower (Figure 5.8), compared to the *neutral* robot ($F(1, 312) = 3.60, p < 0.05$).

Subjective View of the Robot

Participants did not have clear subjective opinion differences about the robots. Although they viewed the *matching affective* robot as more animate, this difference was not statistically significant.

5.3.4 Affective Nonverbal Feedback Discussion

We have support for **H1**, as the accuracy with the *matching affective* robot was significantly higher than with the *neutral* robot. The *matching affective* movements of the robot seemed to help in learning the rule, and this same effect was not observed with the *nonmatching affective* movements. Although both robots were communicating the same information about the task (whether the previous card was sorted correctly), the nonverbal robot movements helped the participants learn the rule better. Our post hoc analysis of the interaction effect gives us additional information that this assistance from affective feedback is particularly useful with a higher task difficulty. This indicates that nonverbal behavior can improve objective task performance.

We also have partial support for **H2**, as participants had a lower perceived difficulty with the *matching affective* robot. The rule seemed easier even if actual **difficulty** was not higher, another indication that the robot's nonverbal behavior seemed to help in the learning process. One possible explanation for this is that the nonverbal behaviors temper the frustration that participants felt when learning the rule was difficult. Including *matching affective* movements, therefore, can not only help with task performance but can also improve subjective experience.

We did not find support for **H3**, as participants did not report a significantly different perception of the two robots. We however show in the following chapter that having a physical robot providing affective feedback does have the desired impact of an improvement in the perception of the robot.

Lastly, when examining the qualitative responses that people provided after each sorting game round, we can see that some people did not see differences in the *neutral* and *matching affective* robots, but they still had a better performance with the *matching affective*. But some participants did mention that they saw the *matching affective* movements as happy or sad (depending on the situation), saying that the felt the robot was 'celebrating' with them after a correct answer. This shows us that even if the robot's nonverbal behavior is not consciously noticed, it can still have a positive impact on the human's performance.

In the next chapter, we address the limitations of this work in two ways. By taking a more complicated task that has more opportunities for nuanced context (beyond simply correct vs. incorrect), the robot will have more opportunities to personalize its behavior. By performing a study in an in-person setting, we can leverage the improved impact of real-time feedback, and we also improve the nonverbal feedback by implementing facial expressions on the robot.

6 Feedback Style Preferences

Summary: Different people respond to feedback and guidance in different ways. Their preferences may even change depending on their mood, fatigue, physical health, etc. We present a robot exercise coach that provides both verbal and nonverbal feedback. We first introduce an exercise evaluation method where the camera feed from the robot is used to evaluate how well people perform exercises. We then present a multimodal feedback controller that uses the exercise evaluation to respond with verbal and nonverbal feedback in different styles (**firm** and **encouraging**). We also evaluate people's preferences for the frequency of verbal and nonverbal feedback during exercise to determine the appropriate feedback cadence for the robot. Our user study found that participants have significantly different performances and subjective experiences with the different styles. These results show that varying feedback styles has an impact and builds the basis for a robot that adapts its style in real time to personalize to the individual presented in Chapter 7.

The contents of this chapter were published in [42, 43].

6.1 Exercise Evaluation

For the robot to provide relevant real-time feedback, it needs to analyze the exercises the human performs. We chose two exercises, bicep curls and lateral raises (Figure 6.1, because they are upper body exercises and have simpler form corrections compared to more complex exercises, such as squats or lunges. However, the methodology that we present is generalizable to other strength training exercises.

Figure 6.1: Bicep curls (left) and lateral raises (right)

The camera mounted on the robot Quori is the Orbbec Astra Mini[1]. We use the Mediapipe library[2] to get 3D joint positions (left elbow, right elbow, etc.) from the camera feed. We then compute the angles between three different positions on the body that are relevant for the exercises. The combinations we chose for these two exercises are: (shoulder - elbow - wrist) and (hip - shoulder - elbow). The robot computes these angles along specific axes (xy, yz, and xz) and for the left and right sides, resulting in a total of $2(3)(2) = 12$ angles. We chose to use the 2D angles, as they result in less noise than a computed 3D angle.

An example of the angles calculated for several sets of bicep curls is shown in Figure 6.2. The repeated pattern at certain angles is clear, especially in the shoulder-elbow-wrist-xy angles, where the elbow angle increases when the forearm extends down and decreases when the forearm moves upward. We also see similar patterns at various angles for lateral raises, where the shoulder angle increases as the arm raises and decreases as the arm lowers.

We divide the exercise evaluation process into two steps: segmentation and comparison. The segmentation step determines the start and end of a repetition of the exercise. The comparison step compares each repetition to recorded reference demonstrations of the exercise, so that the robot can provide corrective feedback.

[1]https://shop.orbbec3d.com/Astra-Mini-S
[2]https://pypi.org/project/mediapipe/

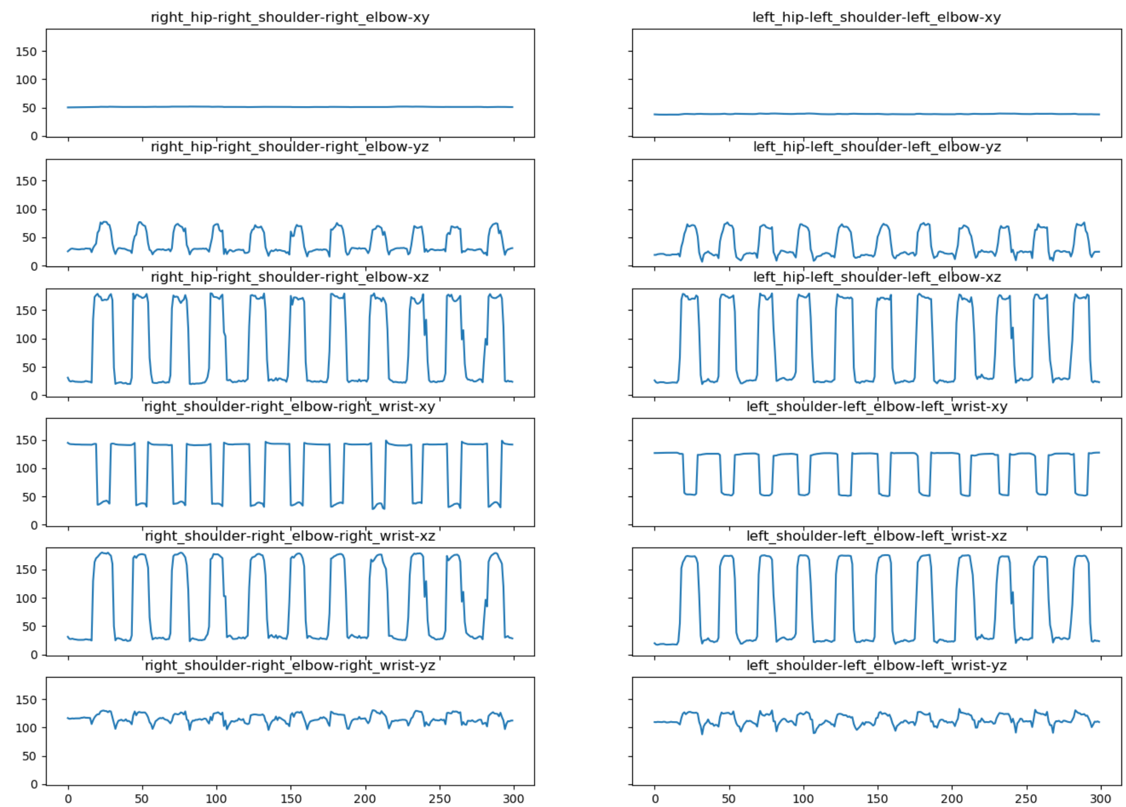

Figure 6.2: Angles computed during bicep curls

6.1.1 Segmentation

Given a series of bicep curls (Figure 6.2), the robot must find where each rep starts and ends. We first need to determine the angles used to segment the reps. For example, the hip-shoulder-elbow-xy angle is not informative, but the shoulder-elbow-wrist-xz angle is (due to the human's positioning in relation to the camera and the nature of the exercise). Examining the data, we chose to segment based on both the hip-shoulder-elbow-xz and shoulder-elbow-wrist-xz, for both the left and right sides.

The algorithm for segmenting the joints is based on finding peaks in the gradients of the chosen angles, where people change the direction of their arms. After finding the peaks, the algorithm performs some additional checks to ensure that peaks found are not too close together and correspond to the same point in the rep (e.g. the beginning rather than the end of a rep). The results of such a segmentation are shown in Figure 6.3 for one of the angles. This algorithm is able to find the end of one rep and the beginning of the next rep in real-time. This method is not 100% accurate, sometimes finding false positives and false negatives. However, since the robot does not give corrective feedback after every repetition, any mistakes in the robot's feedback (based on a pattern of behavior rather than one rep) are not evident from the robot's feedback (our main focus).

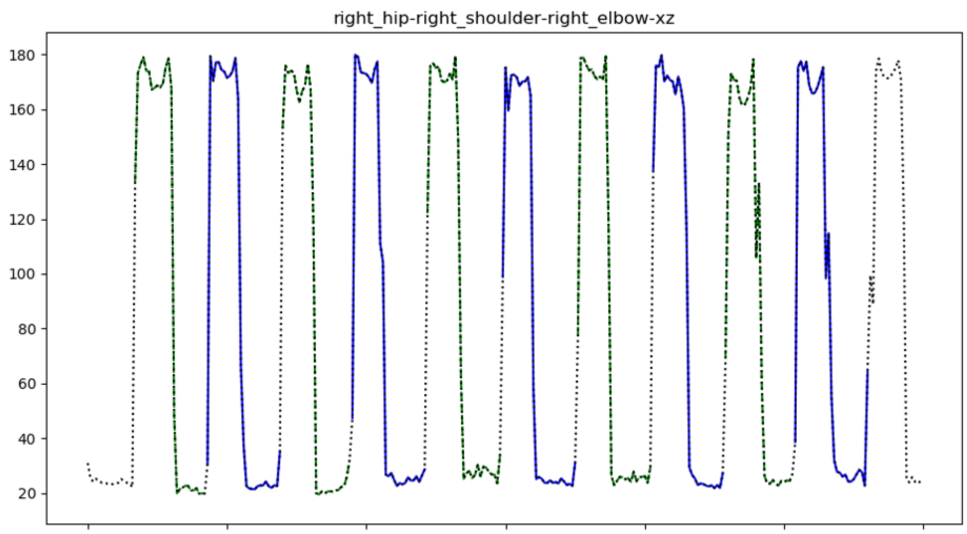

Figure 6.3: One of the bicep curl angles segmented into repetitions with the alternating solid blue and dashed green colors indicating where the segmentation method found the end of one rep and the beginning of another

6.1.2 Comparison to References

After the robot isolates a repetition, it evaluates how well it was performed. To do this, it needs some baselines to use for comparison. We recorded a set of reference demonstrations and ran the segmentation algorithm to obtain a series of reference reps for each exercise. We used demonstrations from three different individuals (researchers associated with the laboratory) who had varying exercise experiences to form this set. This allowed for some natural variation to be included in the demonstration set, as each person may perform the exercises slightly differently but not less correctly.

Each reference rep is labeled either *good* or *bad*. Demonstrations include *bad* form executions of the exercise, with specific common mistakes, so that if a new rep matches a *bad* reference, the robot will have a concrete suggestion to make to fix the issue. For example, if a new rep matches a reference labeled with 'low range of motion on the shoulder joint' for lateral raises, the robot can tell the individual to raise their arms to 90°.

The robot splits the angles of interest for the exercise into groups: right elbow, right shoulder, left elbow, and left shoulder. The robot evaluates each group individually. It compares to references for each group, which allows someone to perform the motion of their right elbow during bicep curls similar to one reference and their left elbow similar to another.

The exercise coach compares new repetitions with references in two ways. First, it compares the speed: is the new repetition fast or slow? If the length of the new rep is within 3 seconds of the average good reference length for that exercise (approximately within 2 standard deviations of the reference length distribution), the robot concludes that it has good speed. Otherwise, it concludes that it is either too slow or too fast.

Second, the robot compares the form: is the new rep close in form to the references? Even if two reps have different time scales (one is slower/faster), their overall shape can still be similar. We use the Fast Dynamic Time Warping (DTW) algorithm [77] to calculate the distances between the new repetition and each reference and take the minimum distance to be the closest reference.

Given the calculated minimum distance, we still need to determine whether that distance is *close enough* to conclude that the new rep has the same label as the closest reference. To do this, we look at the distribution of pairwise distances between reps in our demonstration set that had the same label and different labels. Naturally, comparing two reps that had the same label would result in a lower DTW distance than comparing two reps with different labels, but due to the natural variation in how people perform exercises, the distance between two reps with the same label is typically not zero. Looking at those distributions for both exercises, we set the primary thresholds to 1500 for bicep curls and 1700 for lateral raises. This means that when comparing a new rep to a rep in the demonstration set, we conclude that it has the same label as a demonstration rep if the distance between the two is less than the primary threshold for that exercise. We also have a secondary threshold of 2000 for the case when the new rep is 'somewhat close' to a reference.

Due to computation limitations and the large set of demonstrations (approximately 40 reps per exercise), it was not feasible for the robot to compare each new rep to all 40 demonstrations in real time. Instead we chose a few random demonstrations from the demonstration set from each label to compare the new rep to, thus allowing the robot to complete the comparisons in a timely manner and respond to the human in real-time. There is a trade-off between the accuracy of evaluation (using more demonstrations for comparison) and the timeliness of the feedback (using less demonstrations for comparison), and we experimentally determined the most number of comparisons we could make before compromising the speed of the robot's response. In practice, we found that around 3 comparisons per label (9 comparisons total per repetition) allowed timely feedback by the robot [3].

Table 6.1 shows the correspondence between the distance to the closest reference, the corrective feedback, and the numeric evaluation given to the joint group. If the minimum numeric evaluation over all joint groups is greater than or equal to 0, we say that the repetition has overall good form. In the next section, we will use patterns in the overall evaluation as well as patterns in specific mistakes the human makes to determine the appropriate feedback for the robot to give.

Table 6.1: Correspondence between the corrective feedback, numeric evaluation, and the distance to the closest reference

Closest Reference	**Below Primary Threshold**	**Below Secondary Threshold**	**Corrective Feedback**	**Numeric Evaluation**
Good	Yes	Yes	Good	1
Good	No	Yes	Ok	0
Good	No	No	Bad	-1
Bad	Yes	Yes	Specific Feedback	-1
Bad	No	No	Bad	-1

[3]Quori has a nuc8i7hvk: Intel CoreTM i7-8809G Processor with RadeonTM RX Vega M GH graphics (8M Cache, up to 4.20 GHz). It has a 500 GB SSD and 16 GB of RAM

6.2 Feedback Generation

Once the robot evaluates a new repetition, it should react in a multimodal way. Two domain experts[4] provided guidance on the appropriate types of verbal and nonverbal feedback. Based on their suggestions, we designed a controller that reacts to the feedback generated by the exercise evaluation with three different feedback styles.

Table 6.2: Examples of verbal and nonverbal feedback to different situations for the three feedback styles. The **neutral** style does not have verbal feedback while the human is exercising.

	Firm	
Evaluation	**Verbal**	**Nonverbal**
Last 2 reps slow	Try to speed up	55% sad, lean forward slightly
Last 2 reps low range of motion	Focus on getting a full range of motion in your elbows	55% sad, lean forward slightly
Last 2 reps good speed, previous 2 were slow	Nice speed, keep going	52.5% happy, small upward arms, small backward torso

	Encouraging		**Neutral**
Evaluation	**Verbal**	**Nonverbal**	**Nonverbal**
Last 2 reps slow	Nice job, can you speed up a little on the next few?	25% sad, lean forward slightly less than **firm**	10% happy, random neutral movement
Last 2 reps low range of motion	You are doing great, try to get a full range of motion in your elbows.	25% sad, lean forward slightly less than firm	10% happy, random neutral movement
Last 2 reps good speed, previous 2 were slow	Nice job, great speed!	77.5% happy, large upward arms, large backward torso	10% happy, random neutral movement

[4] Ayotoni Aroyo, ACSM-CPT (Exercise Physiologist and Physical Activity Lead at Emory University's Cognitive Empowerment Program) and Gustavo J Almeida, PT, Ph.D. (UT Health San Antonio)

The first as a **neutral**, baseline coach. This robot style does not respond to the exercise evaluation and serves as a baseline to determine whether people respond well to the contextually specific feedback provided by the other two styles.

The second coach has a **firm** style with minimal encouragement. And the third robot coach has an **encouraging** style withe encouragement included in all its feedback. For these two styles, the robot provides verbal and nonverbal feedback as the human exercises. The feedback these styles choose is based on the evaluation case, the exercise being performed, and how long it has been since the robot has spoken.

Table 6.2 includes examples of verbal and nonverbal feedback in a few situations for the three feedback styles. The following subsections explain how we generate the three modalities (verbal, face, and body) that Quori uses to react and how the feedback controller interacts with those modalities.

6.2.1 Verbal

We first determined how the robot should give verbal feedback during an exercise session. After discussions with our domain experts, we decided on 3 categories of verbal feedback:

1. Positive - α_1 good evals or α_2 good speed in a row
2. Negative - β_1 bad evals with the same message or β_2 fast/slow in a row
3. Improvement - β_1 bad evals with the same message followed by a good eval or β_2 fast/slow in a row followed by a good speed

where $\alpha_1, \alpha_2, \beta_1, \beta_2$ are values to be determined in Section 6.3, where we explore the appropriate feedback cadence for the robot.

The robot gives positive or negative feedback only if it sees a pattern (good or bad) in human behavior. We chose to structure the feedback in this way because our domain experts indicated the importance of ensuring that the robot sees a pattern before intervening. Additionally, they emphasized the importance of rewarding an improvement of behavior (third item above), so that the human knows that they have successfully incorporated the robot's correction.

The robot gives priority to form-related feedback over speed-related feedback, so when multiple feedback cases are detected by the exercise evaluation, the robot chooses what to say according to that priority. Additionally, the robot only says a message if it has not finished a previous message 3 seconds prior (experimentally determined) to not overwhelm the human with verbal feedback. Examples of specific phrases generated for each of the feedback styles are included in Table 6.2. The **firm** style does not have much encouragement and focuses on corrections and pushing the exerciser to do more. This is a style indicated by the domain experts, but they cautioned to not make a coaching style more strict than the **firm** style, as this could be perceived negatively. The **encouraging** style uses encouragement to soften corrections and leaves praise unqualified.

Google's Text-to-Speech library is used to actually generate the voice for Quori. We used an LLM to generate a large variety of responses that the robot could use. The code used to generate responses is included in Appendix A.2. In the work in this chapter, we only use a single fatigue level (low), but we use all three fatigue levels in the work in Chapter 7, when we introduce fatigue estimation. Additionally, we edited the responses provided by the LLM to remove duplicates and rephrase any responses that had awkward phrasing or did not match the prompt well.

6.2.2 Face

The goal for face generation is to have a dynamic face that is not hyperrealistic (to avoid the uncanny valley [63]) and that has the capability to display emotions at varying levels (e.g., slightly happy vs. very happy). We chose to use an existing face generation implementation[5] that takes an animator's approach to generating the facial expressions, by combining pulls from various facial muscles. Sliders for the six main emotions (happy, sad, anger, surprise, disgust, and fear) can be combined with values from 0% to 100% for each. Due to the musculature model, when the face transitions between two specified emotions, it moves smoothly between them. Figure 6.4 shows an example of the face that expresses happy and sad at two different levels (50% and 100%). We created our own Python implementation of this face that allowed us to add an additional feature that was not part of the original work: blinking. The robot face blinks with a randomized frequency, experimentally determined to have some variation, but not to have too little or too much time between blinks.

[5] https://www.mand3l.com/portfolio/facemotion/ by Paul Mandel

Table 6.2 shows examples of how the robot responds with facial expressions in the different feedback styles. The facial expression levels were chosen by experimentation and feedback from pilot participants. The **firm** style has a strong sad reaction to accompany corrections. The **encouraging** style has a truly neutral facial expression to accompany corrections (it has a 30% baseline happy expression, so this is still a change). For positive feedback, the styles have similar reactions, but the **encouraging** style has a more happy facial expression.

