Learning to Partner: Exploring Real-Time Adaptive Feedback via Temporal-Difference Machine Learning for Improved Human-Prosthesis Collaboration

by

Adam S. R. Parker

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Rehabilitation Science

Faculty of Rehabilitation Medicine

University of Alberta

© Adam S. R. Parker, 2024

Abstract

Modern myoelectric artificial limbs are sophisticated devices with many of the degrees of freedom of biological limbs. These devices have great potential to provide function for people with amputations, assisting them in participating in a greater number of activities and tasks of daily living. While tremendous advancements have been made in the control of myoelectric prostheses, the interface between user and device only allows users to scratch the surface of the capability of modern prosthetic devices. Importantly to the interface, feedback of any kind is not widely commercially available from prosthetic devices.

A potential path towards improving user interactions with prosthetic limbs in the current age of artificial intelligence is to view the device not only as a tool being used but as a partner assisting the user in their daily life. That is the goal of this work: to *apply real-time machine learning methods to wearable assistive robotics to promote collaborative partnerships with users.* Most prior work using machine learning in prostheses has been focused on how users control the device. Research has been increasing in providing feedback from devices to users with a focus on communicating sensation, but has not yet begun to explore how to use machine learning to provide and curate feedback. Here the focus is on the application of machine learning to the feedback pathway—signals from the device to the user. This is an important step to enabling bi-directional communication between agents in order to achieve strong collaborative interactions.

First, this dissertation champions viewing a direct-to-body device such as

an upper-limb prosthesis as a partner. It outlines the value of viewing the interaction as a partnership and introduces a framework for the value and evaluation of strong partnerships between a human and a machine. Following that, a set of experiments demonstrate the ability of real-time machine feedback learning methods to learn something of value to a human user in a prosthetic domain for the first time. These experiments also show that machine-learned feedback can be successfully acquired and adapted in real time during human-robot interaction. The final two experiments explore the human side of interactions with a device that is adapting the feedback it provides over time. Evidence in this dissertation suggests that signals coming from devices that adapt as a user interacts with them could be the key to encouraging the user to engage more deeply with their device and initiate positive long-term interactions. Along with these findings, the methods described in this dissertation for understanding the human side of human-prosthesis interaction are important contributions of this work.

Overall this dissertation demonstrates for the first time the reasons and benefits of viewing upper-limb prostheses, and indeed many other assistive technologies, as partners to users rather than resigning them to simply being tools. Temporal-difference learning methods are shown to be capable of learning and adapting feedback signals being sent to users in real-time and are a strong path towards closing the loop between humans and machines to enable collaborations. Combined with methodological contributions that use rich data from both machines and humans, this dissertation proposes a shift in the way we think about human-device interaction in rehabilitation. This shift towards thinking about and creating collaborations between humans and assistive technology, especially in upper-limb prostheses, holds the potential for unlocking greater functional gains wherever humans and devices come together.

Preface

Ethics

This thesis is an original work of Adam S. R. Parker. The various studies done over the course of the Ph.D were conducted in accordance with appropriate ethical oversight.

Chapter 4, "Exploring the Impact of Machine-Learned Predictions on Feedback from an Artificial Limb", was conducted in accordance with the University of Alberta Research Ethics Board, Pro00017008, approved in 2012.

Chapter 5, "Continually Learned Pavlovian Signalling Without Forgetting for Human-in-the-Loop Robotic Control", was conducted in accordance with the University of Alberta Research Ethics Board 2, Pro00085727, approved in March 2019.

Chapter 6, "Assessing Human Interaction in Virtual Reality With Continually Learning Prediction Agents Based on Reinforcement Learning Algorithms: A Pilot Study", was done following a protocol approved with informed consent via DeepMind Technologies Ltd.'s arms-length IRB for human participant studies.

Chapter 7, "Understanding Human Interaction with Real-Time Adaptive Feedback During Simulated Prosthesis Use", was conducted in accordance with the University of Alberta Research Ethics Board 2, Pro00123026, approved in November 2022.

Authorship

This thesis consists primarily of six research manuscripts which have been or are planned to be submitted for publication.

Chapter 2 is available as:

> Parker, A. S. R., & Pilarski, P. M. (2021). Position statement: Assistive technology as partners through machine-learned communication. *Workshop on Reinforcement Learning for Humans, Computer, and Interaction (RL4HCI), ACM CHI 2021*, 1–3. https://sites.google.com/view/rl4hci/position-papers
>
> As first author I conceptualized and wrote this work with the conceptualization, editing, formatting, and figure support from the supervising author, Patrick M. Pilarski.