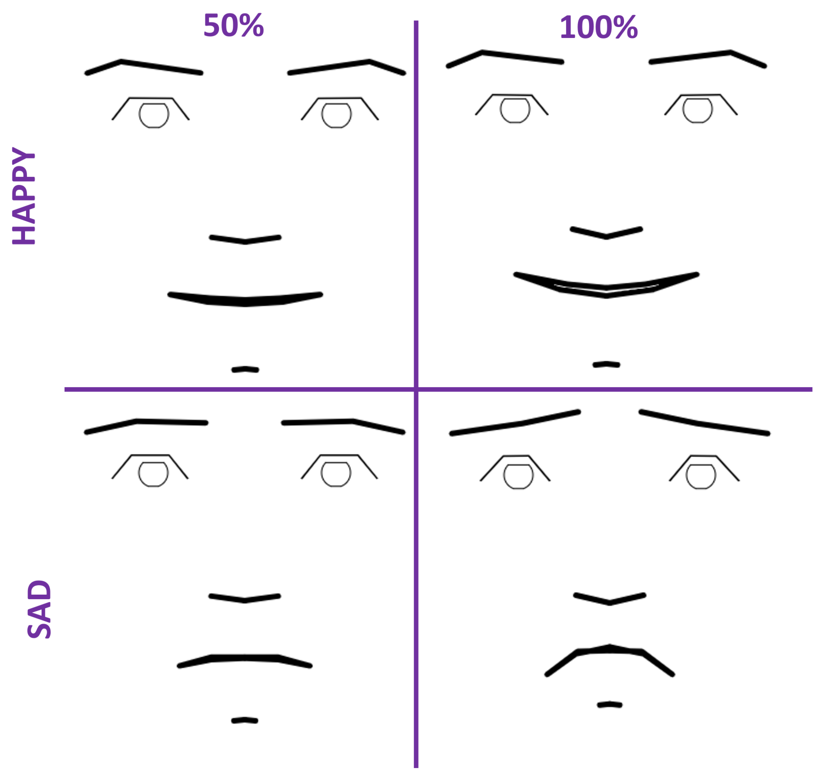

Figure 6.4: An illustration of happy (top row) and sad (bottom row) at two different levels of intensity: 50% (left column) and 100% (right column).

6.2.3 Body

We used the findings from Section 5.1 to create happy, sad, and neutral body movements for the robot. When the robot moves neutrally, it moves its joints slightly in a somewhat random fashion to appear alive and responsive. The robot moves both its arms upward and tilts its torso backward to display happiness, and tilts its torso forward and moves its arms slightly forward at a slow speed to display sadness. We detail the body movements utilized for the specific coaching styles in Table 6.2. With the **firm** style, the robot leans forward more than the **encouraging** style when displaying sadness. The **encouraging** robot lifts its arm more and leans backward more than the **firm** robot when displaying happiness.

After each repetition, the robot reacts nonverbally (face and body) based on whether a positive feedback case (happy reaction) or a negative one (sad reaction) has occurred. When the robot reacts verbally (positive, negative, or improvement), it will also react nonverbally (happy, sad, happy), but when the robot does not react verbally, it chooses whether or not to react nonverbally. We explore the appropriate cadence of nonverbal feedback in Section 6.3.

6.3 Feedback Cadence

In Section 6.2.1, we introduced 3 categories of verbal feedback (positive, negative, and improvement). We also introduced robot nonverbal feedback. An important question is what the frequency of the verbal and nonverbal feedback should be. As this work focuses on helping older adults maintain their physical well-being, we look especially at the preferences of older adults in this online user study.

Verbal Feedback: We define three different cadence levels for verbal feedback: low, medium, and high. As explained in Section 6.2, we have three categories of verbal feedback:

1. Positive - α_1 good evals or α_2 good speed in a row
2. Negative - β_1 bad evals with the same message or β_2 fast/slow in a row
3. Improvement - β_1 bad evals with the same message followed by a good eval or β_2 fast/slow in a row followed by a good speed

where $\alpha_1, \alpha_2, \beta_1, \beta_2$ are all determined by which cadence level (low, medium, or high) is chosen. We chose potential values (based on discussions with our experts) for each parameter for {low, medium, high} verbal feedback: $\alpha_1 = \{4, 3, 2\}$, $\alpha_2 = \{5, 4, 3\}$, $\beta_1 = \{3, 2, 1\}$, and $\beta_2 = \{4, 3, 2\}$.

Nonverbal Feedback: The robot's nonverbal feedback has two forms: facial expressions and body movements. When the robot does not utter anything verbally after a rep, it has a choice of whether to react nonverbally. We wanted to test different frequencies of reacting at these points in the session where the robot is silent. We chose to test three different cadence levels of the nonverbal feedback. For the high cadence, the robot always reacts positively or negatively; the robot reacts only 50% of the time for the medium cadence; and it reacts only 25% of the time for the low cadence. We chose these numerical values as they were distinguishable when tested on the robot and represent a large variety of response frequencies.

6.3.1 Feedback Cadence User Study

Our user study tested the preferences of older adults for different levels of verbal and nonverbal feedback. We focused our attention on preferences of older adults, but future work could explore the preferences of other groups of interest, as we anticipate feedback cadence preference to vary based on the population surveyed.

We chose to test the cadence levels of nonverbal feedback at or above the cadence levels of verbal feedback to reduce the number of combinations of cadences to test. For example, when the verbal cadence is high and the nonverbal cadence is low, the robot would react verbally very frequently, so changes in the nonverbal cadence (which only would change the feedback after the reps where the robot does not react verbally) would be minimal. However, when the verbal cadence is low, there are many opportunities for the nonverbal cadence to affect the number of nonverbal reactions that the human sees, as there are many reps to which the robot does not react verbally.

Table 6.3: Conditions tested in our feedback cadence study design

		Nonverbal Cadence		
		Low	Medium	High
Verbal Cadence	Low	*	*	*
	Medium		*	*
	High			*

For each of the resulting six conditions (Table 6.3), we recorded one session of bicep curls and one session of lateral raises. Each session consisted of a sequence of 4 good form/good speed, 3 low range form/good speed, 3 good form/slow, and 5 good form/good speed, for a total of 15 reps. This resulted in a total of 12 videos (see Figure 6.5 for a video screenshot), 2 per condition.

Figure 6.5: Screenshot for a video from the study with the human exerciser on the left and the robot providing verbal and nonverbal feedback on the right.

Participants began this online study on Qualtrics with a consent form and demographic information. They read an explanation of the videos they would see in the study that included a description of the two feedback modalities. They also completed a training question that asked "Which of the following will the robot change as part of its feedback?" where the required answer was selecting facial expressions, body movements, and what the robot says. This question was intended to ensure that participants paid attention to all aspects of the robot's behavior.

They then saw one video per condition (randomly chosen between the two videos in each condition), with the order of the videos randomized. After seeing each video, the participants completed two sets of questions: one about their impressions of the robot's verbal feedback and one about its nonverbal feedback. On a 5-pt scale from strongly disagree to strongly agree, they evaluated the usefulness, clarity, timeliness, and helpfulness of each type of feedback. They also had the option to explain their choice in a free text format.

Our hypotheses were as follows:

- **H1**: Changing the verbal feedback cadence affects the participants' view of the verbal feedback.
- **H2**: Changing the nonverbal feedback cadence affects the participants' view of the nonverbal feedback.
- **H3**: Changing the verbal feedback cadence affects the participants' view of the nonverbal feedback.
- **H4**: Changing the nonverbal feedback cadence affects the participants' view of the verbal feedback.

Feedback Cadence Study Results

We recruited 100 online participants using Prolific, with a criterion of age greater than 60, as we want to focus on older adults' perceptions of the robot. We converted the Likert-scale results into numeric values of 1-5 and performed a repeated measures ANOVA with the verbal and nonverbal cadences as the two within-subjects conditions, using the Greenhouse-Geisser corrected p-value. Where we found significance, we then performed a pairwise post hoc test to determine which pairwise differences between feedback cadence levels were significant. Figure 6.6 shows the results marked with significant differences.

Usefulness: We did not find any significant effects in the robot's perceived usefulness.

Clarity: We did find a significant difference in the perceived clarity of the robot's verbal feedback ($p < 0.05, F(2, 198) = 3.26$). Specifically, participants thought the high verbal cadence was clearer than the medium verbal cadence.

We also found a significant difference in the perceived clarity of the robot's nonverbal feedback when changing its verbal feedback level ($p < 0.01, F(2, 198) =$

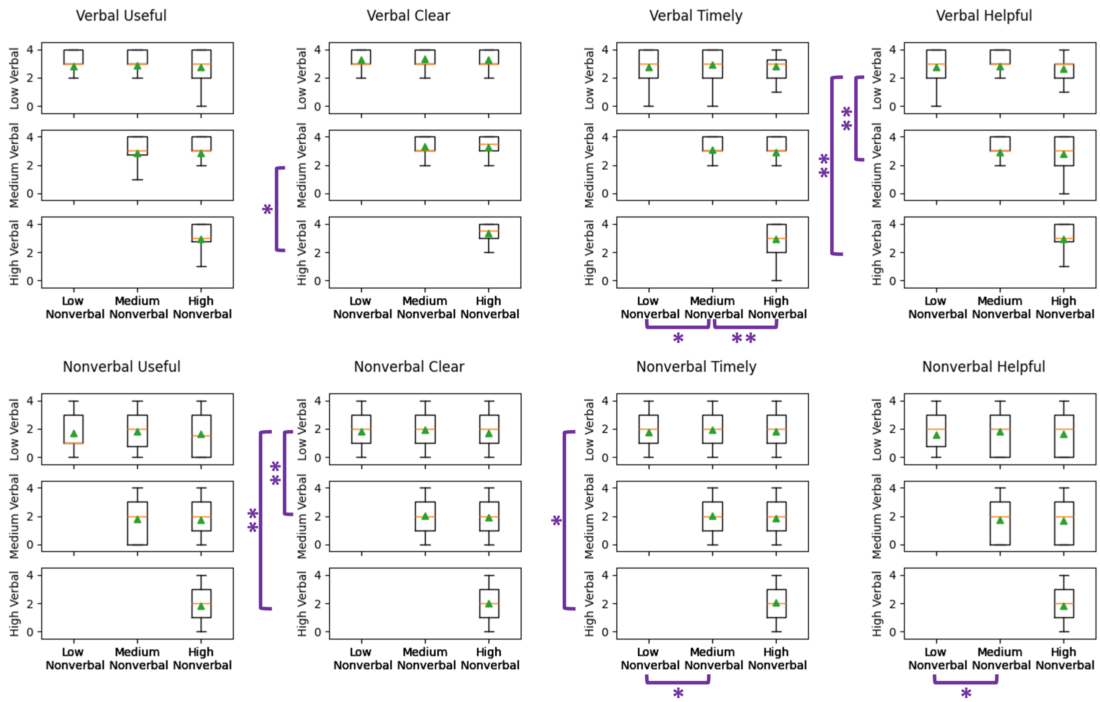

Figure 6.6: Results of all Likert-style measures with perceptions of verbal feedback in the first row and nonverbal feedback in the second row. Significant differences are marked with one asterisk ($p < 0.05$) or two asterisks ($p < 0.01$). Each box plot shows the interquartile range of the data, as well as the min and max of the data. The mean is shown with a green triangle.

4.39). The participants saw the nonverbal feedback as less clear when paired with the low verbal cadence compared to when paired with the medium and high verbal cadence. This is what we call a secondary effect, which we will discuss in more detail when evaluating our hypotheses.

Timeliness: There was a significant difference in the perceived timeliness of the robot's verbal feedback when changing its nonverbal feedback level ($p < 0.05, F(2, 198) = 3.75$), which is also a secondary effect. Participants saw the verbal feedback as more timely with medium nonverbal cadence compared to the low and high levels.

We also found a significant difference in the perceived timeliness of the nonverbal feedback ($p < 0.05, F(2, 198) = 3.70$). The participants saw the low level of nonverbal feedback as less timely than the medium level.

We found a secondary effect with the nonverbal feedback's timeliness; there was a significant difference in the perceived timeliness of the robot's nonverbal feedback when changing its verbal feedback ($p < 0.05, F(2, 198) = 4.36$). Participants saw the nonverbal feedback as more timely with the high verbal level compared to the low verbal level.

Helpfulness: We found a significant difference in the perceived helpfulness of the robot's verbal feedback ($p < 0.01, F(2, 198) = 6.47$). Participants saw the *low* level of verbal feedback as less helpful than the medium and high levels.

We also found a significant difference in the perceived helpfulness of the robot's nonverbal feedback ($p < 0.05, F(2, 198) = 3.42$). Participants saw the low nonverbal level as less helpful than the medium level.

6.3.2 Feedback Cadence Study Discussion

H1: Changing the verbal feedback cadence affects the participants' view of the verbal feedback: We have support for this hypothesis. Changing the verbal feedback cadence has an effect on the perceived clarity and helpfulness of the robot's verbal feedback. Specifically, participants seemed to prefer the high level compared to the medium in terms of clarity and the medium and high over the low in terms of helpfulness. This result illustrates that the low level of verbal feedback is generally not preferred and that the medium and high levels are perceived as more helpful and clear. The more frequent feedback could make the robot feel more responsive and aware of what the human is doing.

H2: Changing the nonverbal feedback cadence affects the participants' view of the nonverbal feedback: We have support for this hypothesis. Changing the nonverbal feedback cadence has an effect on the perceived timeliness and helpfulness of the robot's feedback. Participants preferred a medium level for both of these measures over the low level. The higher frequency of nonverbal reactions could be more engaging for participants, and they could feel that the robot is actually responding to the exercises.

H3: Changing the verbal feedback cadence affects the participants' view of the nonverbal feedback: We have support for this hypothesis. Changing the robot's verbal feedback cadence affects the perception of the nonverbal feedback's clarity and timeliness. Participants thought the medium and high levels of verbal feedback paired with better clarity and timeliness of the nonverbal feedback. This is an interesting secondary effect that shows that the two feedback modalities are inexorably linked. Changing the cadence of one will affect the perception of the other.

Psychology researchers have explored this interaction, and [56] explores many possibilities of how they are linked, including aggressive verbal utterances that signal aggressive nonverbal reactions. In our results, changing the verbal frequency could help participants interpret the nonverbal feedback better and make it appear clearer and more timely.

H4: Changing the nonverbal feedback cadence affects the participants' view of the verbal feedback: We have support for this hypothesis. Changing the robot's nonverbal feedback cadence affects the perception of the verbal feedback's timeliness. Participants preferred the medium level of nonverbal feedback when thinking about the timeliness of verbal feedback. This also agrees with our finding in **H2** that the medium nonverbal cadence is preferred to the low level.

Cadence Recommendations: As a result of this study, we saw that the medium/high verbal cadences were viewed as more helpful, and the medium nonverbal cadence was viewed as more helpful and timely. We chose the medium levels for both verbal and nonverbal feedback for the robot in the feedback style study in Section 6.4. A comment from a participant indicated that "the speed at which the [verbal] correction came when the exercise was not done properly" might be "too fast", and another indicated that the robot's verbal feedback at high frequency was "too repetitive." We used these qualitative comments to choose the medium verbal cadence over the high level.

6.4 Feedback Style Study Design

We ran our main user study to test the robot exercise coach's use of different feedback styles. Although robot exercise coaches have been developed before [27, 31, 55], here we are interested in comparing people's preferences and performance with different coaching styles, similar to how [59] compared responses to different levels of extroversion.

We have three different robot coaches, based on the three styles presented in Section 6.2. The **neutral** style communicates only instructions, such as the start and end of each set, with no verbal or nonverbal feedback while the human is exercising. It has a neutral, unchanging facial expression and slight neutral movements throughout the session. The neutral expression is 10% happy, as our domain experts thought that the robot's facial expression should always be slightly positive.

The second coach has a **firm** style with minimal encouragement. And the third robot style has an **encouraging** style with encouragement included in all its feedback. For these two styles, the robot communicates the verbal instructions (start/end of the set, etc.) as well as corrections. Examples of the verbal and nonverbal feedback for these three coaches can be found in Table 6.2.

Our hypotheses are as follows:

- **H1:** Participants have a better performance with the **firm** and **encouraging** styles compared to the **neutral** robot.
- **H2:** Participants have a better perception of the robot with the **firm** and **encouraging** styles compared to the **neutral** robot.
- **H3:** Participants have a different performance between the **firm** and **encouraging** styles.
- **H4:** Participants have a different perception of the robot between the **firm** and **encouraging** styles.

6.4.1 Study Procedures

Participants began this in-person study with a consent form and brief demographics that include gender, age, ethnicity, familiarity with robots and programming, and level of education. A similar consent form and demographic survey are included in Appendix A.1. The participants were then shown videos of the two exercises: bicep curls and lateral raises, as well as sets of dumbbells (3 lb, 6 lb, and 10 lb) that they had the option of using during the exercises. They were instructed that they could choose whichever set they are comfortable with (if any) and that they could change dumbbells between sets, if desired.

After these explanations, the participants began one of the three rounds of exercise sessions (see the setup in Figure 6.7). In each round, the participant performed four sets of exercises: two bicep curls followed by two lateral raises. At the beginning of each set, the robot said 'Get ready for set {*set number*} out of 2 of {*exercise name*}.' This was accompanied by a lowering of its raised right arm to halfway. After a brief pause, the robot lowered its arm all the way down and said 'Start {*exercise name*}.' Arm movements to accompany the robot's speech were chosen as an additional signal to the verbal instructions for whether the participant should be exercising or resting.

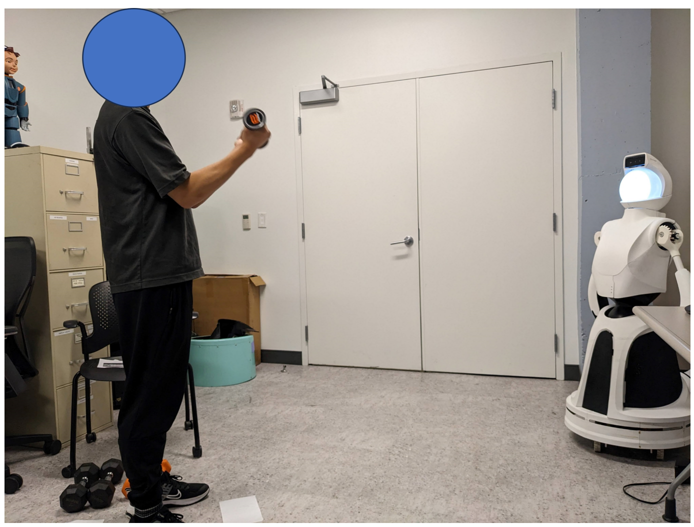

Figure 6.7: Study setup with participant doing bicep curls

During the set, the robot provided feedback based on the style it was using (**neutral**, **firm**, or **encouraging**). When the set was almost done, the robot said 'Almost done,' followed by 'Rest' when the set was done (raising its arm to a high position). Each set was completed if a minimum time of 30 seconds and a maximum time of 50 seconds or more than 8 reps were completed (whichever occurred first).

The participant rested for 40 seconds, with the robot prompting 'Rest for 20 more seconds' at the halfway point. The robot then began the next set, for a total of 4 sets per round.

Each participant saw the **neutral** robot for the first round and then was randomly assigned the **firm** or **encouraging** robot for the second round, with the other option chosen for the third round. We did not fully randomize the study because we wanted the **neutral** style to be the baseline to which participants compared the other two styles. In between each round, the participant completed a short survey. The survey had questions taken from the Godspeed Questionnaire [8] to measure their perception of the robot in terms of animacy, likability, and perceived intelligence. It also had three additional Likert-style questions related to the robot's feedback style: how strict, motivational, and corrective is the robot coach? Appendix A.1 includes the survey presented to participants in this study.

After the participants completed three rounds of exercise followed by surveys, they completed a final survey comparing the styles they saw in rounds 2 and 3. They were asked to decide which round was more lively, interactive, responsive, friendly, kind, pleasant, competent, intelligent, strict, motivational, and corrective, on a 5-point Likert scale, with 3 indicating no difference between the styles.

6.5 Feedback Style Results

Our study protocol was approved by the CMU IRB, and we used the CMU Center for Behavioral and Decision Research to recruit participants from both CMU and non-CMU sources, for a total of 19 participants. The participants were mainly in the 20-49 age range ($\mu = 29, \sigma = 15$). Although a focus of this work is understanding how to provide feedback to older adults during exercise, we recruited a general population for this study due to the difficulty in recruiting older adults. We wanted to determine whether the feedback styles were appropriate and resulted in varying experiences and performances for a general population first before performing further studies with the older adult population, which we did in Chapter 7.

6.5.1 Performance

We first compared the average number of reps, the percentage of good form reps, and the percentage of good speed reps that the participants performed per robot style (averaged over the four sets they performed with each style). We performed an ANOVA for each of these measures between the three different styles and then a Tukey post hoc test.