Chapter 3 has been published as:

> Mathewson, K. W., Parker, A. S. R., Sherstan, C., Edwards, A. L., Sutton, R. S., & Pilarski, P. M. (2023). Communicative capital: A key resource for human–machine shared agency and collaborative capacity. *Neural Computing and Applications*, *35*(23), 16805–16819. https://doi.org/10.1007/s00521-022-07948-1
>
> As second author and lead student of record on the final copy of this work, I contributed heavily to the conceptualization and editing of this paper. More specifically I outlined the conceptualization shown in the figures, which is a key contribution of the paper. I also wrote the section on seeing-eye dogs and contributed heavily to editing the work to properly develop and outline the concepts.

Chapter 4 has been published as:

> Parker, A. S. R., Edwards, A. L., & Pilarski, P. M. (2019a). Exploring the impact of machine-learned predictions on feedback from an artificial limb. *2019 16th IEEE/RAS-EMBS International Conference on Rehabilitation Robotics (ICORR)*, 1239–1246. https://doi.org/10.1109/ICORR.2019.8779424
>
> As first author, I contributed the experimental design, the physical experimental setup, the data analysis, and the core of the writing.

The physical experimental setup here consisted of the workspace, getting the robot to move, and providing visual and vibrotactile feedback, as well as the machine-learning code used. The data was collected by Ann L. Edwards. As supervising author Patrick M. Pilarski assisted with conceptualization and editing, most notably with the machine learning portion of the methods, as well as the final figures.

Chapter 5 is available as:

Parker, A. S. R., Dawson, M. R., & Pilarski, P. M. (2022). Continually learned Pavlovian signalling without forgetting for human-in-the-loop robotic control. *NeurIPS Workshop on Human in the Loop Learning (HiLL). also arXiv:2305.14365 [cs.LG]*, 1–12. https://doi.org/10.48550/arXiv.2305.14365

As first author, I was responsible for the experimental design, implementation, execution, analysis, and the core of the writing. Implementation here included constructing the workspace the experiment was conducted in and modifying the software used to control the robot to take in signals from the workspace as well as autonomously react to those signals. The machine-learning code was adapted from Ann L. Edward's previous work on adaptive switching. Here I modified this code to use different state information, and a different learning target, and to implement the look-ahead version. Micheal R. Dawson previously developed the software and robot that was used and modified for this study. As supervising author Patrick M. Pilarski assisted with conceptualization, editing, as well as the final figures.

Chapter 6 has been published as:

Brenneis, D. J. A., Parker, A. S. R., Johanson, M. B., Butcher, A., Davoodi, E., Acker, L., Botvinick, M. M., Modayil, J., White, A., & Pilarski, P. M. (2022). Assessing human interaction in virtual reality with continually learning prediction agents based on reinforcement learning algorithms: A pilot study. *Adaptive and Learning Agents (ALA) Workshop at AAMAS 2022*, 1–8. https://doi.org/10.48550/arXiv.2112.07774

I was the only student author of this work. I contributed to the thought and design of the experiment as well as the entire qualitative contribution. This included the qualitative design, data collection technique, analysis, and write-up.

Chapter 7 is written for submission to Nature Human Behaviour.

Parker, A. S. R., Williams, H. E., Phelan, S. K., Hebert, J. S., Shehata, A. W., & Pilarski, P. M. (2024). *Understanding human interaction with real-time adaptive feedback during simulated prosthesis use* [In preparation]

As first author of this work, I am largely responsible for the design, execution and writing of this study. The power hand and prosthesis simulation brace were previously developed in the lab by Eric D. Wells and Ben W. Hallworth. The software used to interface with the robot was developed by Micheal R. Dawson. The machine-learning code was adapted from Ann L. Edward's previous work on adaptive switching. Here I modified this code to use selective Kanerva coding, by using a modification of code from Jaden B. Travnik, for the state space. I also changed the learning target and configured the code to generate the audible feedback. The quantitative design was advised by Ahmed W. Shehata and Jacqueline S. Hebert. The mixed-methods and qualitative design were supported by Shanon K. Phelan. Phelan also assisted in question development. Heather E. Williams assisted in data collection by assisting in the development of the checklist, as well as operating the GaMA system. Williams also wrote the portions of the paper that discuss how GaMA works, as well as provided the statistical analysis. As supervising author Patrick M. Pilarski assisted with study conceptualization, and editing.

Additional Works

There are two works that do not contain sufficient writing on my part to be included directly as chapters, but I played a vital role in their creation. The ideas of these works will feature heavily in the introduction and conclusion chapters.

> Dawson, M. R., Parker, A. S. R., Williams, H. E., Shehata, A. W., Hebert, J. S., Chapman, C. S., & Pilarski, P. M. (2024). Joint action is a framework for understanding partnerships between humans and upper limb prostheses. *International Conference on Biomedical Robotics and Biomechatronics (BioRob); also arXiv:2212.14124 [cs.HC]*. https://doi.org/10.48550/arXiv.2212.14124

The concept of joint action was introduced to me at a satellite symposium. I brought the idea back to the lab and discussed with Dr. Pilarski how it might be valuable to our work with human-prosthesis interaction. These discussions were crucial in the development of the above paper.

> Schofield, J. S., Battraw, M. A., Parker, A. S. R., Pilarski, P. M., Sensinger, J. W., & Marasco, P. D. (2021). Embodied cooperation to promote forgiving interactions with autonomous machines. *Frontiers in Neurorobotics*, *15*(661603). https://doi.org/10.3389/fnbot.2021.661603

I was the core collaborator from the BLINC lab on this work. Working closely with Marcus Battraw, the first student author on the paper, I was instrumental in collecting and developing the ideas from the brainstorming session with the full authorship team, as well as brining my on philosophical insight and understanding of human-machine interaction and relationships.

Beyond these there are several other works I was a part of that fit the contributions of this work but would not be appropriate to include directly as chapters. Most of the work listed here is closely related to existing chapters. The exception being "Learned human-agent decision-making, communication and joint action in a virtual reality environment", to which my influence and ideas were deemed sufficiently crucial to include me on the author list.

Parker, A. S. R., Edwards, A. L., & Pilarski, P. M. (2019b). Machine-learned predictions assisting human control of an artificial limb [Abstract and Poster]. *4th Multidisciplinary Conference on Reinforcement Learning and Decision Making (RLDM)*

Parker, A. S. R., Edwards, A. L., & Pilarski, P. M. (2018). Exploring communication as actions in human-machine partnerships [Abstract and Poster]. *12th Annual Canadian Neuroscience meeting Satellite 1: CAPnet/CPS*

Pilarski, P. M., Butcher, A., Johanson, M., Botvinick, M. M., Bolt, A., & Parker, A. S. R. (2019). Learned human-agent decision-making, communication and joint action in a virtual reality environment. *The 4th Multidisciplinary Conference on Reinforcement Learning and Decision Making; also arXiv:1905.02691 [cs.AI]*. https://doi.org/10.48550/arXiv.1905.02691

Pilarski, P. M., Butcher, A., Davoodi, E., Johanson, M. B., Brenneis, D. J. A., Parker, A. S. R., Acker, L., Botvinick, M. M., Modayil, J., & White, A. (2022). The Frost Hollow experiments: Pavlovian signalling as a path to coordination and communication between agents. *arXiv:2203.09498 [cs.AI]*. https://doi.org/10.48550/arXiv.2203.09498

Butcher, A., Johanson, M. B., Davoodi, E., Brenneis, D. J. A., Acker, L., Parker, A. S. R., White, A., Modayil, J., & Pilarski, P. M. (2022). Pavlovian signalling with general value functions in agent-agent temporal decision making. *Adaptive and Learning Agents (ALA) Workshop at AAMAS 2022*. https://doi.org/10.48550/arXiv.2201.03709

To Victoria and Griffon

For giving me a reason to keep going.

Acknowledgements

The author wishes to thank and recognize the people who supported him in this, long, endeavour. Specifically my supervisor Patrick M. Pilarski, and to Micheal R. Dawson and Ahmed W. Shehata for their support and guidance throughout. With special mention to Patrick M. Pilarksi for welcoming me into his research so long ago, and guiding me for so long. Also Richard S. Sutton and Joseph Modayil for early guidance in how I think about science generally. I also wish to acknowledge my advisory committee, Shanon K. Phelan, Kim Adams, and Mahdi Tavakoli, for their wonderful insights, for guiding my thinking down new pathways, and for their patience with my wild ideas. In addition my amazing externals Jim T. Enns and James R. Wright Shanon K. Phelan deserves special mention for opening my eyes to qualitative research and continuing to be my guide on chaotic, messy, important ground. A big thank you to everyone from the BLINC lab at the University of Alberta as well. For all the laughs, celebrations, commiserations, and conversations throughout the years.

Funding for this work was granted by the Natural Sciences and Engineering Research Council of Canada (NSERC), Alberta Innovates, the Sensory Motor Adaptive Rehabilitation Technology (SMART) Network, the Alberta Machine Intelligence Institute (Amii), and the Canada CIFAR AI Chairs program. Their support made this work possible.

Contents

List of Tables

List of Figures

"You need to know where to go," Sanya said.
"Yes."
"And you are going to consult four large pizzas for guidance."
"Yes," I said.
The big man frowned for a moment. Then he said, "There is, I think, humor here which does not translate well from English into sanity."