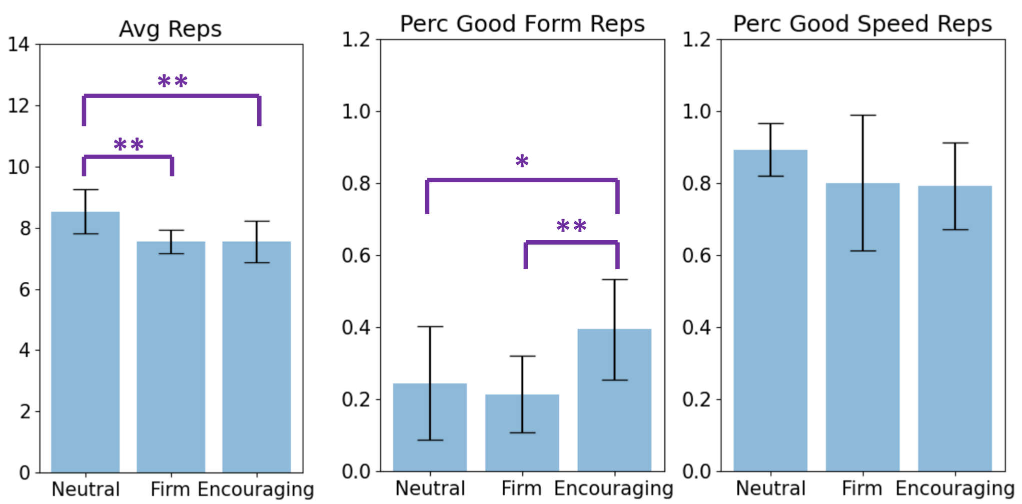

Figure 6.8: Performance measures compared between the three different styles. Statistical differences are shown with one asterisk ($p < 0.05$) and two asterisks ($p < 0.01$).

Figure 6.8 compares the performance measures between the three styles, with values averaged per round. We see that participants performed around 8 reps per set for each round, which is mostly due to the robot stopping each set when either a maximum time was achieved or 8 reps were performed. However, participants performed fewer reps with the **firm** and **encouraging** styles compared to the **neutral** style ($F(2, 42) = 12.2, p < 0.01$).

Participants performed a higher percentage of good-form reps with the **encouraging** style compared to the **neutral** ($F(2, 42) = 5.1, p < 0.01$) and **firm** ($F(2, 42) = 6.9, p < 0.01$) styles. We do not have significant results related to the percentage of good speed reps performed with each style.

6.5.2 Perception of the Robot

We measured subjective preferences in two ways: via the surveys that participants completed after each round and the final survey comparing the robot styles.

For individual surveys, we again used an ANOVA paired with a Tukey post hoc to determine statistical significance. The results indicate that participants saw the **neutral** style as less lively ($F(2, 69) = 28.5, p < 0.01$), less interactive ($F(2, 69) = 22.6, p < 0.01$), less responsive ($F(2, 69) = 32.6, p < 0.01$), less intelligent ($F(2, 69) = 9.3, p < 0.01$), less motivational ($F(2, 69) = 12.0, p < 0.01$), less strict ($F(2, 69) = 9.3, p < 0.01$), and less corrective ($F(2, 69) = 52.1, p < 0.01$) when compared to the **firm** and **encouraging** styles. Participants also perceived the **neutral** style as less friendly than the **encouraging** one ($F(2, 69) = 3.5, p < 0.01$), but that same difference is not perceived with the **firm** style.

For the final survey, participants rated the styles on a scale of 1-5, where the middle value (3) indicated an equal preference for both styles, a value greater than 3 a preference for the **encouraging** style, and a value less than 3 a preference for the **firm** style. We performed a one-sample t-test to determine if the values for each measure are significantly different from the middle value of 3. As shown in Figure 6.9, we can see that participants viewed the **firm** style as more responsive ($t = -2.14, p < 0.05$), strict ($t = -2.36, p < 0.05$), and corrective ($t = -2.12, p < 0.05$), compared to the **encouraging** style.

We also noticed that some participants specifically noticed the robot's nonverbal behavior, with verbal comments like 'Oh, the robot smiled at me' (after a positive comment and smiling facial expression) and 'It seems angry at me' (after a correction and frowning facial expression). Others did not seem to notice changes in the robot's nonverbal behavior (especially between the **firm** and **encouraging** styles), noting the change in the verbal phrases and not mentioning any changes in facial expressions or body movements. This indicates diversity in how people view robot nonverbal behavior.

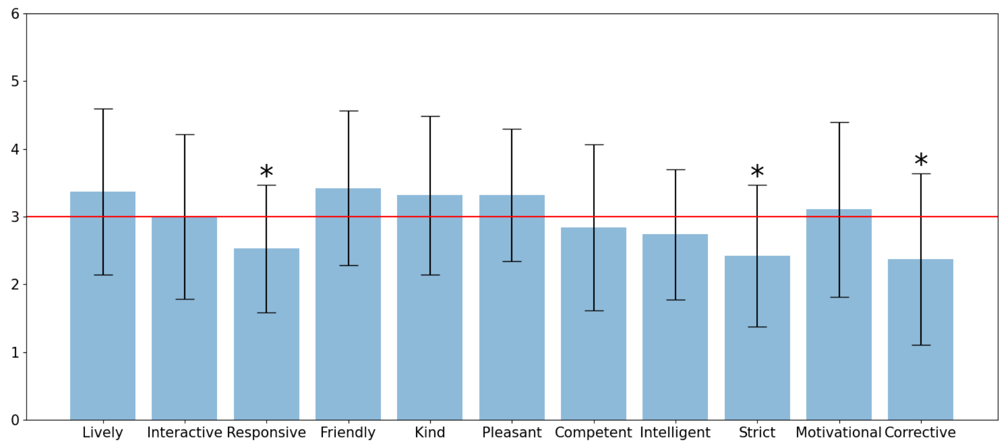

Figure 6.9: Subjective preferences based on Likert-style questions, scored 1-5. Values above 3 are a preference for **encouraging**), and values below 3 are a preference for **firm**. Statistical differences from the center value are shown with one asterisk ($p < 0.05$).

6.6 Feedback Style Discussion

H1: Participants perform better with the firm and encouraging styles compared to the neutral style. We have partial support for this hypothesis. We see that the participants performed a higher percentage of good-form reps with the **encouraging** style compared to **neutral**. However, we do see that participants performed more reps with the **neutral** style compared to the **firm** and **encouraging** styles, which does not support this hypothesis. From our observations, participants spent time listening to feedback and watching the nonverbal reactions of the **firm** and **encouraging** styles, resulting in fewer reps completed compared to **neutral**, but the reps they completed had better form.

H2: Participants have a better perception of the robot with the firm and encouraging styles compared to the neutral style. We have support for this hypothesis. We find that the participants perceived the **neutral** style as less lively, interactive, responsive, intelligent, motivational, strict, and corrective compared to the other two styles. Participants had a less positive perception of the **neutral** style, which is intuitive since the **neutral** style does not react to what the human is doing.

H3: Participants have a different performance between the firm and encouraging styles. We have support for this hypothesis. Participants performed a higher percentage of good-form reps with the **encouraging** compared to the **firm**. Although both styles provided the same content of corrections, the increased encouragement and more positive nonverbal reactions had a significant impact on performance for these participants.

H4: Participants have a different perception of the robot between the firm and encouraging styles. We have partial support for this hypothesis. We have significant results when using the final survey responses; participants viewed the **firm** style as more responsive, strict, and corrective. The **firm** style is designed to be more strict and corrective, as it does not soften the blow of corrections with encouragement, so this result is not surprising. However, it is surprising that the **firm** style was viewed as more responsive than the **encouraging** style, since both styles have the same conditions for feedback and the content of feedback; the only difference is the way feedback is conveyed. It is possible that the difference is perceived because the stronger criticism of the **firm** style stands out in the minds of the participants, so it is perceived as more responsive. Another possible explanation is that participants performed more good form reps with the **encouraging** style, so they received fewer comments (less corrections), which led the robot to appear less responsive.

Summary: In this work, we presented the formulation of a robot exercise coach that provides contextually aware feedback to the human in real time in two styles, **firm** and **encouraging**. For these two styles, we showed that people have different performances and subjective experiences, which sets the stage for the question: “How should the robot choose which style to use when?”, which we explore in the next chapter.

7 Personalized Feedback Style

Summary: Using the exercise coach presented in Chapter 6, we explore a contextual bandit approach that enables the robot to learn the best style to use over time to optimize human performance. We first present the contextual bandit and test it using a human model in complex contextual situations. Next, we collect a dataset of older adults to add to the data collected in Section 6.5 to test the contextual bandit on real-world data. After showing that the bandit shows promise in choosing the performance-optimizing feedback style on human data, we design an in person study to test this approach with participants, comparing the **adaptive** approach to choosing only one style (**firm** or **encouraging**). We show that the **adaptive** approach is successful in determining the performance-optimizing style to choose in real-time and performs well for humans who perform better with either one style or equally well with both.

The contents of this chapter will be published in [44].

7.1 Adaptive Feedback Model

We introduce an adaptive feedback model using a contextual bandit to allow a robot to learn in real time which feedback style to use when. The feedback style someone might perform best with may be dependent on a variety of factors, and we include fatigue estimation in this work as an important feature, since someone's performance with encouragement may vary based on how tired they are.

In this approach (Figure 7.1), the robot first observes the context from the human.

For our work, we assume that this context c is a fatigue estimate, where the robot estimates whether the human has low, moderate, or high fatigue. Next, the robot queries an Upper Confidence Bound policy Π for the best action (feedback style), given c. The chosen action a is then used to generate multi-model feedback in the chosen style. The human observes this feedback and performs the next rep of the exercise with either good or bad form. The robot can observe this reward (0 or 1) and trains on the combination of context, action, and reward to improve the policy (and minimize regret over time). The major difference between a contextual bandit and a simple bandit is that a simple bandit does not observe the context and trains a policy on only actions and rewards. The goal of this approach is for the robot to choose the action that maximizes human performance given the context.

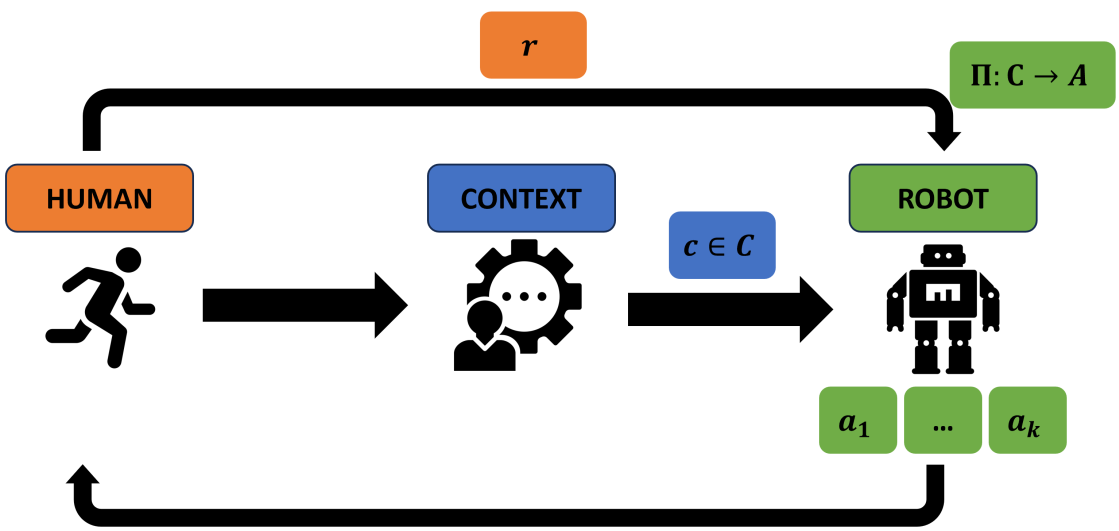

Figure 7.1: The robot observes the context c from the human and chooses an action a. The human responds to the robot's action, and their performance r forms the reward for the robot to train policy Π.

When translating this approach to the robot exercise coach, we note that the robot does not provide verbal feedback at every repetition (only reacting when a pattern is observed, such as 3 good reps in a row). Reacting verbally every rep would be very overwhelming for the human, and they would not be able to process the feedback at such a high frequency. In practice, we assume that any reps following a verbal utterance with a particular style have the same style (e.g., an encouraging phrase has an effect for the few reps after that phrase is uttered), and the style of the robot can change when the conditions have been met for the next verbal utterance. This will allow the robot to adapt to the human's performances and will allow the human to react to the robot's feedback. However, in the simulation formulation, we assume that the robot can change its feedback style every repetition.

We developed a human simulation and fatigue model to determine whether the contextual bandit can learn the feedback style with which the simulated human performs best. First, we created a model of fatigue as the human is exercising; this is the context the robot will observe. We simulate an exercise session where the human performs sets of 10 reps, where fatigue is low for the first 5 reps, moderate for the next 3 reps, and high for the last 2 reps. This approximates a human whose fatigue increases as they perform each set and resets to low fatigue after a rest period following each set.

Next, we determined the feedback styles, or actions, the robot can choose between. We chose five different feedback styles: **very firm**, **firm**, **neutral**, **encouraging**, and **very encouraging**. We have thus far implemented two of these styles (**firm** and **encouraging**) on the actual robot, as discussed in Chapter 6, but we wanted to explore a larger range of styles in simulation to see if the bandit approach can still learn the best style in a more complicated scenario.

The bandit learns a policy on each of these feedback styles as a function of the context observed, and treats each of the styles as unrelated. However, the styles are not truly independent, as **very firm** is a more extreme version of **firm** and **very encouraging** a more extreme version of **encouraging**. We tried to update the policy of related actions after viewing the reward, similar to the pseudo-reward approach presented in [18]. For example, a reward after **very firm** feedback may tell us something about the **firm** action, but found that the pseudo-rewards did not significantly improve our results. Therefore, we treat each of the styles as independent.

Lastly, we chose to construct a reward that is simply based on performance: 0 for bad form and 1 for good form. Our simulated human model performs the next rep correctly based on a probability $p(a_k, f)$, where a_k is the feedback style chosen by the robot after each rep and f is the fatigue level. We assume that the simulated human has a base probability of performing the next repetition correctly without fatigue after seeing feedback from each style. We also assume that fatigue can impact their performance in two ways. The first is a general reduction in performance as fatigue increases. The second is a more complex interaction between style and fatigue; for some people, they may perform better with one style as their fatigue increases (e.g., encouragement having a larger impact when tired). Based on these intuitions, we construct the following equation to compute the simulated human's probability of completing a repetition correctly ($p(a_k, f)$) after seeing feedback from a particular style k with a specific level of fatigue f:

$$p(a_k, f) = p(a_k)(1 - \gamma f e(a_k)) \tag{7.1}$$

- a_k is the action taken by the bandit, with a_0 being **very firm**, a_1 being **firm**, etc.
- $p(a_k, f)$ is the probability the human performs the next rep correctly after viewing feedback style a_k with fatigue f (bounded between 0.05 and 0.95)
- $p(a_k)$ is the probability the human performs the next rep correctly with no fatigue after seeing feedback style a_k
- γ is 0 if there is no fatigue dependence in performance, and 1 if there is a large fatigue dependence
- f is 0, 0.3, or 0.6 for low, medium, and high fatigue, respectively
- $e(a_k)$ is an action-dependent fatigue factor (e.g., whether high fatigue reduces or improves performance with a specific feedback style)

We can capture three different fatigue dependence scenarios using this equation, and to understand them, let us consider someone who performs well without fatigue with the **very firm** style ($p(a_0) = 0.8$), which means that they perform 80% of their reps with good form after seeing **very firm** feedback.

The first case *no fatigue dependence* covers the situation where the human performance does not change with fatigue ($\gamma = 0$). For this person, that would mean that $p(a_0, f) = p(a_0) = 0.8$. They always perform 80% of their reps with good form after seeing **very firm** feedback, regardless of their fatigue.

The second level of complexity (*basic fatigue dependence*) covers the case where human performance decreases with fatigue, but there are no action-dependent fatigue factors ($e(a_0) = 1$). For example, if $\gamma = 1$, then the final term in the equation has a large impact on the human's performance with **very firm**. If their fatigue was high for a particular rep ($f = 0.6$), their performance $p(a_0, f)$ with high fatigue would become $0.8(1 - 0.4) = 0.48$, which is 48% of their reps performed with good form after seeing **very firm** feedback with high fatigue, compared to 80% with low fatigue.

The most complex situation (*complex fatigue dependence*) covers the case where the human's performance has an action-dependent fatigue factor. If they had an action-dependent fatigue factor of ($e(a_0) = 0.5$), which means their preference for **very firm** feedback increases with higher fatigue, their performance with high fatigue would be $p(a_0, f) = 0.8(1-(0.4 \times 0.5)) = 0.64$. This means that since their preference for **very firm** feedback increases with fatigue, their drop in performance with high fatigue is less severe (64% instead of 48% without the action-dependent fatigue factor). When $e(a_k) > 1$, fatigue reduces performance more for style k, and when $e(a_k) < 1$, fatigue reduces performance less.

7.1.1 Contextual Bandit Simulation Results

We ran a contextual bandit simulation in these three different scenarios. Experiment 1 explores no fatigue dependence, where the human's performance does not change with fatigue ($\gamma = 0, p(a_k, f) = p(a_k)$). Experiment 2 explores basic fatigue dependence, where the human's performance reduces with fatigue, but the style they perform best with does not change with fatigue ($\gamma = 0.5, e = [1, 1, 1, 1, 1]$). The vector e represents the action-dependent fatigue factor for each of the five styles: **very firm**, **firm**, **neutral**, **encouraging**, and **very encouraging**. Experiment 3 is the most complicated scenario of complex fatigue dependence, where the style the human performs best with changes based on their fatigue.($\gamma = 0.5, e = [4, 2, 1, -0.25, -2]$). In this scenario, performance with **very encouraging** and **encouraging** increases with fatigue, and performance with **very firm** and **firm** decreases with fatigue. These experiments will demonstrate the efficacy of this approach in learning the optimal feedback style in increasingly complicated scenarios. We implement the contextual bandit model using the *bayesianbandits*[1] package.

For all these experiments, let us consider someone who performs best with **very firm** feedback and performs worst with **very encouraging** feedback. Specifically,

[1]https://bayesianbandits.readthedocs.io/en/latest/index.html

$p(a_k) = [0.8, 0.6, 0.5, 0.4, 0.3]$, where the human performs 80% of their reps correctly without taking fatigue into account after viewing **very firm** feedback, 60% with **firm**, etc. For each experiment, we ran 20 sets of 10 reps. Figure 7.2 shows the average reward per set for 30 runs of each experiment, as well as the optimal expected reward the bandit could achieve. The recorded reward can exceed the optimal reward in some cases, as the reward received is 0 or 1 and the expected reward is a probability (e.g., 0.8).

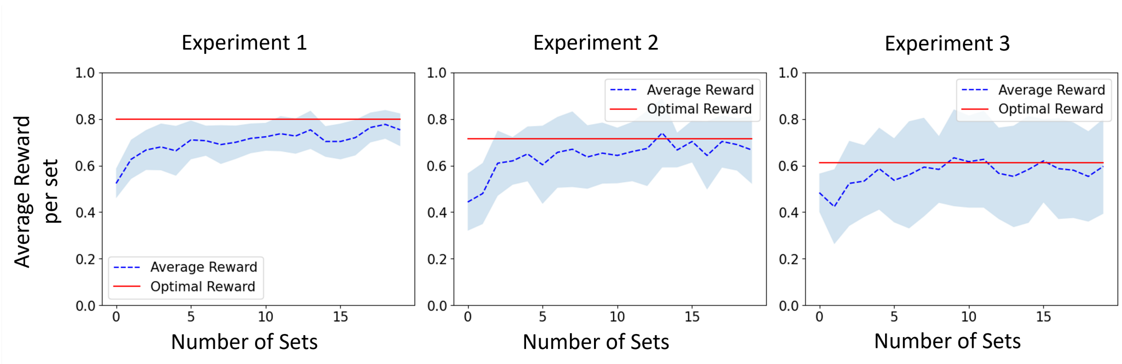

Figure 7.2: The average reward received shown in blue for each experiment (over 30 trials) with one standard deviation shaded in light blue. The average reward per set increases as the bandit sees more data, with the 20th set closer to the expected optimal reward (shown in red) compared to the 1st set.

Experiment 1 (No Fatigue Dependence)

We first performed an experiment with 20 sets of 10 reps each where human performance is not fatigue-dependent ($\gamma = 0$, $p(a_k, f) = p(a_k)$). The bandit's optimal action for all reps is the **very firm** action. Figure 7.3 shows that the bandit chose the **very firm** action 77% of the time over the 20 sets, and learned by the later sets to choose **very firm** for all the reps. As seen in Figure 7.2, the bandit quickly approaches the expected optimal reward of 80%. Since the reward the bandit receives is sampled, even when choosing the optimal action it will only receive a reward 80% of the time. Therefore, if by chance, it does not receive a reward with the optimal action for a few repetitions, it may begin to experiment with suboptimal actions before converging on the optimal action again. This can be seen in Sets 9 and 10.