– Jim Butcher, Changes

Chapter 1

Introduction

1.1 Learning Devices as Partners in Rehabilitation

A core goal of rehabilitation interventions is for the patient to gain function (World Health Organization, 2007). The purpose of this function is often to enable people to participate in activities of daily living such as brushing their teeth, folding laundry, cooking, or even getting around. Such activities are easy to take for granted, but make a huge difference on people's quality of life. According to contemporary clinical thinking, it is also important that the function added, as well as how it is added, is done using an individual, patient-centred approach (Castellini et al., 2014). It is the role of assistive technologies to change the functioning of users in activities that involve the device. Ideally, these devices provide function across a wide range of tasks, and in a way that is easy and comfortable for the participant (Arthanat et al., 2007). Examples of assistive technologies are wheelchairs, seeing-eye dogs, and prosthetic limbs (Kalinowska et al., 2023).

As one example, wheelchairs increase mobility for people with many different conditions. When a user receives a wheelchair there is a period of adjustment, and the wheelchair itself can be adapted to better suit the user's needs (Papadimitriou, 2008; Winance, 2006). This is sometimes described as a process of extending their sense of self to include the wheelchair, despite there being little anatomical similarity between human locomotion and the action of the wheelchair (Papadimitriou, 2008). If the person struggles to get around because of a sight issue one option is a seeing-eye dog. Seeing-eye dogs are independent agents that are trained to act in certain ways. They take responsibility for actions, including disobedience when required, and users are trained in how to use, interpret, and understand the dogs. Communication and interaction between the user and the dog is taught to both parties and adapted in a supervised way (Fishman, 2003; Pfaffenberger, 1976). So, the tool (the wheelchair) is adapted to best assist, and the agent (the seeing-eye dog) is, at least initially, coupled to the human in a predetermined way but uses its agency and decision-making to assist. What this implies is that a fully independent agent with well-understood communication and interaction pathways can be a valuable assistant to a user, and tools can be adapted to best fit a user.

Another assistive technology is prosthetic, or artificial, limbs. Such devices are intended to assist users in many different tasks of daily living. Of particular interest in this work are upper limb prostheses. Modern robotic upper limb prostheses have a vast amount of capability to execute biologically accurate motion and in some cases the capacity to provide feedback. Despite research into improving control and feedback for users of prosthetic limbs, *we are still not close to achieving the full potential of modern devices* (Biddiss & Chau, 2007b; Cordella et al., 2016; Yamamoto et al., 2019). Taking the perspective of adding function, and drawing on inspiration from wheelchairs and seeing-eye dogs, another path appears—to work towards collaborative partnerships between humans and devices as we explore in the chapters that follow. One possible way to enable such partnerships could be through adding agency to the device, such as in the case of the seeing-eye dog. This would allow the device to adapt itself to user needs, similarly to how with the help of experts users adapt wheelchairs to their own needs (Winance, 2006). With learning methods such as those proposed in this dissertation, this could be the gateway to getting more out of the devices for users while ensuring user-centred care.

Adding agency to artificial limbs has been explored in the context of prosthetic limbs previously (Edwards, Dawson, et al., 2016; Pilarski, Dick, et al., 2013). These studies utilize machine learning techniques to learn about an interaction with a user in real-time, and the system takes some action to assist a user based on this learning. When discussing training times for dexterous hands, Castellini et al. (2014) called for improved adaptive methods to reduce learning times for users (Castellini et al., 2014). It is not a significant extension to suggest that this learning continues during use in real-time.

The machine-learning technique of choice for most of this work is temporal-difference (TD) learning techniques from the field of reinforcement learning (RL) (Sutton, 1988; Sutton & Barto, 2018). Such techniques are well suited to rehabilitation applications and the addition of agency to devices for several reasons (Pilarski, Dawson, Degris, Carey, et al., 2013). Most importantly to rehabilitation, such methods can learn in real-time from interactions with the user and the environment (Pilarski, Dawson, Degris, Carey, et al., 2013). These methods learn predictions about the target of learning that are associated with those interactions. This means devices enabled with TD methods should be able to learn from a specific user and adapt the assistance they provide in a user-centered way. While these techniques are from the field of reinforcement learning, TD methods can be used with signals from the environment as the target of prediction learning rather than reward (Sutton et al., 2011; White, 2015)

Building on this capability to make models of partners and predict interactions with them, we can frame interactions between users and prostheses using joint action (Dawson et al., 2024). the study of joint action includes the study of perception and action in a human, social, context. Humans coordinate with each other regularly using cues, both verbal and non-verbal, and predictive models of the partner they are working with (Brennan & Enns, 2015; Sebanz et al., 2006; Sebanz & Knoblich, 2009). Temporal-difference learning methods, which learn predictions about the world from a real-time sensor stream, are therefore interestingly positioned to take advantage of the pathways that already exist in human users to coordinate actions. Brennan and Enns (2015) has an especially interesting finding that users who worked on a task with a friend performed better than the combination of each part (Brennan & Enns, 2015). It stands to reason, then, that close partnerships between human and machine (prosthetic) agents could lead to better outcomes.

Schofield et al. (2021) further posit that the pursuit of embodiment of prosthetic limbs does not have to be separate from collaborations between humans and devices (Schofield et al., 2021). There are biological systems in our bodies that act autonomously to allow complex coordination of muscles to achieve grip. Despite this, when it comes to devices and mechanisms external to ourselves we can be quick to blame the "other" (Jackson, 2002). The pursuit of knowledge to improve collaboration between user and device, then, could be crucial to truly connecting people with their devices.

Enabling bi-directional communication between user and device may be a pivotal step to achieving collaborative partnerships between users and devices, especially in human-robot interaction such as exists with prostheses (Kalinowska et al., 2023). While it is likely that a user can achieve an understanding of a device, depending on the device's complexity, without receiving feedback directly from the device, such feedback has inherent advantages (Kalinowska et al., 2023). For example, if something unexpected or new happens, having well-established bi-directional communication between user and device would be of great value. The goal of designing systems that adapt the feedback they provide in real-time, then, is to tune the communication to be the most effective for an individual user, that partnership, in that time. This level of individualization would be challenging to achieve on a person-by-person basis with a fixed, non-learned, form of feedback.

Therefore this dissertation explores how to apply temporal-difference methods to learn in real-time to provide signals to users to assist and promote collaborations between humans and machines. The signals used in these explorations are first vibration, then audible feedback. Prosthetic limbs present a unique challenge to this goal. While they are intended for a wide range of tasks, there's often a significant gap between a device's capabilities and the functions users actually experience. Prosthetic limbs also have the added challenge of being compared to previously existing functions that were fully natural and intuitive to the user. Progress in achieving collaborative relationships between the user and the device in this field should give tremendous insight that could be used across assistive technologies used in rehabilitation to improve outcomes for users.

With this in mind, one of the key contributions in the latter chapters of this dissertation is a mixed-methods approach to study the interaction between a human user and a machine-learning agent that is adapting signals being sent to the user in real time. Mixed methods are of core importance to the rehabilitation philosophy of patient-centred care as they bring the patient's interests, needs, and experiences to the forefront. A key component of a qualitative study, or the qualitative portion of a mixed methods study, is the methodology that is used to frame the work (Ponterotto, 2005). A common methodology used with prosthesis users is phenomenological analysis. Such methodologies seek to understand a core, shared, "qualia" or "experiences of being" that relate to living with limb difference and using prosthetic limbs (Murray, 2004; Wilding & Whiteford, 2005) Since the concern in this dissertation is to understand the interaction, and perhaps the unfolding relationship between the user and device, a different approach is taken here. The methodology that these studies in the latter chapters of this dissertation were approached with varied slightly between them, but could be categorized as pragmatism with constructivist leanings. What this means is that truth is seen as "constructed" by individuals, and even co-constructed between agents when they interact with each other (Schwandt, 1994). That said, when conducting studies involving humans, we can consider two sets of observations. One set comes from a shared, external space where we can measure and record numerical data. The other data set comes from an individual's interpretation of this shared space and the reality and truth they construct from that interpretation (Dewey, 1908; Goldkuhl, 2012; Kaushik & Walsh, 2019).

At the outset of the research of this dissertation, it was thought that discourse analysis would be used for the qualitative portions of these studies. Discourse analysis would provide insight into how the interactions between the person and the device were discussed. This is done by first examining the context in which the comments are made. This context is composed of the the social or cultural influences surrounded the participant and activity Ballinger and Payne, 2000. While this approach could provide valuable insight, studying and understanding the context of the participant and activity fell outside of the scope of this work. As such, descriptive thematic analysis modified from Braun and Clarke's reflexive thematic analysis was used instead Braun and Clarke, 2021a, 2021b.

Understanding these two sets of observations—objective data and subjective experience—is crucial when studying human interaction with assistive technologies. These devices that are intended to provide function for users and assist them throughout their day with activities of daily living. Some of the devices are tools, some have agency and are trained to work with users to assist them. Machine learning, specifically real-time machine learning methods can be used to add agency to tools, allowing them to adapt to users to provide even more specific care for individual users. Bringing agency to artificial limbs is an additional pathway to unlock the full potential of modern robotic prostheses, and has great capacity to keep up with future developments as they occur. This would be accomplished by developing methods of creating human-machine partnerships in these devices. Collaborations between agents to provide increased function to users is not unheard of in rehabilitation, and the addition of agency through machine learning has been explored even in the prosthetic domain (Edwards, Dawson, et al., 2016; Fishman, 2003; Pilarski, Dawson, Degris, Carey, et al., 2013). Strong collaborations between users and devices, prosthetic limbs or otherwise, utilizing innate human mechanisms for joint action, are a promising interaction method to unlock the full potential of assistive technologies. Therefore, this dissertation seeks to take crucial steps toward improving patient interactions with assistive technologies by demonstrating how we can apply real-time machine learning to encourage collaborative interactions.