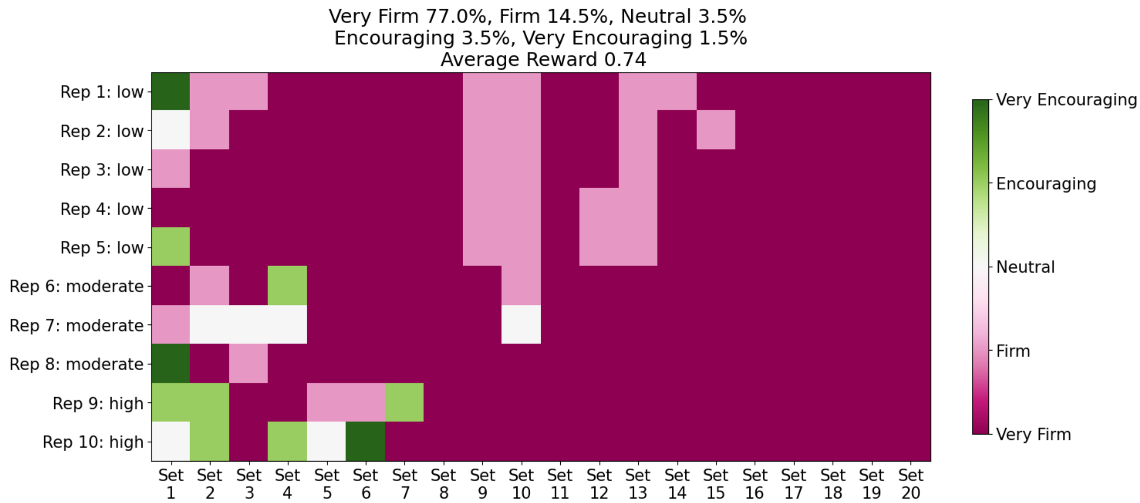

Figure 7.3: Experiment 1: Simulation results with no human fatigue dependence in performance. The human's performance is the same across all fatigue levels, and the optimal action for the bandit is **very firm** across all reps.

Experiment 2 (Basic Fatigue Dependence)

We next ran an experiment over 20 sets of 10 reps each where the human performance reduces with fatigue, but the style they perform best with does not change with fatigue ($\gamma = 0.5, e = [1, 1, 1, 1, 1]$). Figure 7.4 shows that the bandit chose the optimal action of **very firm** 83% of the time over the 20 sets, and learned by the later sets to choose **very firm** for all the reps (with some exploration in Sets 10-11 and in Set 19). As seen in Figure 7.2, the optimal reward is lower than in the first experiment because human performance decreases with fatigue, but the bandit still approaches the expected optimal reward around set 10.

Experiment 3 (Complex Fatigue Dependence)

We lastly ran an experiment over 20 sets of 10 reps each where the style the human performs best with changes based on their fatigue ($\gamma = 0.5, e = [4, 2, 1, -0.25, -2]$). This vector e indicates that performance with the **very encouraging** and **encouraging** styles increases with fatigue, and performance with **very firm** and **firm** decreases with fatigue. The bandit should optimally choose **very firm** for low fatigue, is relatively indifferent between all styles for moderate fatigue (probabilities range from

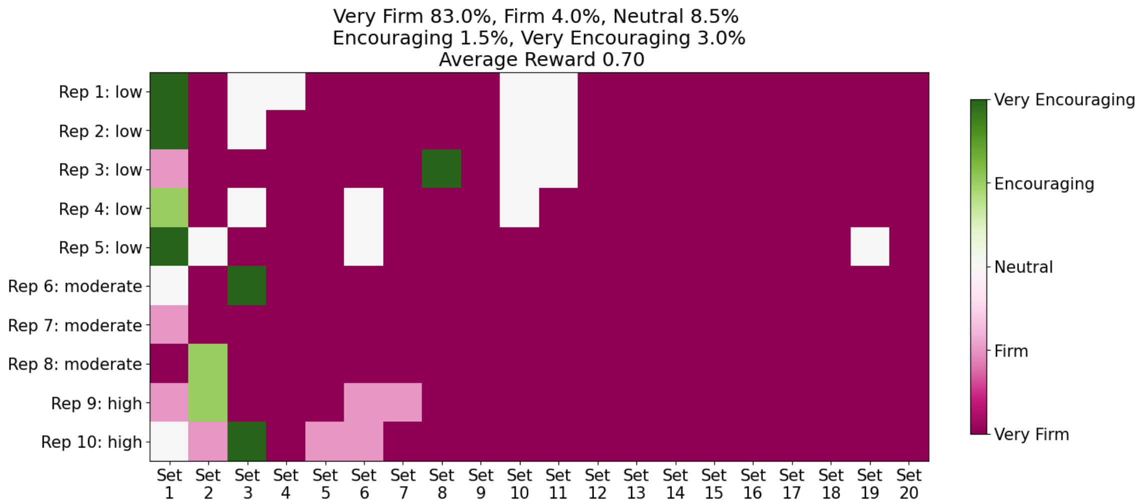

Figure 7.4: Experiment 2: Simulation results with basic human fatigue dependence in performance. The human's performance decreases as fatigue increases, but the optimal action for the bandit is **very firm** across all reps.

39% - 43%), and should choose **encouraging** or **very encouraging** for high fatigue. This extreme case of fatigue dependence illustrates how the style the robot should choose could completely change as fatigue changes.

Figure 7.5 shows that the bandit can still learn the optimal action in the complex fatigue scenario. It learns to choose **very firm** for low fatigue and **very encouraging** for high fatigue. As seen in Figure 7.2, the average bandit reward approaches the expected optimal reward as the number of sets increases, the expected optimal reward being much lower than in the previous experiments.

Contextual vs. Simple Bandit

In these experiments, we show that the contextual bandit can accurately learn the optimal feedback style even with complex fatigue dependence, but we also want to compare how a simple bandit without access to the context would perform in these three experiments.

Figure 7.6 illustrates the distributions of the average reward for the three fatigue situations with and without context after 30 simulation runs. We performed a Mann-Whitney U test with a one-sided alternative to determine significant differences.

For no fatigue dependence, the bandit performed slightly better without context

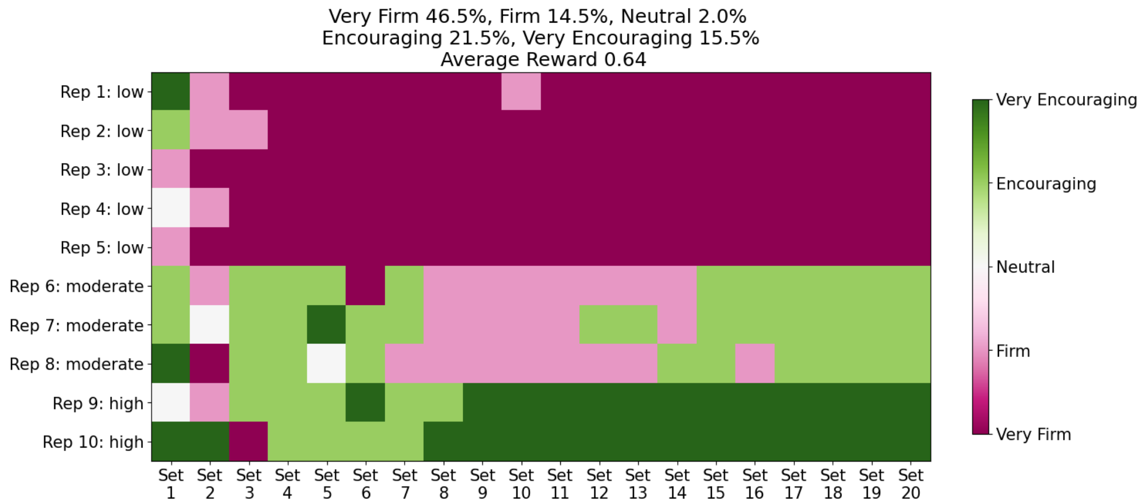

Figure 7.5: Experiment 3: Simulation results with complex fatigue dependence in performance. The bandit should choose **very firm** for low fatigue, is relatively indifferent between all styles for moderate fatigue, and should choose **encouraging** or **very encouraging** for high fatigue.

(average reward of 0.75) compared to with context (average reward of 0.70). Since the context does not impact performance for this case, the contextual bandit is learning identical models for each fatigue level and is therefore learning slightly slower than the simple bandit. However, even with this difference, the contextual bandit performs quite similarly to the simple bandit for this fatigue case.

In the case of basic fatigue dependence, the simple bandit performed slightly better again (average reward of 0.66) compared to the contextual bandit (average reward of 0.62). Again, context does not affect the choice of feedback style in this contextual situation, so the contextual bandit learns slightly slower than the simple bandit (but not statistically significantly).

Complex fatigue dependence is where we anticipate seeing the most difference in performance with context, as the context actually changes the optimal choice of the bandit. We can see that the contextual bandit performs significantly better (average reward of 0.56) than the simple bandit (average reward of 0.46) for complex fatigue ($p < 0.001$). For people who fit this contextual situation, having the context is vital for performance, as the simple bandit performed significantly worse without the context to inform its learning.

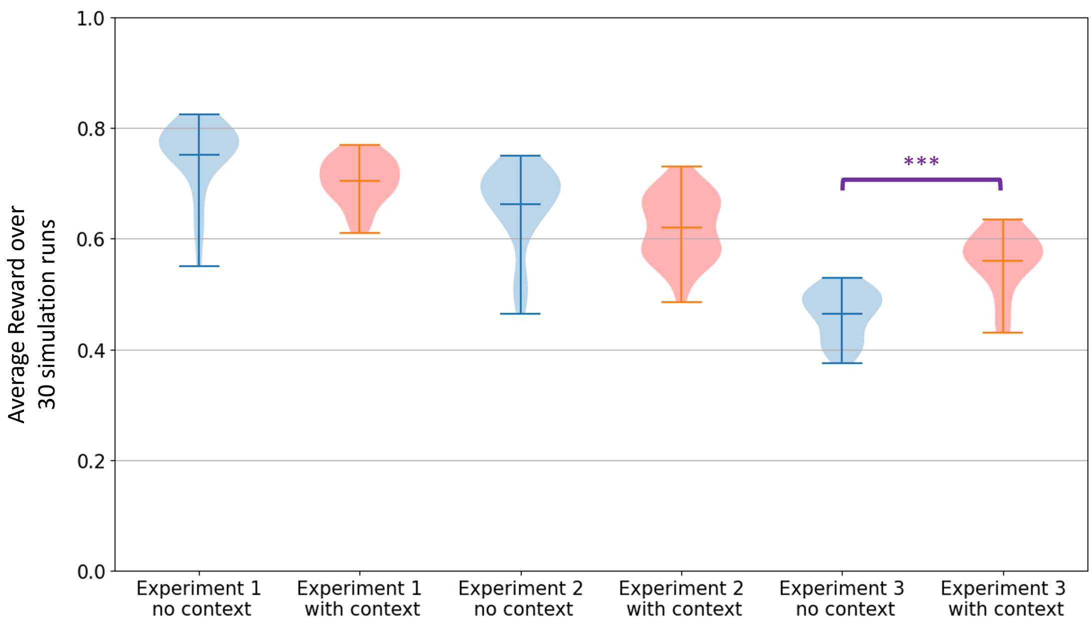

Figure 7.6: Comparison of average reward for all three fatigue situations with a simple bandit and a contextual bandit. Significant differences are indicated with $***$ ($p < 0.001$), and the mean of each distribution is also marked.

7.2 Analysis of Dataset

In the study conducted in Section 6.5, 19 participants performed three rounds of exercise of 4 sets each (2 bicep curls and 2 lateral raises). They saw a baseline robot in the first round that did not provide any feedback, and then a **firm** and **encouraging** robot in the second and third rounds (with the order of these two rounds randomized). After each round, the participants completed a short survey to measure their perception of the robot in terms of animacy, likability, and perceived intelligence (see Section A.1 for the survey questions). Most of these 19 participants were young adults ($\mu = 28.5, \sigma = 14.7$, with 1 participant 60 years or older.

We implemented similar study procedures with our older adult study: two counter-balanced rounds of exercise, one with **firm** and one with **encouraging**, and surveys with the same questions. In this study, we took Quori to an assisted living facility (Vincentian Schenley Gardens in Pittsburgh, PA) and ran the study protocol with 8 older adults there. We also added two more participants in a laboratory setting, resulting in a total of 10 participants ($\mu = 77.4, \sigma = 10.6$) with 9 participants 60 years or older (one participant was 59).

Combining the two sets of data, we have a total of 29 participants, 11 of whom are 60 years or older. We aimed to create a more age-diverse dataset as most traditional recruiting methods skew younger in age, and exercise has value across all age ranges. Even with a specific effort to add older adults, we still have an unbalanced dataset, but it is much more balanced than the original data set (38% older adults rather than 5%).

There are two major differences between these two studies. The first is that in the older adult study, we asked participants before beginning the exercise rounds whether they preferred a **firm** or **encouraging** feedback style. As we will show below, the style that participants state before the interaction and the style that they prefer (rate higher on the surveys) do not always match the style they perform best with. This highlights the need for an **adaptive** approach to learn in an online way which style each person performs best with.

The second difference is that in the older adult study, we estimated fatigue using a heart rate monitor (Polar Verity Sense), calculating the heart rate reserve using the following equation[2]:

$$HRR = \frac{HR - RHR}{MaxHR - RHR} \tag{7.2}$$

- HRR is the heart rate reserve, where we set less than 0.2 to be low fatigue, 0.2-0.4 to be moderate fatigue, and 0.4 and above to be high fatigue after pilot testing
- HR is the participant's current heart rate in beats per minute
- RHR is the participant's resting heart rate, computed as an average during the introduction to the exercise session
- $MaxHR$ is the participant's estimated maximum heart rate, computed using $(220 - Age)$[3]

which allows us to determine how high an individual's heart rate is, proportional to their resting and max heart rates. [19] found that heart rate reserve provided an accurate prediction of exercise intensity, and reported RPE (Rate of Perceived

[2]https://my.clevelandclinic.org/health/articles/24649-heart-rate-reserve
[3]https://www.heart.org/en/healthy-living/fitness/fitness-basics/target-heart-rates

Exertion) levels were slightly less accurate than recorded HRR.

For this study, we also included an adjustment to the verbal phrases the robot uttered based on fatigue (see Section A.2 for the LLM prompt used to generate the verbal phrases). For example, if the robot was using the **firm** style and was giving a correction to improve the range of motion on the left side, it might say 'Focus on a full range of motion on the left side, keep pushing!' if the fatigue was low and 'Focus on a full range of motion on the left side.' if fatigue was moderate. This allows for the fatigue information to be incorporated into the robot's verbal phrases.

We do not have fatigue information for the 19 participants in the first study, so after observing the data from the older adult study, we estimated that participants have low fatigue for the first 90% of each set and moderate fatigue for the remainder of each set.

7.2.1 Performance and Preference Groups

We first assigned a performance group to each of the 29 participants: those who performed better with the **encouraging** style, those who performed better with the **firm** style, and those who performed approximately the same with the two styles. To do this, we computed the difference d in performance (percentage of good-form reps) between the two styles for each participant (**encouraging** - **firm**). We then calculated the mean and standard deviation of those differences. To compute a 95% confidence interval, we used the formula $\mu \pm (t \times \sigma)$, where t is the critical value from the t-distribution and σ is the standard error (standard deviation divided by $\sqrt{n}$). This resulted in an interval, where we could take the performance difference of each participant and determine whether it was within the interval (no performance difference), less than the minimum of the interval (perform better with **firm**), or greater than the maximum of the interval (perform better with **encouraging**).

We also calculated a subjective measure for each participant from their survey responses after experiencing each feedback style. We averaged the 1-7 Likert scores that each participant completed for each robot style (lively, interactive, responsive, friendly, kind, pleasant, competent, intelligent). We subtracted the **firm** score from the **encouraging** one to obtain a single value where < 0 indicates a preference for **firm** and > 0 a preference for **encouraging**. We then performed the same grouping procedure as for the performance scores to determine which participants preferred the **firm** style, preferred the **encouraging** style, or did not have a style preference. The grouping method allowed us to determine which values were "close enough" to 0 to indicate no style preference.

Table 7.1 includes the number of participants with each combination of the performance and preference groups. We can see that 76% of the participants lie off the diagonal, where the diagonal indicates agreement between the style someone prefers and the style they perform best with. Not taking into account the groups with no preference or performance difference, 70% of the remaining participants had complete disagreement in group assignment (e.g., preferring the **firm** style but performing best with the **encouraging** style).

Table 7.1: Performance and preference groups for the 29 participants. The diagonal indicates agreement between the style preference and the style the participant performs best with.

	Prefers Firm	**Prefers Encouraging**	**No Style Preference**
Performs Best with Firm	1	2	6
Performs Best with Encouraging	5	2	2
Performs Equally with Both	3	4	4

For the 10 participants in the older adult study, we additionally have the participants' stated style to compare to the style they perform best with. Table 7.2 shows the performance groups separated by what the participants stated as their style preference. We can see that 3/10 of the participants stated the style that they performed best with, but 4/10 of the participants stated the opposite style to the one they performed best with. This shows that simply asking people their preference does not always help the robot choose the feedback style to optimize performance. Some potential reasons for the discrepancy include limited exercise experience, unfamiliarity with feedback styles, or selecting responses based on perceived rather than actual preferences.

Table 7.2: Performance and stated style preference for 10 participants in the older adult study

	Stated Firm	Stated Encouraging
Performs Best with Firm	2	4
Performs Best with Encouraging	0	1
Performs Equally with Both	0	3

We also compared how older adults performed with different feedback styles compared to those less than 60 years of age. We can see in Table 7.3 that the distribution of the performance groups is statistically different ($p < 0.05$ after performing an ANOVA) between adults (18) and older adults (11). In particular, a higher percentage of older adults appears to perform better with the **firm** feedback style, demonstrating the importance of collecting data from multiple age groups.

Table 7.3: Feedback style that participants performed best with, split by age group.

	Adult (< 60)	Older Adult (≥ 60)
Performs Best with Firm	16.7%	54.5%
Performs Best with Encouraging	33.3%	27.3%
Performs Equally with Both	50%	18.2%

7.2.2 Model Results

Section 7.1.1 explored the use of a contextual bandit approach using a complex simulated human model, but now we want to test the efficacy of the approach with our dataset of real human data. Using data from the 29 participants described in Section 7.2, we set the context in those data to be the estimated fatigue (low, moderate, or high). The model has the choice of two feedback styles that humans experienced: **firm** or **encouraging**. In our studies, participants experienced one round of exercise in the **firm** style and one with the **encouraging** style.

The goal of this test on real data is to see the actions the contextual bandit *would have* chosen given the observed context, and then estimate the reward it *would have* received based on the participants' performance. We can then compare that estimated reward to the participants' actual performance when they experienced one round of each feedback style. This should give us a sense of how well the bandit would have performed for this participant compared to a strategy of choosing roughly half of each style to present to the participant.

Let us assume we have a hypothetical Participant X who performed two rounds of exercise, one with each style. They performed 40 repetitions with the **firm** style first, followed by 35 repetitions with the **encouraging** style. Note that the participants did not perform the same number of reps with each style since rounds are time-based. They performed 80% of reps with the **firm** robot correctly and 70% with the **encouraging** robot correctly, with low fatigue.

We then start with the first repetition that they performed, which was with the **firm** style, and let us assume that the first repetition had good form. We give the model the context, which for the first repetition is low fatigue. We query the model as to the action it thought it should take, given the context. Let us assume the model chooses the **encouraging** style. We now train the model on what the human actually experienced (on the true data, not the action the model chose); specifically, the (context, action, reward) of (low fatigue, **firm**, good performance). We then continue this process through the rest of the repetitions. For each repetition, we compute the expected reward that the robot would have received had it chosen the style outputted by the bandit, for this fatigue level. In this example, we set the expected reward received by the bandit for this rep as 0.7, since the human is expected to perform 70% of the repetitions correctly with the **encouraging** style (the action chosen by the bandit). Note that we train the bandit on the true data (**firm**) and compute the bandit's expected reward based on the bandit's chosen action (**encouraging**).

We computed two different metrics to evaluate this approach. The first is the *actual reward* which is the total number of good performance reps the human actually performed through both rounds. The second is the *expected bandit reward* which is the sum of the expected rewards over all rounds. We hypothesize that the *expected bandit reward* will be higher than the *actual reward*, since the model should be optimizing for the human's best performance.

Figure 7.7 illustrates an example of this approach with one participant who performed much better with the **encouraging** style. The bandit quickly adapted to this performance difference and chose the **encouraging** style very frequently, which is optimal for this participant.

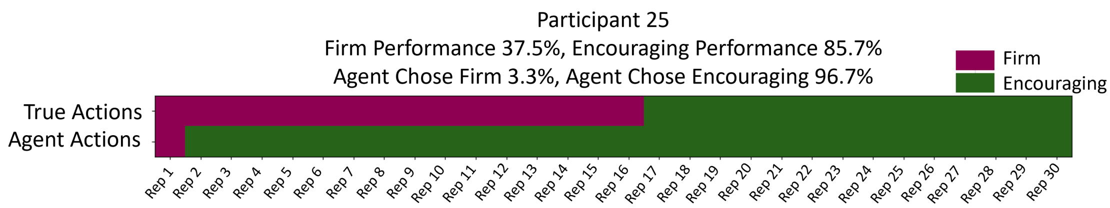

Figure 7.7: Example of model results for one participant. This participant performed much better with the **encouraging** style (37.5% with **firm**, 85.7% with **encouraging**), which the model quickly learns by choosing the **encouraging** style 96.7% of the time.

We calculated the difference in the *expected bandit reward* and the *actual reward* for each of the 29 participants. If this difference is positive, then the bandit performed better than the true data, which we call choosing *static styles*. If the difference is negative, then the bandit performed worse than choosing static styles. We used a similar grouping procedure as in Section 7.2.1 to find a margin of error based on a 95% confidence interval for the differences to be significantly different from 0. Figure 7.8 shows the reward differences for the 29 participants (sorted by difference) with the cut-offs between groups marked in red.

The adaptive agent performed better than choosing static styles for 15 participants, performed equally well as the static styles for 12 participants, and performed worse than the static styles for 2 participants. For the participant in Figure 7.7, the bandit had an expected reward that was 25% higher than the human's actual performance, which illustrates how an **adaptive** approach can learn the style with which someone performs best and use that style more frequently to optimize human performance.

Examining the participants for which the bandit performed worse, we can see that this occurs when the human performance with the two styles is very close. This means that the expected reward the bandit received at each trial is approximately the same and that combined with a limited number of iterations over which to learn (e.g., 38 reps instead of the 200 we performed in simulation) caused the bandit to not learn the best style to use for the human fast enough. For example, for Participant

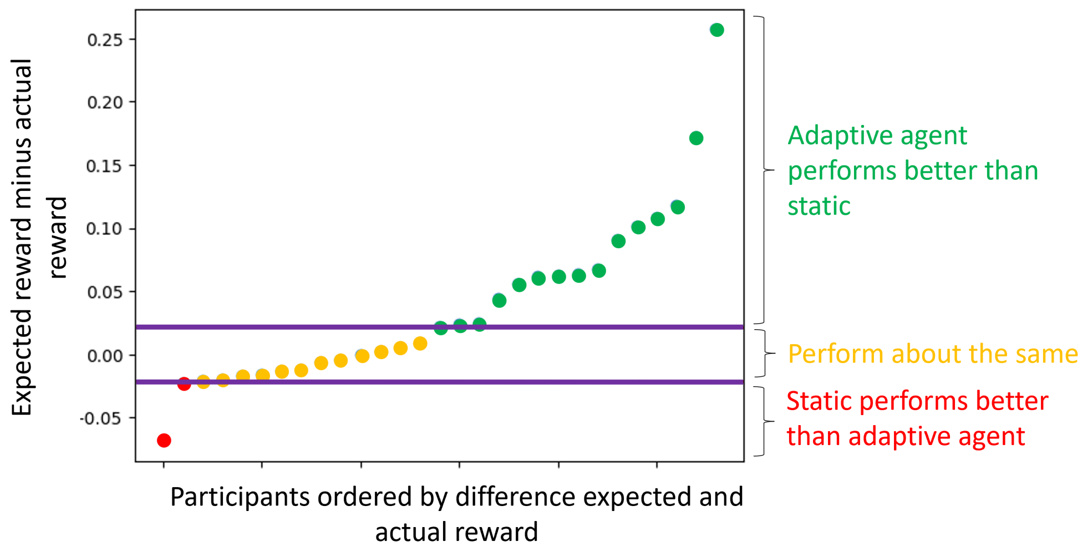

Figure 7.8: Expected bandit reward - actual reward for all 29 participants (x-axis is participant sorted by difference). The bandit outperforms the true data for most participants.

22 (Figure 7.9), their performance was 53% with the **firm** style and 68% with the **encouraging** style. At each rep, the bandit was receiving 0.53 or 0.68, and for the first 20 reps, it kept switching between the two styles. However, for the last 12 reps, it began to choose the **encouraging** style more frequently, which is the optimal choice for this participant.

Another reason this could occur is that the bandit only received a reward from one of the styles (first round of exercise) until the 20th rep, so it did not know the benefit of the other style until the second round. One approach we considered was interleaving the two rounds of exercise the participant performed for the bandit to learn on (choose the first rep from the first round, first rep from the second round, second rep from the first round, and so on). This could cause the bandit the learn much better because it is seeing examples of the human's performance on both styles earlier in the learning process. However, we do not believe this is a valid test for the bandit's ability because the human is also learning to perform the exercises when moving through the exercise session, and to remove the temporal element from the reps and present them to the bandit in a different order than what was actually

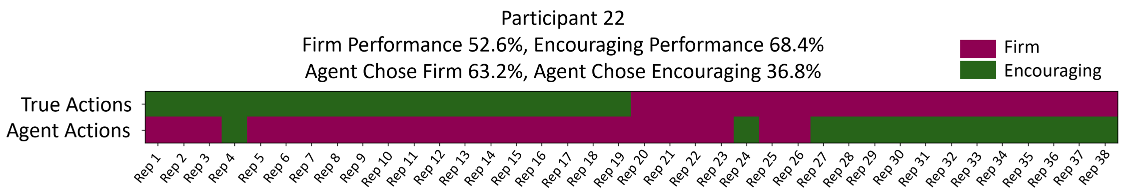

Figure 7.9: Example of model results for one participant, who performed approximately equally with the two styles (52.6% with **firm** and 68.4% with **encouraging**). The model does not learn their preferences as quickly since their performance with the two styles is quite close.

performed is not a true representation of how the human would exercise. We note, however, that the only way to truly determine if the bandit performs better than static styles is a user study, which we present in Section 7.3.

Additionally, we can compare the bandit's performance with no context (simple bandit) and fatigue as context (contextual bandit), similar to the comparison performed for simulated data in Figure 7.6. We calculated the difference between the contextual bandit reward and simple bandit reward for each participant and performed the same grouping procedure, with results shown in Figure 7.10. The two bandits performed roughly the same for most participants (17/29), and the contextual bandit outperformed the simple bandit for 6 participants and underperformed the simple bandit for 6 participants.

When the contextual bandit outperforms the simple bandit, it could be that the participant's performance was fatigue dependent, and the simple bandit did not have access to the fatigue in order to learn that fatigue dependence (resulting in a lower performance).

When the simple bandit outperforms the contextual bandit, it could be due to two possible reasons. One is that the fatigue estimate for the participants where we assumed 90% low fatigue was inaccurate, so the contextual bandit was receiving inaccurate contextual information over which to learn. The second reason could be that the participant's performance was not fatigue-dependent. The contextual bandit will naturally take longer to learn than the simple bandit in this case as it is learning three identical models (one for each fatigue level), and the simple bandit is only learning one.

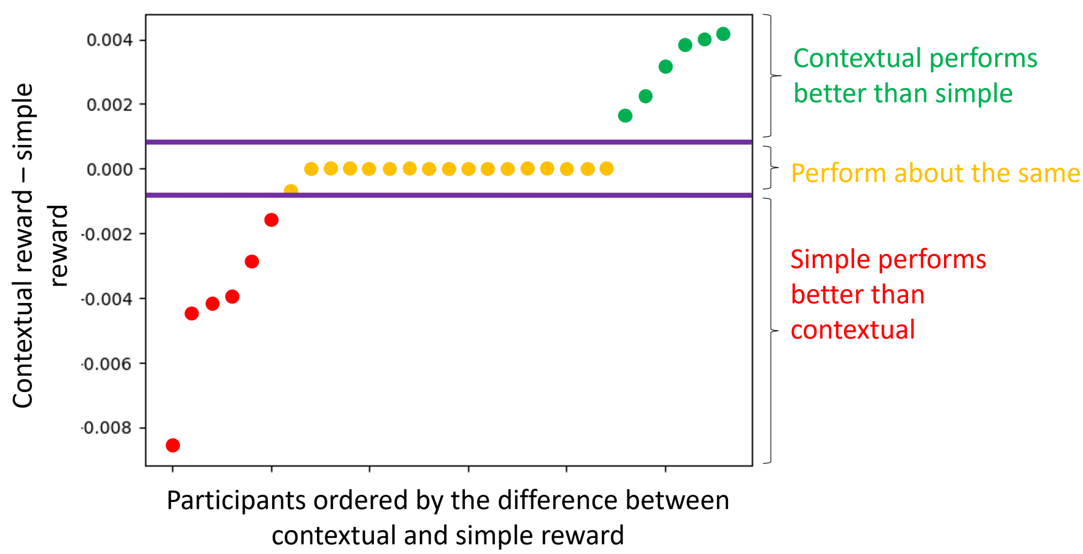

Figure 7.10: Expected contextual bandit reward - expected simple bandit reward for all 29 participants (x-axis is participant sorted by difference).

7.3 Adaptive Study Design

The previous sections showed that the contextual bandit approach performs well in simulation with complex contextual situations and performs well with human data (compared to approximately 50% with each style). We ran our main user study in this chapter to test whether this **adaptive**, contextual bandit approach for choosing feedback styles in real-time would outperform choosing a static baseline. We implemented three different coaching styles. The first two are **firm** and **encouraging**, implemented identically to Chapter 6, with the small change of fatigue-informed verbal phrases (as explained in Section 7.2.1).

The third is **adaptive**, where the robot uses a contextual bandit to determine which style to use when giving feedback. As mentioned in Section 6.3, the robot does not give verbal feedback after every repetition, as that would be overwhelming to the human. Instead, the robot provides verbal feedback when certain conditions have been met (e.g., 3 good reps in a row, 2 mistakes in a row). As the human is exercising, the robot observes their performance, and when a condition for verbal feedback has been met, it first observes the context. We chose fatigue as our context, as explained in Section 7.2, where we categorize the human's fatigue as low, moderate, or high according to their heart rate. Once the robot observes the context, it chooses a feedback style (**firm** or **encouraging**) for the feedback, and then observes the human's performance on the next repetition. It trains on the (context, action, reward) combination to improve its action choice the next time the robot needs to provide verbal feedback. Additionally, for the reps that the human performs until the next verbal feedback, we assume that the effect of that verbal feedback continues, so we train on the human's performance on those repetitions as well.

The robot's nonverbal feedback is designed in the same way as in Chapter 6. The robot selects an action/style only every time it is giving verbal feedback, and its nonverbal feedback style adjusts along with its verbal feedback style when that action is chosen. For example, if the robot chooses the **encouraging** style for its next verbal utterance, it will also adjust its facial expressions and body movements to be in the **encouraging** style. Table 6.2 includes several examples of how the **firm** and **encouraging** styles differ in the same contextual situation for verbal and nonverbal feedback.

We define the *effective style* for a participant to be the style that matches their performance group and the *ineffective style* to be the opposite style. For example, a participant in the **firm** performance group would have an effective style of **firm** and an ineffective style of **encouraging**. For people who perform equally well with both styles, we consider **firm** and **encouraging** to both be *effective styles*.

Our hypotheses are as follows:

- **H1:** The style with which the participants perform best does not necessarily match what they rate the highest or what they state as their preference.
- **H2:** Participants perform better on average with the **adaptive** style than the style they state as their preference.
- **H3:** Participants perform better with the **adaptive** style compared to their ineffective style and perform equally well with the **adaptive** style and their effective style.
- **H4:** Some participants will match each of the three contextual situations explored in Section 7.1.1 (no fatigue dependence, basic fatigue dependence, and complex fatigue dependence).

7.3.1 Study Procedures

Participants began this study by wearing a heart rate monitor and completing a consent form and brief demographics that included gender, age, ethnicity, familiarity with robots and programming, and level of education (see Section A.1 for the form and demographics administered). They were then introduced to the two exercises, bicep curls and lateral raises, as well as a set of dumbbells (3 lb each) that they had the option of using during the exercises. Ideally, we would have included several dumbbell choices as in the study in Section 6, but we only had this one set available to us when running this study. They were instructed that they could choose whether to use the dumbbells and could use them for part of the session, according to their comfort. Before starting the exercise session, they were asked about their feedback preference with the following question: "The robot wants to be the best coach it can for you. Which of these types of feedback would you prefer"

- I prefer feedback like "Focus on improving your range. Keep pushing!" when I make a mistake
- I prefer feedback like "Amazing effort! Let's aim for full range next time" when I make a mistake
- I have no preference between these two options

After these explanations and introductory questions, the participants began one of the three rounds of exercise (see the setup in Figure 6.7). In each round, they performed four sets of exercises: two bicep curls followed by two lateral raises (very similar to the study procedures in Section 6.4.1). During each set, the robot provided feedback based on the style it was using, where the style for each round was randomized between (**firm**, **encouraging**, and **adaptive**) using a Latin square. Note that the **adaptive** style only learned within the **adaptive** round, so even if the participant saw the **adaptive** style in the third round, it would not learn from the previous two rounds performed.

Each set was completed in 45 seconds, and participants rested for 40 seconds between sets. This is a shift from the study presented in Chapter 6 as each set was strictly based on time rather than based on the number of reps the participant performed in addition to time. We found that the complex condition of time and number of reps resulted in participant confusion as to how long each set would be.

In between each round of 4 sets, participants completed a short survey (the same survey as in Section 6.4.1 and is included in Section A.1), which has questions to measure their perception of the robot in terms of animacy, likability, and perceived intelligence. They also had the option to write about their perception of the robot in that round.

7.4 Adaptive Study Results

Our study protocol was approved by the CMU IRB, and we used the CMU Center for Behavioral and Decision Research to recruit participants from both CMU and non-CMU sources, for a total of 24 participants. We had a mix of 6 older adults (60+) and 18 younger adults (< 60), with $\mu = 34.5, \sigma = 20.2$.

7.4.1 Performance vs. Ratings vs. Stated Preference

As seen in Section 7.2.1, the style that someone performs best with may not align with the style they rate the highest or the style they state that they prefer. Figure 7.11 illustrates this concept, with participants divided into younger and older adults. The data is also aggregated in Tables 7.4 and 7.5, showing the differences split by older adults, younger adults, and all participants.

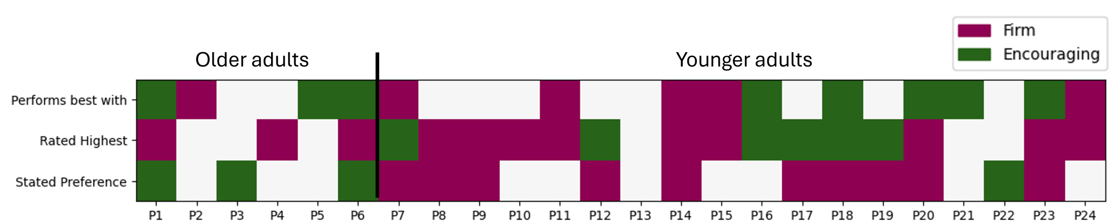

Figure 7.11: The style the participant performs best with, rates the highest, and states as their preference before exercising, separated out by younger and older adults. **Firm** is shown in maroon, **Encouraging** in green, and neither/no preference in white.

We first compute the percentage of good performance reps for each participant per round. We then compute the difference between their **encouraging** and **firm** performance and assign each participant to one of three groups: **Firm**, **Encouraging**,

Table 7.4: The style the participant performs best with compared to the style they rated the highest. The data are split by older adults, younger adults, and all participants.

	Older Adults			Younger Adults			All Participants		
	Rated Firm Highest	Rated Encouraging Highest	Rated both equally	Rated Firm Highest	Rated Encouraging Highest	Rated both equally	Rated Firm Highest	Rated Encouraging Highest	Rated both equally
Performs Best with Firm	0	0	1	4	1	0	4	1	1
Performs Best with Encouraging	2	0	1	2	2	1	4	2	2
Performs Equally with Both	1	0	1	3	3	2	4	3	3

or **Neither** (using the same grouping procedure as in Section 7.2.1, which calculates a 95% confidence interval from the performance difference values). This results in 6 participants in the **Firm** group, 8 participants in the **Encouraging** group, and 10 participants in the **Neither** group. The first row of Figure 7.11 indicates the performance groups computed for each participant: **Firm** (maroon), Encouraging (green), or **Neither** (white).

We performed the same grouping procedure for the differences in subjective scores between the **firm** and **encouraging** rounds, resulting in **Firm**, **Encouraging**, and **Neither** groups based on the participants' likability ratings of the **firm** and **encouraging** styles (second row of the figure). The third row indicates what each participant chose when asked which style they prefer (**firm**, **encouraging**, or no preference). An interesting trend in these results is that all older adults who indicated a style preference stated **encouraging**, but all those who had a preference after experiencing the robots rated the **firm** style higher. Furthermore, almost all young adults who stated a style preference stated **firm**, but they had mixed results when rating styles after experiencing them.

Only considering a choice of **firm** or **encouraging** (ignoring the neither group), we can see that only 4 participants performed best with the style they stated at the beginning of the session, and 3 participants stated the opposite style to what they

Table 7.5: The style the participant performs best with compared to the style they stated as their preference. The data are split by older adults, younger adults, and all participants.

	Older Adults			Younger Adults			All Participants		
	Stated Firm	Stated Encouraging	Stated no preference	Stated Firm	Stated Encouraging	Stated no preference	Stated Firm	Stated Encouraging	Stated no preference
Performs Best with Firm	0	0	1	2	0	3	2	0	4
Performs Best with Encouraging	0	2	1	3	0	2	3	2	3
Performs Equally with Both	0	1	1	5	1	2	5	2	3

performed best with. Additionally, 6 participants performed best with the style they rated the highest, and 5 performed best with the opposite style they rated highest. Specifically among older adults, only two participants stated the style with which they performed best and none rated the style they performed best with the highest. This shows us that simply asking people their style preference or even analyzing survey results after they experience the styles does not give the robot a sense of which style it needs to choose to optimize performance, especially for older adults. The adaptation to performance using the contextual bandit is required for the robot to know which style to choose.

7.4.2 Stated Style vs. Adaptive Performance

Next we compare each participant's performance with the style they stated at the beginning and with the **adaptive** style. Figure 7.12 shows the two distributions for only the participants who indicated a preference for **firm** or **encouraging** (excluding the 10 participants who stated no preference). The average performance with the **adaptive** style was 90.3%, and the average performance with the stated style was 88.3%.

Although the differences in the two distributions are not statistically significant, 9 of the 14 participants represented in the figure performed an average of 8.1% better with **adaptive** than their stated preference with the performance improvement ranging from 0.8% to 15.2%. The remaining 5 participants performed an average of 8.9% better with their stated preference, with the performance improvement ranging from 1.7% to 16.7%. This illustrates that, for most participants, choosing their stated preference does not perform better than our **adaptive** method.

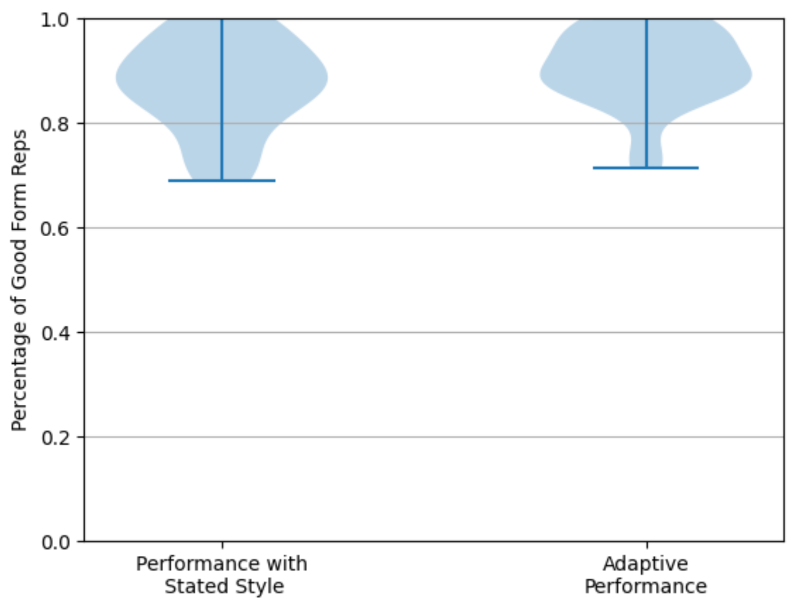

Figure 7.12: Performance comparison with stated style preference and with the **adaptive** style for the 14 participants who chose either **firm** or **encouraging** for their stated preference at the beginning of the session.