1.2 Research Contributions and Aims

At the onset of this research, feedback in upper limb prostheses was of growing interest. There was little work using machine learning to adapt feedback in real time, and even less relating to prosthetic use. There was also some suggestion of viewing an upper limb prosthesis and its users from a multi-agent standpoint, but this was at the time considered a fringe idea. My dissertation therefore set out to accomplish three main things. The first was to make a case for upper limb prostheses as partners, with the goal of the partnership being to provide function to users. The second was to discover if simple learning methods could, in real-time, learn something that is of value to the vastly more powerful thinking agent, the human, that is using it. The third was to discover how users' experience interacting with devices is different when the device is sending signals that are adapting in real-time as opposed to the device sending a fixed signal. Therefore, this work sought to, for the first time to our knowledge, mix rich quantitative (recordings of robot data, motion, and gaze) with rich qualitative (guided journalling and semi-structured interviews) data. This provides a fuller picture of a user's interactions with a device that is learning and adapting the feedback the user receives, all in real time. By understanding how real-time machine learning can be used by people who are using assistive robotics, we can learn how to facilitate strong human-machine collaboration and unlock greater potential from our rehabilitation technologies.

1.3 Thesis Structure

This thesis, in keeping with the University of Alberta paper-based thesis format, is primarily constructed of works that are published, or intended to be published in the near future. As such, background information for each chapter is contained in each chapter.

Beginning in Chapter 2, an argument is established for the view of upper limb prostheses as partners. Following this in Chapter 3 some examples of how existing work could be viewed in this way are discussed and a framework for thinking about how we can add function to users through partnerships with devices is established. Chapter 3 also proposes how we might characterize the benefits of human-agent partnerships. Next, the question of the value of machine-learned feedback to a stronger agent, namely a human user, is examined in Chapter 4. This chapter uses the temporal-difference learning methods that will be commonly used throughout this work. The following chapter, Chapter 5, examines a shortfall of the previous study that is of great importance to the future of machine learning in prosthetic limbs in general—that is, having the device adapt in real-time. The final two major chapters both bring the human user into the examination more closely. Chapter 6 sees a user collaborating with several different machine agents in virtual reality. The final chapter, Chapter 7, is the culmination of the years of thought and research behind the previous work of this dissertation. It presents a detailed study of humans using a wearable robot hand and interacting with feedback that is adapting in real time. Finally, the overall conclusions of this dissertation as a whole are presented along with future directions in the final concluding chapter. The flow between these elements can be seen in Fig. 1.1