7.4.3 Performance per Performance Group

We compare the performances with each style of the participants in each performance group (Figure 7.13). To determine whether one performance distribution is statistically significantly greater than another, we use a Mann-Whitney U with a single-sided alternative. To determine whether two performance distributions are statistically similar, we first calculate the mean difference between the two performance distributions and perform an independent t-test to obtain the standard error of the difference in their means. We calculate the critical t-value for a 95% confidence level and compute the margin of error and the confidence interval. If the difference in means is within the confidence interval, we conclude with 95% confidence that the two distributions are similar in mean.

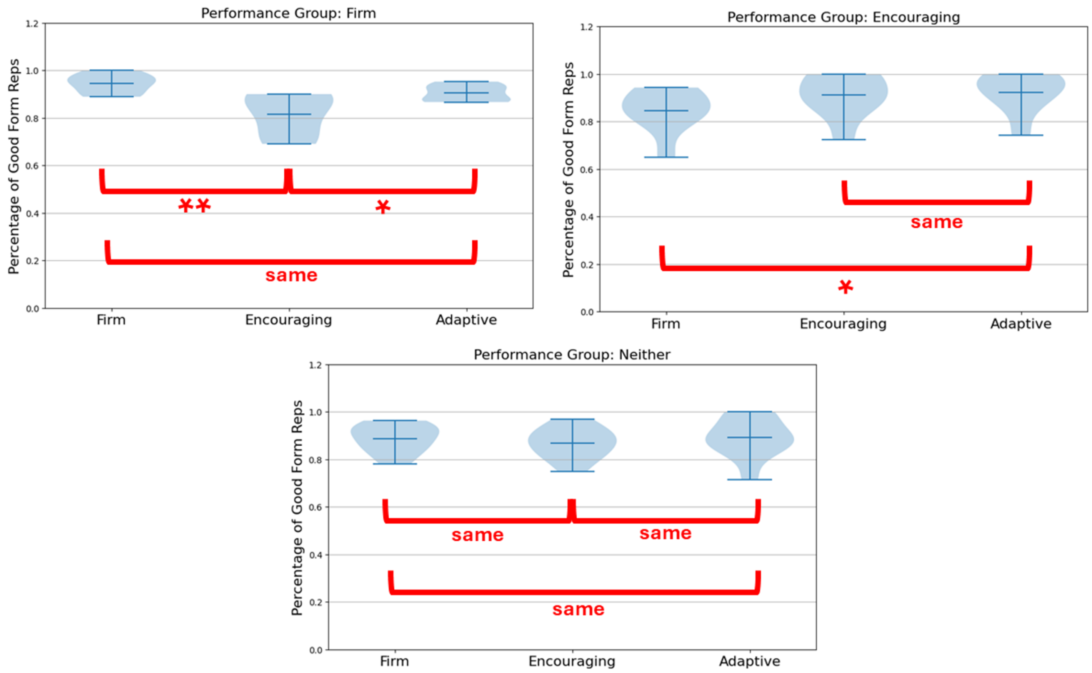

Figure 7.13: Performance comparison by performance group for each style. Statistical differences are marked at the $p < 0.05$ and $p < 0.01$ level with $*$ and $**$ respectively, and groups that are similar at a 95% confidence level are marked as "same"

For the participants in the **firm** performance group, their performance with **firm** was better than their performance with **encouraging** ($p < 0.01, U = 33.5$), and their performance with **adaptive** was better than their performance with **encouraging** ($p < 0.05, U = 27.0$). Furthermore, their performance with **firm** and **adaptive** was statistically similar at a 95% confidence level. This means that for those who performed better with **firm**, the bandit correctly adapted to perform at least as well as the effective style for them.

For the **encouraging** performance group, their performance with **encouraging** was not statistically higher than that with **firm** ($p = 0.09, U = 45.0$), although it was trending in that direction. Their performance with **adaptive** was better than their performance with **firm** ($p < 0.05, U = 51.0$). Their performance with **encouraging** and **adaptive** were statistically similar at a 95% confidence level. This means that for those who performed better with **encouraging**, the contextual bandit was able to perform at least as well as if it knew beforehand that they performed better with **encouraging**.

Lastly, for the **neither** performance group, all three performances were statistically similar at the 95% confidence level. The bandit performed at least as well as choosing either static style.

7.4.4 Performance by Context

In Section 7.1.1, we explored different fatigue scenarios: no fatigue dependence, basic fatigue dependence, and complex fatigue dependence. We showed that the contextual bandit could handle those different contextual situations with simulated data, so we can see if those situations occur in the real data. Figure 7.14 illustrates each participant's performance split by fatigue level, by looking at their performance difference between the **firm** and **encouraging** rounds. The first row shows whether the participant's performance during low fatigue indicates performing better with the **firm** style, the **encouraging** style, or **neither**. The second row illustrates the same for moderate fatigue, and the third row for high fatigue (no participant reached a high fatigue during this study). Additionally, some participants did not reach moderate fatigue in both **encouraging** and **firm** rounds, so their performance group is automatically assigned as **neither** (white).

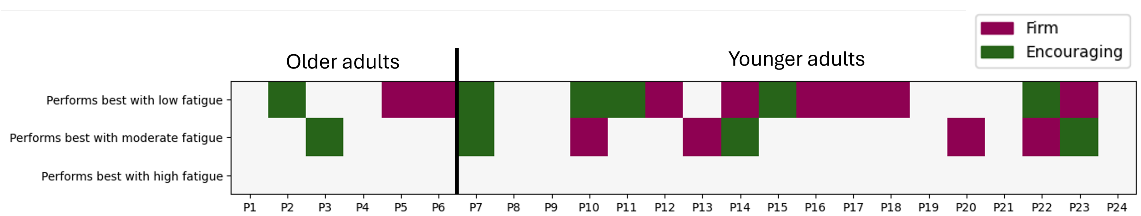

Figure 7.14: The style the participant performs best with across all fatigues as well as by fatigue level, separated out by younger and older adults. **Firm** is shown in maroon, **encouraging** in green, and neither/no performance difference in white.

We can see that 8 participants were assigned to the same performance group from low to moderate fatigue. P7 has consistent good performance with the **encouraging** style, with an approximately equal performance with low and moderate fatigue. This is an example of no fatigue dependence, where performance is not affected by increasing fatigue level. Taking a look at P9 for example, their performance with both styles decreases slightly when going from low fatigue to moderate fatigue (95% → 88% for **firm** and 90% → 89% for **encouraging**), which is an example of basic fatigue dependence. The optimal choice for the bandit does not change with fatigue, but the human's performance does decrease overall as fatigue increases.

Complex fatigue dependence, where the optimal style choice changes with fatigue, occurs for several participants. P23 is an example of when the robot should choose the **firm** style for low fatigue and should choose the **encouraging** style for moderate fatigue. The bandit picked up on this fatigue dependence and chose **firm** 90% of the time for low fatigue and chose **encouraging** 67% of the time for moderate fatigue.

P22 is an example of when the robot should choose the **encouraging** style for low fatigue and should choose the **firm** style for moderate fatigue. The bandit in this case did not recognize this fatigue dependence and chose **encouraging** 26% of the time for low fatigue and 38% of the time for moderate fatigue. However, the performance for both styles for P22 was quite close over fatigue levels (within about 10%), so the bandit might have needed more data in order to properly learn this subtle contextual difference.

7.5 Adaptive Study Discussion

H1: The style with which the participants perform best does not necessarily match what they rate the highest or what they state as their preference. We have support for this hypothesis. In Section 7.2, we found that for the participants in the collected data set that the style with which they performed best with did not necessarily match what they rated the highest or what they stated as their preference. As shown in Figure 7.15, this trend was also observed in this user study. Only 9% of participants were consistent across all 3 groups, and 9% of participants differed across 3 groups. This indicates that simply asking the human about their preferences will generally not result in a choice of feedback style to optimize performance. This could be due to limited exercise experience, unfamiliarity with feedback styles, or selecting responses based on perceived rather than actual preferences.

This could also be due to the human preferring one style of feedback, but performing better with another style. For example, someone may like a lot of encouragement

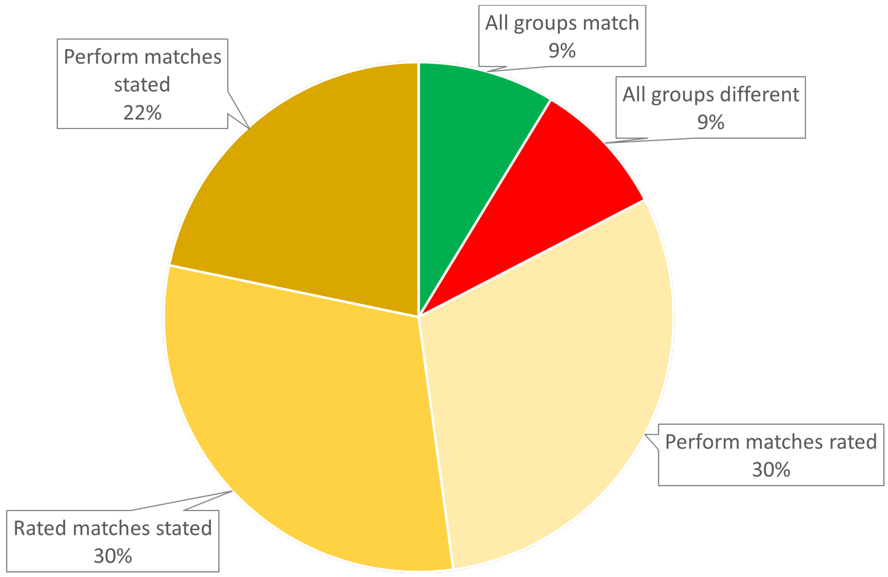

Figure 7.15: Pie chart indicating the percentage of people that matched across 3 groups: style they performed best with, style they rated the highest, and style they stated as their preference. This chart aggregates the information shown in Figure 7.2.1

while exercising, but they would perform better when being pushed by firmer feedback. However, something to note here is that, over the long term, a robot that continuously chooses the feedback style to optimize performance rather than preference could suffer from low engagement or a growing dislike of the robot. In a long-term scenario, it may be beneficial to include other variables into the robot's reward that take the human's enjoyment and preferences into the account to promote long-term use of the robot (perhaps with the downside of lower short-term performance).

H2: Participants perform better on average with the adaptive style than the style they state as their preference. We have support for this hypothesis. For the 14 participants who chose either **firm** or **encouraging** as their stated style preference, they performed on average 2% better with the **adaptive** over their stated style. 9 of the 14 participants performed strictly better with the **adaptive**, with an average improvement of 8.1%. This illustrates that for most people, simply choosing a style that they state at the beginning of a session is not necessarily optimal, and that our approach can perform better than that static choice. Furthermore, as seen in Figure 7.11, of the 10 participants who did not declare preference for one of the two styles, 4 performed better with **firm** and 3 with **encouraging**. The **adaptive** model should pick up on that performance difference that the participants themselves were not aware of.

In fact, when examining what the **adaptive** style did for these participants, it chose **encouraging** most often for 2/3 of the participants who performed better with encouraging and chose **firm** most often for 3/4 of the participants who performed better with **firm**. The two remaining participants had less than a 10% difference between their **encouraging** and **firm** performances, which could mean that the bandit needed more time to learn to pick up on their small performance differences.

H3: Participants perform better with the adaptive style compared to their ineffective style and perform equally well with the adaptive style and their effective style. We have support for this hypothesis. Participants in the **firm** performance group performed statistically similar with **firm** and **adaptive**, and performed statistically better with **adaptive** compared to **encouraging**. Additionally, participants in the **encouraging** performance group performed statistically similar with **encouraging** and **adaptive**, and performed statistically better with **adaptive** compared to **firm**. This shows that the bandit performed as well as the *effective style* for these participants and better than the *ineffective* style

For participants with no significant performance difference between the two styles, the **adaptive** approach performed equally as well as the static baselines of **firm** and **encouraging**. The choice of the bandit is not particularly important for these participants, but it at least performed as well as choosing one of the two static styles.

H4: Some participants will match each of the three contextual situations explored in Section 7.1.1 (no fatigue dependence, basic fatigue dependence, and complex fatigue dependence). We have support for this hypothesis. Figure 7.16 shows the percentage of participants who matched each of these three contextual situations. Some participants did not reach multiple fatigue levels in all rounds, so we had insufficient data to analyze whether they matched a particular contextual situation. As we can see from the pie chart, we have participants who match each of the different types of fatigue dependence.

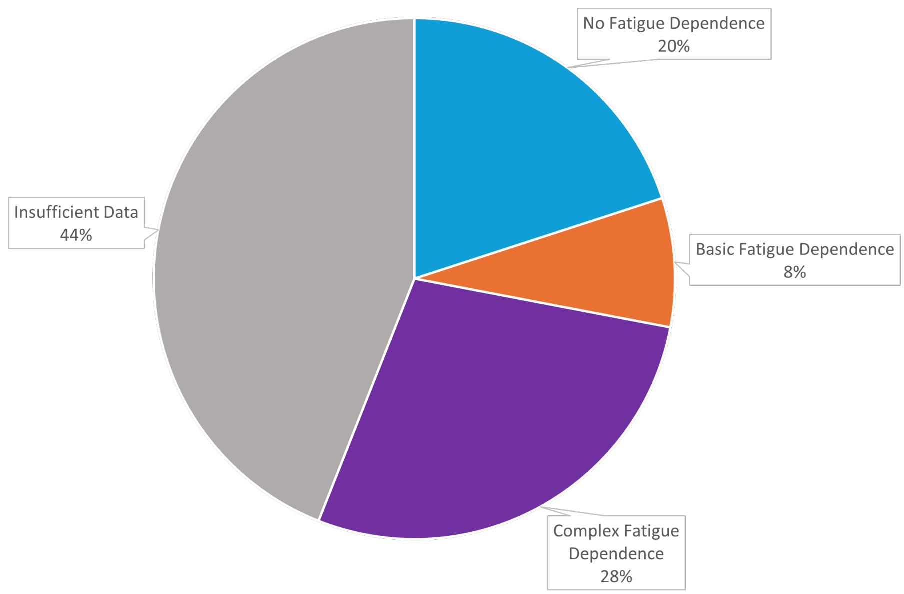

Figure 7.16: Pie chart indicating the percentage of people that match each of the three contextual situations explored in simulation: no fatigue dependence, basic fatigue dependence, and complex fatigue dependence. Participants that did not reach multiple fatigue levels in all rounds are marked as insufficient data.

As discussed in Section 7.4.4, we have participants who followed the pattern of no fatigue dependence (such as P7), where their performance was unaffected by the context. If they performed best with a particular feedback style, they performed best with that style for all fatigue levels with no decrease in performance as fatigue increased. We also have participants who followed the pattern of basic fatigue dependence (such as P9), where their performance decreased as a result of increasing fatigue, but the bandit's optimal choice of style did not change. This could indicate a higher likelihood of sloppier form as they became more tired, but their preference for **firm** or **encouraging** feedback not being affected by that fatigue.

Lastly, we have some participants who followed the most complex scenario, where their optimal style choice changed with fatigue level. For some of these participants, such as P23, the bandit learned this complex dependence on fatigue, accurately learning the choice of the other feedback style when the human became fatigued. However, for others, such as P22, the robot did not pick up on that contextual nuance, though this may be simply due to not enough data to train the bandit for the moderate fatigue level. The existence, however, of some participants in this complex fatigue-dependent scenario, justifies the use of a contextual bandit, rather than one that does not take context into account. The fatigue information was important for the robot to take into consideration when choosing the feedback style for the human, and the bandit was able to learn this fatigue dependence in an online fashion.

Summary: In this work, we show that a contextual bandit approach to choosing robot feedback style to optimize exercise performance effectively learns the optimal feedback style choice for people who perform well with either a **firm** style, **encouraging** style, or equally with both. We show supporting results for complex contextual situations (in this case, varying fatigue levels and fatigue-dependent performance) using a simulated human model, as well as results from the bandit using data on a real-world dataset. Lastly, we show that this **adaptive** approach performs well in a user study compared to simply choosing static baselines.

8 Conclusion and Future Directions

We conclude with insights from our thesis on personalized context-aware robot feedback, as well as limitations and future directions for this work.

8.1 Key Attributes of Applicable Problems and Design Considerations

We explored problems in two domains, education and exercise, in this work, but our general approach of adaptation using a contextual bandit can be applied to a wider range of applications. The agent need not necessarily be a physical robot; it could also be a simulated robot or a virtual agent. Specifically, for the adaptation explored in Chapter 7 to be appropriate, the problem explored must have the following:

1. One agent providing feedback to one human (or a group of humans treated as a singular entity over which to learn preferences)
2. The human completing a series of similar tasks, allowing the agent to estimate performance and provide corrections
3. The agent controlling at least one nonverbal modality, with some examples of these modalities provided below
    - Visual modalities: facial expressions, gestures, eye gaze, body language, lights, and colors
    - Auditory modalities: tones and haptic feedback
    - Tactile modalities: force feedback and vibration feedback

- Proxemics: orientation, speed of motion, and approach/retreat

Given these criteria, here are some examples of problems and domains in which this adaptation could be applicable.

- Physical Therapy Assistance: a virtual agent or robot can guide patients through rehabilitation exercises
- Cognitive Stimulation Program: a virtual agent could assist with intellectual exercises for older adults to help prevent cognitive decline
- Surgical Training Simulations: an agent could provide feedback to doctors learning how to perform simulated surgical procedures
- Customer Service Training: an agent could provide feedback to those training to handle customer interactions
- Music Learning Assistance: a virtual tutor could analyze piano playing technique and provide personalized feedback

After choosing a problem of interest, there are many design considerations to explore, especially with regard to context. Throughout this work, we have emphasized the importance of context in the robot's feedback generation. Chapter 4 showed that a mixture of task-related and facial features were most effective for the prediction of engagement early during an educational activity, and Chapter 5 found that nonverbal robot behavior that took into account the human's performance on the previous round actually improved human learning, even when the nonverbal behavior did not provide additional task-related information to the human (simply reacting affectively). And finally, the exercise coach presented in Chapters 6 and 7 used a plethora of contextual information including human fatigue estimates, task information (e.g., which exercise the human is performing), and human performance not only to provide relevant multimodal feedback, but also to adapt the robot feedback style to personalize to what the human will perform best with.

When designing a context-aware human-robot interaction for another domain, there are many questions to ask as a researcher that can help in the design process. Some of these questions include:

- What sensing and processing capabilities does the robot platform have?
- What feedback modalities does the robot platform have?

- What is the role of the robot in the interaction?
- What is the role of robot feedback in the interaction?
- What factors are important for the robot to incorporate into its feedback?
- If the robot is going to adapt over time to personalize its behavior, what factors are important for the robot to know for that personalization and how does that personalization change the robot's behavior?

Answering these questions can help the researcher determine what is useful to include as context for the robot and what information is feasible for the robot to compute. It also aids in understanding how the robot can translate that context into varying behaviors.

8.2 Verbal and nonverbal feedback

Both the sorting game robot (Chapter 5) and the exercise coach robot (Chapters 6 and 7) used a combination of verbal and nonverbal feedback to communicate with the human. As especially demonstrated in Chapter 5, nonverbal feedback plays a crucial role in the human's enjoyment, subjective experience, and performance. The exercise coach robot reacted to the human both verbally and nonverbally, with those reactions combined in a way where the nonverbal reaction emphasized the verbal reaction (smiling along with an encouraging verbal utterance) and filled the gaps between verbal utterances (smiling after the human performs the last rep with good form even if the robot does not say anything verbally to acknowledge). We also noticed that some people particularly noticed the robot's nonverbal behavior, exclaiming when the robot smiled at them. Others did not seem to notice changes in the robot's behavior, but this does not mean the difference was not subconsciously affecting their perception of the robot. For example, many participants in the study in Section 6.5 reported not noticing differences between the **firm** and **encouraging** robot styles, yet their ratings and performance varied with the two styles. Whether the robot's nonverbal behavior is actively noticed or not, it plays a vital role in the human's interaction with the robot, and paired with verbal feedback, it can enhance both the human's experience and performance. Based on the results from 5 and 6, we recommend ensuring that both verbal and nonverbal modalities are used and aligned to the contextual situation to improve both performance and subjective experience.

8.3 Balancing Competing Objectives

One takeaway from Chapter 7 (seen in Tables 7.1 and 7.2 and Figure 7.11) is that the feedback style that optimizes performance for an individual does not necessarily match the style they prefer. This leaves the robot with a decision: adapt to improve performance to the detriment of subjective experience, or adapt to improve subjective experience to the detriment of performance? We chose to adapt in order to improve performance, and this approach was successful. The contextual bandit determined the style that optimized performance for most participants, even in complex, context-dependent scenarios.

However, if we were to deploy a robot using this type of personalization in a longer-term scenario, it might not have the desired impact. If the robot continues to use the style that optimizes short-term performance but is not preferred by the human, they may lose interest and stop exercising. If the robot's larger goal was to promote long-term exercise performance, it could fail, as the human may stop using the robot if it continues giving feedback in a style they dislike. To address this issue, the robot's reward could be modified to include subjective measures, optimizing for a combination of performance and preference over the long term.

This could involve adding terms to the reward based on positive comments made by the human about the robot's feedback (participants, especially older adults, tended to respond verbally back to the robot after feedback) or even more subtle signals like facial expressions. Another approach is to create pseudo-rewards based on the human's ratings of the robot. For example, if the robot used mostly **firm** feedback in a particular set, and the human indicated dissatisfaction in a survey immediately afterward, we could reduce the reward for the robot in the next set. This reduction would apply each time the robot used the **firm** style, compensating for the lower human enjoyment.

Once an enjoyment measure is obtained, a new challenge emerges: How do we determine the correct balance between performance and enjoyment for each individual? For example, if we constructed a reward that weighted performance with α and enjoyment with $1 - \alpha$, α becomes another parameter to learn. One approach could be to weight performance only ($\alpha = 1$) for the first session, then adjust α based on whether the style that optimizes the human performance differs from their preferred style. Taking a more longitudinal approach could allow the robot to establish a baseline during the first session and then use those estimates to adapt for the remaining sessions.

8.4 Limitations and extensions of the contextual bandit formulation for an exercise coach

The contextual bandit formulation presented in Chapter 7 effectively demonstrated the robot's ability to adapt in real time, selecting the feedback style that optimized the human's performance. There are several extensions of this work that address some of the simplifications we made in this formulation. First, we set the reward the bandit received as 0 for bad performance and 1 for good performance. However, incorporating human enjoyment or style preference could improve the long-term impact of the exercise coach, as mentioned earlier. Additionally, including a more nuanced reward, with performance values between 0 and 1 (e.g., right side had good form, but left side had bad form resulting in a reward of 0.5), could improve the coach's adaptation. Other extensions include adding more feedback styles (e.g., very firm, very encouraging, neutral) to diversify the robot's feedback, and introducing additional exercises to increase variety in the exercise session itself. Additionally, the robot's verbal feedback in Chapter 7 was generated offline by an LLM, and having the LLM generating context-aware phrases in real-time would further increase the robot feedback's variety.

Another possibility is to add additional contextual variables to the context on which the bandit trains. This could include more nuanced performance information, exercise type, etc. When increasing the number of contextual variables (e.g., the number of fatigue levels), it is important to note that the more contexts the bandit has to model, the slower the learning process will be. Therefore, careful consideration should be given to which contextual variables are important for the robot to use when adapting its feedback style. For example, knowing the exercise the human is performing is important for generating accurate verbal feedback, as it helps the robot interpret performance and provide relevant corrections. For instance, if the human was performing lateral raises and the evaluation indicated incomplete range of motion on the last repetition, the robot could say, "Make sure you raise your arms all the way to 90°". The choice of exercise informs the robot's verbal phrase; however, it may not be necessary for the bandit to train a separate model for each exercise, as the style of feedback likely remains the same regardless of the exercise. Choosing the right contextual variables for adaptation is critical. Care should be taken to include only relevant variables that affect the choice of feedback style, avoiding unnecessary variables that could slow down the learning process.

In general, the contextual bandit approach forms a solid foundation for context-aware adaptation of robot feedback and remains flexible for these extensions, as well as further investigation into the long-term effects of this personalization approach.

8.5 Other Future Directions

Lastly, there are further extensions of this approach that could improve both the feedback and the overall human experience. Adding a measure of the human's subjective experience to the human model, and possibly incorporating it into the reward function for the bandit, could help the robot balance the competing outcomes of performance and enjoyment. This approach treats the human-robot interaction as unidirectional: the robot provides feedback to the human, but does not incorporate any feedback the human gives about their experience and preferences (besides context estimates used to generate and adapt the feedback). Adding a social and conversational aspect, where the robot and human engage in back-and-forth conversation (even during breaks between tasks), could improve the human's experience and provide the robot with valuable information, such as the human's mood and motivation level. The robot could also converse with the human about topics unrelated to the task, building rapport and increasing the human's enjoyment of the interaction. Lastly, undetrsanding how human's long term progression and enjoyment should influence robot behavior is another important extension of this work.

8.6 Conclusion

In this work, we present personalized context-aware multimodal robot feedback that focuses on tailoring the robot's responses to individual users, while considering the dynamic context that shapes the way humans interact. We first considered estimating *context* in Chapter 4, modeling key aspects of the human state related to engagement when approximately 80% of an educational activity remained, which could allow an agent to provide feedback early enough to improve the human experience. We then generated *nonverbal affective* robot behavior in Chapter 5 and showed that aligning the robot's conveyed emotion with affective movement positively impacted the human's performance in a sorting game. In Chapter 6, we designed a physical robot exercise coach and demonstrated changes in human perception and performance with different robot feedback styles. In Chapter 7, we developed a *personalized context-aware* robot that used a contextual bandit approach to dynamically adapt its feedback style to optimize human performance, learning over time which style to apply and when. Lastly, in this chapter, we presented key takeaways and discussed future directions of this work.

As robots become increasingly prevalent in human-facing domains, developing personalized context-aware robots that provide multimodal feedback will foster engaging and productive human-robot interactions. We believe that our work is a step toward achieving that goal.

A Appendix

A.1 Study Form

This participant form was given to participants for the study presented in Section 7.3.

A.1.1 Consent Form for Participation in Research

Principal Investigator Reid Simmons The Robotics Institute, NSH 3213 Pittsburgh, PA, 15213 Email: rsimmons@andrew.cmu.edu (412) 268-2621

Sponsor(s) : NSF (National Science Foundation) âĂŞ AI-CARING âĂŞ Georgia Tech

Purpose of this Study The purpose of the study is to evaluate and compare different coaching behaviors for the Quori robot to design a personalized, context-aware robot exercise coach.

Procedures You will be asked to perform an exercise session with the Quori robot. The robot will converse with you before the session begins. The robot will ask you to perform a series of exercises (see information sheet), with rest times included after each set of the exercise is completed. You can use the light dumbbells available to make the exercise as difficult as you are comfortable, but you always have the option to use only your bodyweight and can drop or switch the exercise equipment whenever desired. In all cases, the robot will not be mobile and will only move its arms and torso during the session. You will always be a safe distance away from the robot.

To gather data, you will be asked to fill out questionnaires to rate your experience of the tasks and perception of the robot. Video cameras and microphones will record you during the study. We will be using a third-party transcriber on the audio recorded. We may also ask you to wear a heart rate monitor to help the robot monitor your fatigue. These videos will be used to analyze your performance and responses, which provide data for assessing our robot technology. Experimenters may assist the robot with some parts of the study. This assistance may not be visible to you. If you do not wish to be audio and video recorded, you should not participate in this study.

The study will take up to 45-60 minutes.

Participant Requirements Participants must be adults age 18 or older. You must be comfortable performing basic exercise motions and not have significant mobility limitations that would prevent the execution of the exercises

Risks The risks and discomfort associated with participation in this study are no greater than performing exercises with low resistance at home or at a gym. As seen on the information sheet, you can utilize bodyweight or dumbbells to do the exercises at your preferred difficulty level. You can make modifications to the exercises (range of motion, speed of movement, etc.) based on your comfort and fatigue level. If at any point you need an additional rest or would like to stop the session, you may do so.

Benefits You will receive the health and wellness benefits associated with a short exercise session. Your participation will be beneficial to the our research and will help us evaluate the robot exercise coach.

Compensation and Costs You will be compensated with $20 in cash upon completion of the study, and we may offer transportation compensation if necessary. You are expected to complete the study, but in the case of partial completion, partial payment will be remitted.

There will be no cost to you if you participate in this study.

Future Use of Information In the future, once we have removed all identifiable information from your data, we may use the data for our future research studies, or we may distribute the data to other investigators for their research studies. We would do this without getting additional informed consent from you. Sharing of data with other researchers will only be done in such a manner that you will not be identified.

Confidentiality By participating in the study, you understand and agree that Carnegie Mellon may be required to disclose your consent form, data and other personally identifiable information as required by law, regulation, subpoena or court order. Otherwise, your confidentiality will be maintained in the following manner:

Your data and consent form will be kept separate. Your research data will be stored in a secure location on CMU property or via secure electronic means and in the control of CMU. By participating, you understand and agree that the data and information gathered during this study may be used by Carnegie Mellon and published and/or disclosed by Carnegie Mellon to others outside of Carnegie Mellon. However, your name, address, contact information and other direct personal identifiers in your consent form will not be mentioned in any such publication or dissemination of the research data and/or results by Carnegie Mellon.

The researchers will take the following steps to protect participantsâĂŹ identities during this study: (1) Each participant will be assigned a number; (2) The researchers will record any data collected during the study by number, not by name; (3) Only members of the research group will view collected data in detail; (4) Any recordings or data files will be stored in a secured location accessed only by authorized researchers.

The sponsors listed on page 1 may also review identifiable research records.

Rights Your participation is voluntary. You are free to stop your participation at any point. Refusal to participate or withdrawal of your consent or discontinued participation in the study will not result in any penalty or loss of benefits or rights to which you might otherwise be entitled. The Principal Investigator may at his/her discretion remove you from the study for any of a number of reasons. In such an event, you will not suffer any penalty or loss of benefits or rights which you might otherwise be entitled.

Right to Ask Questions and Contact Information If you have any questions about this study, you should feel free to ask them now. If you have questions later, desire additional information, or wish to withdraw your participation please contact:

Dr. Reid Simmons Carnegie Mellon University Pittsburgh, PA, 15213 Email: rsimmons@andrew.cmu.edu (412) 268-2621

If you have questions pertaining to your rights as a research participant; or to report concerns to this study, you should contact the Office of Research Integrity and Compliance at Carnegie Mellon University. Email: irbâĂŘreview@andrew.cmu.edu . Phone: 412-268-4721.

Voluntary Consent By signing below, you agree that the above information has been explained to you and all your current questions have been answered. You understand that you may ask questions about any aspect of this research study during the course of the study and in the future. By writing your name below, you acknowledge that this is equivalent to a signature and you agree to participate in this research study.

Please write your name below as your signature

Optional Permission I understand that the researchers may want to use any video or audio recording for illustrative reasons in presentations of this work online, in print for scientific or educational purposes, or as part of a publicly available database. Please choose one of the following options.

- I grant full permission for the use of audio and video recordings as described above.
- I grant permission provided that my face is de-identified.
- I decline this optional permission.

A.1.2 Demographics Form

1. Your gender
2. Your age
3. How many children do you have?

4. How many grandchildren do you have?
5. Your ethnicity
6. What is your highest level of education?
7. What is your marital status?
8. How much knowledge do you have of robotics?
9. How much computer language programming experience do you have?
10. Have you ever interacted with a robot before?
11. If yes, what robot did you interact with? (name or type)

A.1.3 Survey completed after each round of exercise

Rate the robot on these scales

1. Stagnant to Lively (1-7)
2. Inert to Interactive (1-7)
3. Apathetic to Responsive (1-7)
4. Unfriendly to Friendly (1-7)
5. Unkind to Kind (1-7)
6. Unpleasant to Pleasant (1-7)
7. Incompetent to Competent (1-7)
8. Unintelligent to Intelligent (1-7)

Answer these questions about the robot coach

1. How strict did you feel the robot coach was? (1 not strict to 7 very strict)
2. How motivational did you feel the robot coach was? (1 not motivational at all to 7 very motivational)
3. How corrective did you feel the robot coach was? (1 not corrective to 7 very corrective)
4. How well did the robot coach's feedback match the kind of feedback you prefer? (1 not well at all to 7 extremely well)

5. Any additional comments?

A.2 LLM Prompt for Generating Robot Verbal Feedback

```
from openai import OpenAI
import pickle
import json

API_KEY = ''
client = OpenAI(api_key=API_KEY)

system_prompt = '''
You are a physical therapy assistant, and your client is a physical
                                  therapy patient. Your client is
                                  performing repetitions of var:
                                  exercise name and the last
                                  repetition var:evaluation. You see
                                   that your client var:fatigue.

You can give feedback in five different ways:

The very encouraging style has the most encouragement and is the
                                  most positive even when the client
                                   makes some mistakes
The encouraging style has a little less encouragement than very
                                  encouraging, but is still very
                                  positive
The neutral style is neither very encouraging nor very firm
The firm style has very small amount of encouragement and pushes the
                                   client to fix mistakes
The very firm style has even less encouragement and pushes the
                                  client even more to fix mistakes

For the repetition, you choose to give some feedback to them in a
                                  var:style style.
```

```
Make sure that you adjust the feedback that you provide to take into
                                       consideration the client's
                                      fatigue, how their their last
                                      repetition went, the type of
                                      exercise they are doing, and what
                                      style of feedback you are
                                      providing.

For example, here are a few examples of output.

Great effort! Keep going, youâĂŹre doing awesome!
Concentrate on the right shoulder. Work towards a better range.
'''

user_prompt = 'Provide a one sentence statement of feedback in a var
                                      :style style to the client, and
                                      your response should be less than
                                      10 words.'

exercise_list = ['bicep curls', 'lateral raises']

fatigue_list = ['low', 'moderate', 'high']
fatigue_message_list = ['is not very tired, and has a lot of energy'
                                      , 'is a little tired and has less
                                      energy.', 'is very tired and
                                      fatigued']

style_list = ['very firm', 'firm', 'neutral', 'encouraging', 'very
                                      encouraging']

m1 = ['had a lower range of motion than ideal on the left side, but
                                      the right side had good form', '
                                      had a lower range of motion than
                                      ideal on the right side, but the
                                      left side had good form', 'had a
                                      lower range of motion than ideal
                                      on both sides']
e1 = ['low_range left side', 'low_range right side', 'low_range both
                                       sides']
```

```
m2 = ['had bad form on the left side, but the right side had good
                                    form', 'had bad form on the right
                                    side, but the left side had good
                                    form', 'had bad form on both sides
                                    ']
e2 = ['bad left side', 'bad right side', 'bad both sides']

m3 = ['corrected a mistake in the previous repetition of lower range
                                     of motion on the left side', '
                                    corrected a mistake in the
                                    previous repetition of lower range
                                     of motion on the right side', '
                                    corrected a mistake in the
                                    previous repetition of lower range
                                     of motion on both sides']
e3 = ['corrected low_range left side', 'corrected low_range right
                                    side', 'corrected low_range both
                                    sides']

m4 = ['corrected a mistake in the previous repetition of bad form on
                                     the left side', 'corrected a
                                    mistake in the previous repetition
                                     of bad form on the right side', '
                                    corrected a mistake in the
                                    previous repetition of bad form on
                                     both sides']
e4 = ['corrected bad right side', 'corrected bad left side', '
                                    corrected bad both sides']

m5 = ['had good form']
e5 = ['good form']

m6 = ['had a higher range of motion than ideal on the left side', '
                                    had a higher range of motion than
                                    ideal on the right side.', 'had a
                                    higher range of motion than ideal
                                    on both sides']
```

```
e6 = ['high_range left side', 'high_range right side', 'high_range
                                    both sides']

m7 = ['corrected a mistake in the previous repetition of higher
                                    range of motion on the left side',
                                     'corrected a mistake in the
                                    previous repetition of higher
                                    range of motion on the right side'
                                    , 'corrected a mistake in the
                                    previous repetition of higher
                                    range of motion on both sides']
e7 = ['corrected high_range left side', 'corrected high_range right
                                    side', 'corrected high_range both
                                    sides']

m8 = ['was slower than it should be', 'was faster than it should be'
                                    ]
e8 = ['fast', 'slow']

m9 = ['corrected a mistake in the previous repetition of going too
                                    fast', 'corrected a mistake in the
                                     previous repetition of going too
                                    slow']
e9 = ['corrected fast', 'corrected slow']

m10 = ['had good speed']
e10 = ['good speed']

evaluation_message_list = m1 + m2 + m3 + m4 + m5 + m6 + m7 + m8 + m9
                                     + m10
evaluation_list = e1 + e2 + e3 + e4 + e5 + e6 + e7 + e8 + e9 + e10

num_prompts = 0
for exercise in exercise_list:
    for fatigue, fatigue_message in zip(fatigue_list,
                                        fatigue_message_list):

        response_dict = {}
        #New file for each fo these combinations
```

```
    for message, evaluation in zip(evaluation_message_list,
                                        evaluation_list):
        response_dict[evaluation] = {}
        for style in style_list:
            response_dict[evaluation][style] = []
            cur_system_prompt = system_prompt.replace('var:
                                            exercise',
                                            exercise)
            cur_system_prompt = cur_system_prompt.replace('var:
                                            evaluation',
                                            message)
            cur_system_prompt = cur_system_prompt.replace('var:
                                            fatigue',
                                            fatigue_message)
            cur_system_prompt = cur_system_prompt.replace('var:
                                            style', style)
            cur_user_prompt = user_prompt.replace('var:style',
                                            style)

            for ii in range(5):
                completion = client.chat.completions.create(
                model="gpt-3.5-turbo",
                messages=[
                    {"role": "system",
                    "content": cur_system_prompt},
                    {"role": "user", "content": cur_user_prompt}
                ]
                )

                response = completion.choices[0].message.content
                response_dict[evaluation][style].append(response
                                                )
                num_prompts += 1
                if num_prompts % 10 == 0:
                    print(num_prompts)
    print(exercise, fatigue)
    print(num_prompts)
    with open('{}_{}.json'.format(exercise, fatigue), 'w') as f:
        f.write(json.dumps(response_dict, indent=4))
```

Bibliography

[1] Henny Admoni, Thomas Weng, Bradley Hayes, and Brian Scassellati. Robot nonverbal behavior improves task performance in difficult collaborations. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 51–58. IEEE, 2016. ISBN 1-4673-8370-8. 2.3

[2] Mansi Agarwal and Jack Mostow. Semi-supervised Learning to Perceive Children's Affective States in a Tablet Tutor. In *Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI-20): Tenth Symposium on Educational Advances in Artificial Intelligence(EAAI)*, New York, NY, 2020. 4.2, 4.2.1

[3] Fady Alnajjar, Massimiliano Cappuccio, Abdulrahman Renawi, Omar Mubin, and Chu Kiong Loo. Personalized robot interventions for autistic children: an automated methodology for attention assessment. *International Journal of Social Robotics*, 13:67–82, 2021. 2.5

[4] Amir Aly and Adriana Tapus. Towards an intelligent system for generating an adapted verbal and nonverbal combined behavior in human–robot interaction. *Autonomous Robots*, 40(2):193–209, 2016. 2.5

[5] Brandon Amos, Bartosz Ludwiczuk, and Mahadev Satyanarayanan. Openface: A general-purpose face recognition library with mobile applications. *CMU School of Computer Science*, 6, 2016. 4.2.1

[6] Elisha Babad. TeachersâĂŹ nonverbal behavior and its effects on students. *The scholarship of teaching and learning in higher education: An evidence-based perspective*, pages 201–261, 2007. 2.1

[7] Elisha Babad. Teaching and nonverbal behavior in the classroom. In *International handbook of research on teachers and teaching*, pages 817–827. Springer, 2009. 2.1

[8] Christoph Bartneck, Dana Kulić, Elizabeth Croft, and Susana Zoghbi. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International journal of social robotics*, 1(1):71–81, 2009. 5.3.2, 6.4.1

[9] Aryel Beck, Lola Cañamero, and Kim A. Bard. Towards an affect space for

robots to display emotional body language. In *19th International Symposium in Robot and Human Interactive Communication*, pages 464–469. IEEE, 2010. ISBN 1-4244-7990-8. 2.2

[10] Tony Belpaeme, James Kennedy, Aditi Ramachandran, Brian Scassellati, and Fumihide Tanaka. Social robots for education: A review. *Science robotics*, 3(21), 2018. 2.2

[11] Veronika Bogina, Tsvi Kuflik, and Osnat Mokryn. Learning item temporal dynamics for predicting buying sessions. In *Proceedings of the 21st International Conference on Intelligent User Interfaces*, pages 251–255, 2016. 4.5