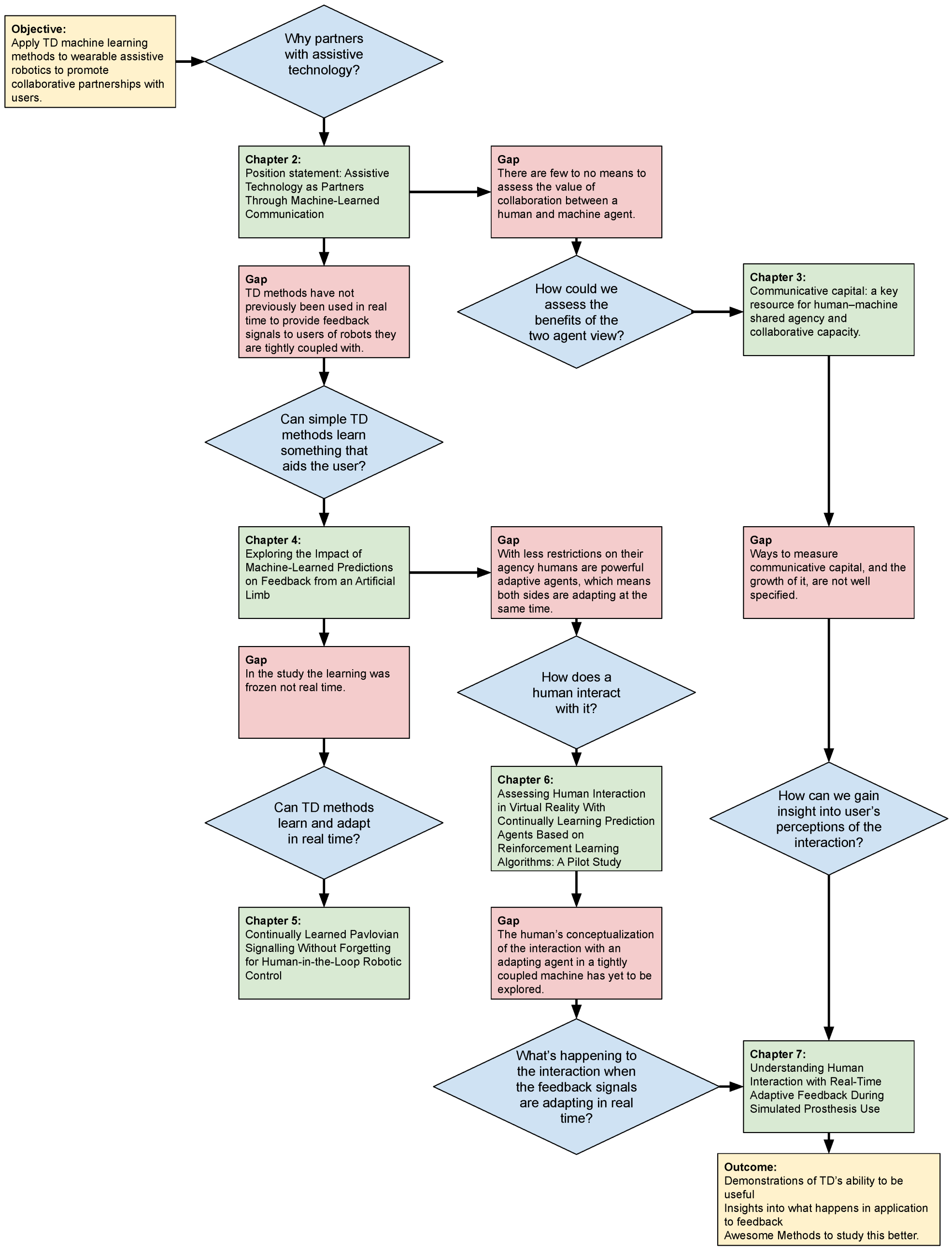


Figure 1.1: The flow and connection between the chapters that compose the bulk of this thesis.

Chapter 2

Position Statement: Assistive Technology as Partners Through Machine-Learned Communication

As machine-learning is increasingly applied to technologies we interact with, devices become agents. These agents can learn, adapt, make decisions and take actions. When our devices can do these things, it is natural to begin to look at human-machine interaction in a different way than how it has been historically viewed. Specifically, we can, and should, begin to examine human-machine interaction in the same was as we do human-human interaction, as a *collaboration between agents*. Artificial limbs provide an especially interesting domain to study and apply this viewpoint as they represent an intimate pairing of human and machine, and artificial limbs are traditionally considered a replacement rather than a partnership. As a primary contribution of this chapter, we briefly outline the perspective that artificial limbs can and should be viewed as collaborative agents, and provide arguments in favour of using approaches from the field of reinforcement learning to achieve tightly-coupled collaborations between human and machine. We further outline a set of recommended first steps for how we might make progress towards achieving human-machine collaborations by adopting the viewpoint of collaboration as a real-time, continual learning interaction on the part of both a person and a machine.

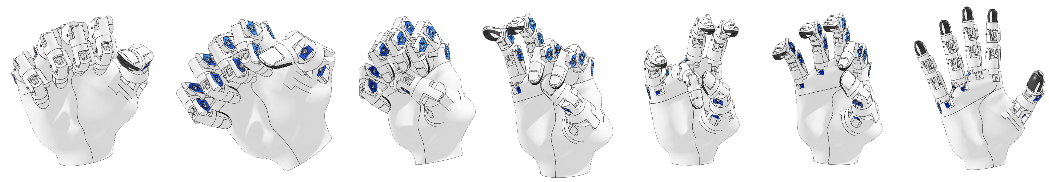


Figure 2.1: Interactions between humans and their world are now enabled by tightly coupled partnerships with increasingly complex machine agents, as perhaps best viewed in the domain of robotic artificial limbs.

2.1 Machine Partners

Our daily interactions with devices are becoming increasingly numerous, and the devices themselves are becoming more computationally powerful and adaptive to perceived human need. As these capabilities continue to increase and approach the point of resembling animal or even human-level intelligence we can begin to examine human-machine interaction in different ways than it has been traditionally viewed. In order to improve how people use such devices, we suggest that one natural approach would be to leverage innate human social instincts for cooperation. Further, a core position put forth by the present work is that an ideal area to study human-machine interaction is artificial limbs (Figs. 2.1–2.2). Prosthetic devices are an ideal setting for studying human-machine cooperation as they represent a tightly coupled, immediate, real-time interaction between a human and a device (Castellini et al., 2014). The interaction is not abstract, like using your smartphone to search the internet or manage your calendar, but a direct and persistent sensorimotor interaction. The human and the device exist in the same space, they move and affect changes on the world in the same location, and they are directly impacted by the actions taken by the other entity. Further, in terms of social relevance, prosthetic limbs currently face many challenges with how they are operated as well as with adoption rates which may be, at least in part, related to how the challenges of the interaction are typically approached as replacing a highly integrated part of a biological entity with mechanics (Castellini et al., 2014; Peerdeman et al., 2011).