[12] Jean-David Boucher, Ugo Pattacini, Amelie Lelong, Gerard Bailly, Frederic Elisei, Sascha Fagel, Peter F. Dominey, and Jocelyne Ventre-Dominey. I reach faster when I see you look: Gaze effects in human–human and human–robot face-to-face cooperation. *Frontiers in neurorobotics*, 6:3, 2012. 2.3

[13] Djallel Bouneffouf, Irina Rish, and Charu Aggarwal. Survey on applications of multi-armed and contextual bandits. In *2020 IEEE Congress on Evolutionary Computation (CEC)*, pages 1–8. IEEE, 2020. ISBN 1-72816-929-1. 3.4

[14] Stéphanie Buisine, Matthieu Courgeon, Aurélien Charles, Céline Clavel, Jean-Claude Martin, Ning Tan, and Ouriel Grynszpan. The role of body postures in the recognition of emotions in contextually rich scenarios. *International Journal of Human-Computer Interaction*, 30(1):52–62, 2014. 2.5, 5.1, 5.1

[15] Judee K. Burgoon, David B. Buller, Jerold L. Hale, and Mark A. de Turck. Relational messages associated with nonverbal behaviors. *Human Communication Research*, 10(3):351–378, 1984. 4.2.1

[16] Paul C. Burnett. Elementary students' preferences for teacher praise. *The Journal of Classroom Interaction*, pages 16–23, 2001. 1, 2.1

[17] Mina Choi, Rachel Kornfield, Leila Takayama, and Bilge Mutlu. Movement matters: Effects of motion and mimicry on perception of similarity and closeness in robot-mediated communication. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 325–335, 2017. 2.4

[18] Ku-Chun Chou, Hsuan-Tien Lin, Chao-Kai Chiang, and Chi-Jen Lu. Pseudo-reward algorithms for contextual bandits with linear payoff functions. In *Asian Conference on Machine Learning*, pages 344–359. PMLR, 2015. 7.1

[19] Sheri R Colberg, David P Swain, and Aaron I Vinik. Use of heart rate reserve and rating of perceived exertion to prescribe exercise intensity in diabetic autonomic neuropathy. *Diabetes care*, 26(4):986–990, 2003. 7.2

[20] Mark Coulson. Attributing emotion to static body postures: Recognition accuracy, confusions, and viewpoint dependence. *Journal of nonverbal behavior*, 28

(2):117–139, 2004. 5.1, 5.1

[21] Nele Dael, Marcello Mortillaro, and Klaus R. Scherer. Emotion expression in body action and posture. *Emotion*, 12(5):1085, 2012. 2.1, 5.1, 5.1

[22] Tatiana Dobrescu. The role of non-verbal communication in the coach-athlete relationship. *Procedia-Social and Behavioral Sciences*, 149:286–291, 2014. 2.1

[23] Paul Ekman. Facial expression and emotion. *American psychologist*, 48(4):384, 1993. 2.1

[24] Paul Ekman and Wallace V. Friesen. *Facial Action Coding Systems*. Consulting Psychologists Press, 1978. 2.1, 4.2.1, 5.1

[25] Brittney A. English, Alexis Coates, and Ayanna Howard. Recognition of gestural behaviors expressed by humanoid robotic platforms for teaching affect recognition to children with autism-a healthy subjects pilot study. In *International Conference on Social Robotics*, pages 567–576. Springer, 2017. 2.2

[26] Mahtab Jahanbani Fard, Ping Wang, Sanjay Chawla, and Chandan K. Reddy. A bayesian perspective on early stage event prediction in longitudinal data. *IEEE Transactions on Knowledge and Data Engineering*, 28(12):3126–3139, 2016. 4.1

[27] Juan Fasola and Maja J. Matarić. A socially assistive robot exercise coach for the elderly. *Journal of Human-Robot Interaction*, 2(2):3–32, 2013. 6.4

[28] Hassan Ismail Fawaz, Germain Forestier, Jonathan Weber, Lhassane Idoumghar, and Pierre-Alain Muller. Deep learning for time series classification: A review. *Data Mining and Knowledge Discovery*, 33(4):917–963, 2019. 4.1, 4.5

[29] Sophie Fyfe, Claire Williams, Oliver J. Mason, and Graham J. Pickup. Apophenia, theory of mind and schizotypy: Perceiving meaning and intentionality in randomness. *Cortex*, 44(10):1316–1325, 2008. 5.2

[30] Goren Gordon, Samuel Spaulding, Jacqueline Kory Westlund, Jin Joo Lee, Luke Plummer, Marayna Martinez, Madhurima Das, and Cynthia Breazeal. Affective personalization of a social robot tutor for children's second language skills. In *Thirtieth AAAI Conference on Artificial Intelligence*, 2016. 2.5

[31] Binnur Görer, Albert Ali Salah, and H. Levent Akın. An autonomous robotic exercise tutor for elderly people. *Autonomous Robots*, 41(3):657–678, March 2017. ISSN 1573-7527. doi: 10.1007/s10514-016-9598-5. 6.4

[32] Gerda Hagenauer, Tina Hascher, and Simone E. Volet. Teacher emotions in the classroom: Associations with students' engagement, classroom discipline and the interpersonal teacher-student relationship. *European journal of psychology of education*, 30(4):385–403, 2015. 2.1

[33] Madison Heimerdinger and Amy LaViers. Influence of environmental context on recognition rates of stylized walking sequences. In *International Conference on*

Social Robotics, pages 272–282. Springer, 2017. 2.5

[34] Noora Hirvonen, Heidi Enwald, Peter A. Bath, Riitta Pyky, Raija Korpelainen, and Maija-Leena Huotari. Individual Factors Affecting Preferences for Feedback Message Tactics in the Contexts of Physical Activity. *Journal of Health Communication*, 20(2):220–229, February 2015. ISSN 1081-0730, 1087-0415. doi: 10.1080/10810730.2014.925015. 2.1

[35] Bahar Irfan, Nathalia Céspedes, Jonathan Casas, Emmanuel Senft, Luisa F. Gutiérrez, Mónica Rincon-Roncancio, Carlos A. Cifuentes, Tony Belpaeme, and Marcela Múnera. Personalised socially assistive robot for cardiac rehabilitation: Critical reflections on long-term interactions in the real world. *User Modeling and User-Adapted Interaction*, pages 1–48, 2022. 2.5

[36] Melissa L. Kamins and Carol S. Dweck. Person versus process praise and criticism: Implications for contingent self-worth and coping. *Developmental psychology*, 35 (3):835, 1999. 2.1

[37] Roshni Kaushik and Amy LaViers. Imitating Human Movement Using a Measure of Verticality to Animate Low Degree-of-Freedom Non-humanoid Virtual Characters. In *Social Robotics*, volume 11357, pages 588–598, Cham, 2018. Springer International Publishing. ISBN 978-3-030-05203-4 978-3-030-05204-1. doi: 10.1007/978-3-030-05204-1_58. 2.2

[38] Roshni Kaushik and Amy LaViers. Robots Imitating Humans: Human Perception and Preference of Low Degree-of-Freedom Simulated Robot Motion Driven by Various Human Capture Mappings. *International Journal of Social Robotics*, 2019. 2.2

[39] Roshni Kaushik and Reid Simmons. Early Prediction of Student Engagement-Related Events from Facial and Contextual Features. In *International Conference on Social Robotics*, pages 308–318. Springer, 2021. 4

[40] Roshni Kaushik and Reid Simmons. Perception of Emotion in Torso and Arm Movements on Humanoid Robot Quori. In *Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*, pages 62–66, 2021. 5

[41] Roshni Kaushik and Reid Simmons. Affective Robot Behavior Improves Learning in a Sorting Game. In *2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, pages 436–441. IEEE, 2022. ISBN 1-72818-859-8. 5

[42] Roshni Kaushik and Reid Simmons. Effects of Feedback Styles on Performance and Preference for an Exercise Coach. In *2024 33st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, 2024. 6

[43] Roshni Kaushik and Reid Simmons. Older Adults' Preferences for Feedback

Cadence from an Exercise Coach Robot. In *2024 33st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, 2024. 6

[44] Roshni Kaushik and Reid Simmons. Choosing robot feedback style to optimize human exercise performance. In *2025 20th edition of the IEEE/ACM International Conference on Human-Robot Interaction*, 2025. 7

[45] Spencer D. Kelly, Dale J. Barr, R. Breckinridge Church, and Katheryn Lynch. Offering a hand to pragmatic understanding: The role of speech and gesture in comprehension and memory. *Journal of memory and Language*, 40(4):577–592, 1999. 1, 2.1

[46] James Kennedy, Paul Baxter, Emmanuel Senft, and Tony Belpaeme. Higher nonverbal immediacy leads to greater learning gains in child-robot tutoring interactions. In *International Conference on Social Robotics*, pages 327–336. Springer, 2015. 2.3

[47] Mark L. Knapp, Judith A. Hall, and Terrence G. Horgan. *Nonverbal Communication in Human Interaction*. Cengage Learning, 2013. ISBN 1-285-49984-0. 2.5

[48] Heather Knight and Reid Simmons. Expressive motion with x, y and theta: Laban effort features for mobile robots. In *The 23rd IEEE International Symposium on Robot and Human Interactive Communication*, pages 267–273. IEEE, 2014. ISBN 1-4799-6765-3. 2.2

[49] Heather Knight and Reid Simmons. Laban head-motions convey robot state: A call for robot body language. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pages 2881–2888. IEEE, 2016. ISBN 1-4673-8026-1. 2.2

[50] Nathan Koenig and Andrew Howard. Design and use paradigms for gazebo, an open-source multi-robot simulator. In *2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)(IEEE Cat. No. 04CH37566)*, volume 3, pages 2149–2154. IEEE, 2004. ISBN 0-7803-8463-6. 5.3.1

[51] Rudolf Laban and Lisa Ullmann. The mastery of movement. 1971. 2.2

[52] Evgeny Levinkov and Mario Fritz. Sequential bayesian model update under structured scene prior for semantic road scenes labeling. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 1321–1328, 2013. 4.3.2

[53] Daniel Leyzberg, Samuel Spaulding, and Brian Scassellati. Personalizing robot tutors to individuals' learning differences. In *2014 9th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 423–430. IEEE, 2014. ISBN 1-4503-2658-7. 2.5

[54] Manja Lohse, Reinier Rothuis, Jorge Gallego-Pérez, Daphne E. Karreman, and Vanessa Evers. Robot gestures make difficult tasks easier: The impact of gestures on perceived workload and task performance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1459–1466, 2014. 2.3

[55] Ahmad Lotfi, Caroline Langensiepen, and Salisu Wada Yahaya. Socially assistive robotics: Robot exercise trainer for older adults. *Technologies*, 6(1):32, 2018. 6.4

[56] O. Ivar Lovaas. Interaction between verbal and nonverbal behavior. *Child Development*, pages 329–336, 1961. ISBN: 0009-3920 Publisher: JSTOR. 6.3.2

[57] Maxim Makatchev and Reid Simmons. Perception of personality and naturalness through dialogues by native speakers of American English and Arabic. In *Proceedings of the SIGDIAL 2011 Conference*, pages 286–293. Association for Computational Linguistics, 2011. ISBN 1-937284-10-7. 2.5

[58] Takahiko Masuda, Phoebe C. Ellsworth, Batja Mesquita, Janxin Leu, Shigehito Tanida, and Ellen Van de Veerdonk. Placing the face in context: Cultural differences in the perception of facial emotion. *Journal of personality and social psychology*, 94(3):365, 2008. 2.5

[59] Maja J. Matarić, Jon Eriksson, David J. Feil-Seifer, and Carolee J. Winstein. Socially assistive robotics for post-stroke rehabilitation. *Journal of neuroengineering and rehabilitation*, 4(1):1–9, 2007. 6.4

[60] Derek McColl and Goldie Nejat. Recognizing emotional body language displayed by a human-like social robot. *International Journal of Social Robotics*, 6(2): 261–280, 2014. 2.2

[61] Andrew A. McReynolds, Sheba P. Naderzad, Mononito Goswami, and Jack Mostow. Toward Learning at Scale in Developing Countries: Lessons from the Global Learning XPRIZE Field Study. In *Learning @ Scale*. ACM, August 2020. ISBN 978-1-4503-7951-9. doi: 10.1145/3386527.3405920. 4, 4.1, 4.2, 4.5

[62] Albert Mehrabian. *Silent Messages*, volume 8. Wadsworth Belmont, CA, 1971. ISBN 0-534-00059-2. 2.1

[63] Masahiro Mori. The uncanny valley: The original essay by Masahiro Mori. *IEEE Robots &*, 2017. 6.2.2

[64] Usue Mori, Alexander Mendiburu, Eamonn Keogh, and Jose A. Lozano. Reliable early classification of time series based on discriminating the classes over time. *Data mining and knowledge discovery*, 31(1):233–263, 2017. 4.1

[65] Lilia Moshkina. Improving request compliance through robot affect. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 26, 2012. ISBN 2374-3468. 2.3

[66] TIMOTHY P. Mottet and VIRGINIA P. Richmond. Student nonverbal communication and its influence on teachers and teaching. *Communication for teachers*, pages 47–61, 2002. 1, 2.1

[67] Tomasz Niznikowski, Jerzy Sadowski, and Andrzej Mastalerz. The effectiveness of different types of verbal feedback on learning complex movement tasks. *Human movement*, 14(2):148–153, 2013. ISBN: 1899-1955 Publisher: Termedia. 2.1

[68] Sunghyun Park, Lilia Moshkina, and Ronald C. Arkina. Recognizing nonverbal affective behavior in humanoid robots. In *Intelligent Autonomous Systems*, volume 11, page 12, 2010. 2.2

[69] Ung Park, Minsoo Kim, Youngeun Jang, GiJae Lee, KangGeon Kim, Ig-Jae Kim, and Jongsuk Choi. Robot facial expression framework for enhancing empathy in human-robot interaction. In *2021 30th IEEE International Conference on Robot & Human Interactive Communication (RO-MAN)*, pages 832–838. IEEE, 2021. ISBN 1-66540-492-2. 2.5

[70] André Pereira, Catharine Oertel, Leonor Fermoselle, Joseph Mendelson, and Joakim Gustafson. Effects of different interaction contexts when evaluating gaze models in HRI. In *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, pages 131–139, 2020. 2.5

[71] Rifca Peters, Joost Broekens, and Mark A. Neerincx. Robots educate in style: The effect of context and non-verbal behaviour on children's perceptions of warmth and competence. In *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pages 449–455. IEEE, 2017. ISBN 1-5386-3518-6. 2.4, 2.5

[72] Maria Reichelt, Frauke Kämmerer, Helmut M Niegemann, and Steffi Zander. Talk to me personally: Personalization of language style in computer-based learning. *Computers in Human behavior*, 35:199–210, 2014. 2.5

[73] Virginia P. Richmond, James C. McCroskey, and Mark Hickson. *Nonverbal Behavior in Interpersonal Relations.* Allyn & Bacon, 2008. ISBN 0-205-48669-X. 2.1

[74] Astrid M. Rosenthal-von der Pütten, Nicole C. Krämer, and Jonathan Herrmann. The effects of humanlike and robot-specific affective nonverbal behavior on perception, emotion, and behavior. *International Journal of Social Robotics*, 10 (5):569–582, 2018. 2.2, 2.4

[75] Maha Salem, Katharina Rohlfing, Stefan Kopp, and Frank Joublin. A friendly gesture: Investigating the effect of multimodal robot behavior in human-robot interaction. In *2011 Ro-Man*, pages 247–252. IEEE, 2011. ISBN 1-4577-1573-2. 2.4

[76] Nicole Salomons, Tom Wallenstein, Debasmita Ghose, and Brian Scassellati.

The Impact of an In-Home Co-Located Robotic Coach in Helping People Make Fewer Exercise Mistakes. In *31st IEEE International Conference on Robot & Human Interactive Communication*, 2022. 2.3

[77] Stan Salvador and Philip Chan. Toward accurate dynamic time warping in linear time and space. *Intelligent Data Analysis*, 11(5):561–580, 2007. 6.1.2

[78] Jyotirmay Sanghvi, Ginevra Castellano, Iolanda Leite, André Pereira, Peter W. McOwan, and Ana Paiva. Automatic analysis of affective postures and body motion to detect engagement with a game companion. In *Proceedings of the 6th International Conference on Human-robot Interaction*, pages 305–312, 2011. 4.4.2

[79] Shane Saunderson and Goldie Nejat. How robots influence humans: A survey of nonverbal communication in social human–robot interaction. *International Journal of Social Robotics*, 11(4):575–608, 2019. 1, 2.2

[80] Mayank Saxena, Rohith Krishnan Pillai, and Jack Mostow. Relating Children's Automatically Detected Facial Expressions to Their Behavior in RoboTutor. In *Thirty-Second AAAI Conference on Artificial Intelligence*, 2018. 4.2

[81] Sofia Serholt and Wolmet Barendregt. Robots tutoring children: Longitudinal evaluation of social engagement in child-robot interaction. In *Proceedings of the 9th Nordic Conference on Human-Computer Interaction*, page 64. ACM, 2016. ISBN 1-4503-4763-0. 2.4

[82] Megha Sharma, Dale Hildebrandt, Gem Newman, James E. Young, and Rasit Eskicioglu. Communicating affect via flight path: Exploring use of the laban effort system for designing affective locomotion paths. In *Proceedings of the 8th ACM/IEEE International Conference on Human-robot Interaction*, pages 293–300. IEEE Press, 2013. ISBN 1-4673-3055-8. 2.2

[83] Mohammad Shayganfar, Charles Rich, and Candace L. Sidner. A design methodology for expressing emotion on robot faces. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 4577–4583. IEEE, 2012. ISBN 1-4673-1736-5. 5.1

[84] Qiming Shen, Kerstin Dautenhahn, Joe Saunders, and Hatice Kose. Can real-time, adaptive human–robot motor coordination improve humans' overall perception of a robot? *IEEE Transactions on Autonomous Mental Development*, 7(1):52–64, 2015. 2.5

[85] Howard A. Smith. Nonverbal communication in teaching. *Review of Educational Research*, 49(4):631–672, 1979. 2.1

[86] Samuel Spaulding, Goren Gordon, and Cynthia Breazeal. Affect-aware student models for robot tutors. In *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems*, pages 864–872, 2016. 4.4.2

[87] Andrew Specian, Nick Eckenstein, Mark Yim, Ross Mead, Braden McDorman, Simon Kim, and Maja Matarić. Preliminary system and hardware design for Quori, a low-cost, modular, socially interactive robot. In *Workshop on Social Robots in the Wild*, 2018. 5.1, 5.3.1

[88] Ioanna-Ourania Stathopoulou and George A. Tsihrintzis. Emotion recognition from body movements and gestures. In *Intelligent Interactive Multimedia Systems and Services*, pages 295–303. Springer, 2011. 5.1

[89] Adriana Tapus and Maja J. Matarić. User personality matching with a hands-off robot for post-stroke rehabilitation therapy. In *Experimental Robotics*, pages 165–175. Springer, 2008. 2.5

[90] Elisabeth T. Van Dijk, Elena Torta, and Raymond H. Cuijpers. Effects of eye contact and iconic gestures on message retention in human-robot interaction. *International Journal of Social Robotics*, 5(4):491–501, 2013. 2.3

[91] Michael L. Walters, Kerstin Dautenhahn, René Te Boekhorst, Kheng Lee Koay, Christina Kaouri, Sarah Woods, Chrystopher Nehaniv, David Lee, and Iain Werry. The influence of subjects' personality traits on personal spatial zones in a human-robot interaction experiment. In *ROMAN 2005. IEEE International Workshop on Robot and Human Interactive Communication, 2005.*, pages 347–352. IEEE, 2005. ISBN 0-7803-9274-4. 2.5

[92] Anita E. Woolfolk and Douglas M. Brooks. Chapter 5: Nonverbal communication in teaching. *Review of research in education*, 10(1):103–149, 1983. 2.1, 2.5

[93] Anita E Woolfolk and Douglas M Brooks. The influence of teachers' nonverbal behaviors on students' perceptions and performance. *The Elementary School Journal*, 85(4):513–528, 1985. 2.1

[94] Zhengzheng Xing, Jian Pei, and S. Yu Philip. Early classification on time series. *Knowledge and information systems*, 31(1):105–127, 2012. 4.1

[95] Junchao Xu, Joost Broekens, Koen Hindriks, and Mark A. Neerincx. *Mood Expression through Parameterized Functional Behavior of Robots*. IEEE, 2013. ISBN 1-4799-0509-7. 2.2

[96] Junchao Xu, Joost Broekens, Koen Hindriks, and Mark A. Neerincx. Effects of bodily mood expression of a robotic teacher on students. In *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2614–2620. IEEE, 2014. ISBN 1-4799-6934-6. 2.4