Artificial limbs fall within the setting of rehabilitation medicine. In rehabilitation there is increasing emphasis being placed on enabling patients to accomplish tasks, rather than restoring something that is missing or getting the patient to some predefined “normal” function (World Health Organization, 2007). With that in mind, we propose viewing assistive technology, especially the human-computer interaction effected through artificial limbs, as *adding* function. As artificial limbs now begin to incorporate machine learning in their regular operation (c.f., Castellini 2014), we submit that their added function should be in fact viewed as providing the user with an assistant—one that ideally learns and improves in an ongoing way through a process of interaction and feedback (c.f., Fig. 2.2).

When real-time machine learning is involved, the interaction between a user and their prosthesis is now similar to providing a person who is visually impaired with a cane, or a seeing-eye dog. There are several features of the person-agent relationship in the seeing-eye dog case that are of particular interest. One feature is that the seeing-eye dog is expected to take responsibility for actions and situations where it has different information than the human. Similarly, the dog is expected to refuse commands that may endanger the owner. Interestingly, both partners, the dog and the human, need to be trained in this partnership through a process of reward and feedback. The dog is trained in commands as well as general mannerisms and expectations, and the human needs to be taught how to interact and communicate with the dog, in part through the success or failure of their intended daily-life tasks (Pfaffenberger, 1976) (pp. 85, 88).

The human-human interactions studied in the field of joint action also provides insight into a collaborative, learning interaction between users and devices (Pesquita et al., 2018; Sebanz & Knoblich, 2009). One study suggested that participants doing a joint task performed better than the sum of their independent parts when they had a pre-existing relationship; when the participants were friends they performed better than pairings of strangers. This effect was lost when the partners were visually obscured from each other (Brennan & Enns, 2015). Joint action research also outlines a possible framework for human collaboration on tasks, of which an important aspect is having and maintaining a number of predictions of the other agent, their actions, and their responses to the first agents actions.

With this context in mind, we now return to the core posit of the present work: that we can best improve artificial limbs by making them cooperative agents, assistants to humans that can learn how to interact and communicate on a personal level through reinforcement and feedback. More generally, through the recommendations that follow, we aim to encourage exploration into how machine agents can learn to use natural human tendencies for co-operation to make interactions with machines better across a broad range of domains. If successful, users might be expected to achieve a sense of embodiment or "oneness" with the device as a skilled rider does with their favoured horse. This pursuit may also provide insight and techniques to further studies of communication and co-operation in human-human, human-machine, and machine-machine agent partnerships. If we can satisfactorily claim that a human-machine interaction is acting as a collaboration between agents, this opens up new possibilities for research in other collaborative pairings.

2.2 Recommendations: Collaboration Through Learned Communication

To ground our suggested line of investigation in human-machine interaction, we turn to the reinforcement learning (RL) literature and specifically methods of learning by way of a temporal-difference (TD) error (Sutton, 1988). The RL problem formulation and related algorithms are ideally suited to our proposed line of study into tightly coupled human-prosthesis interaction for two main reasons. First, RL approaches are well suited to learning to predict online and in real-time, such as during an ongoing interaction between a machine and its environment (Sutton, 1988). This is well in line with the predictive joint-action model of human cooperation (Pesquita et al., 2018). Second, RL methods can adapt their predictions about a diverse set of signals in real-time from the sensory-motor data stream during deployment (e.g., via the TD learning of generalized value functions, GVFs; (Pilarski, Dawson, De-

Figure 2.2: An individual with an amputation interacting with their environment in collaboration with an electromechanical prosthesis—in this case the Bento Arm (Dawson et al., 2014) and brachI/Oplexus software (Dawson et al., 2020), a robotic limb suite integrating reinforcement learning algorithms and classical prosthetic control solutions.

gris, Carey, et al., 2013; Sutton et al., 2011)). This is vital for tightly coupled cases of human-machine interaction. Previous work by our group and others with RL techniques applied to artificial limbs and other neuroprostheses have indicated the natural fit to the human-machine setting, with areas of improvement including control, feedback, and lessening the cognitive load (Castellini & Van Der Smagt, 2009; Dalrymple et al., 2020; Edwards, Dawson, et al., 2016; Pilarski, Dawson, Degris, Carey, et al., 2013; Pilarski, Dick, et al., 2013).

Recommendation 1: As a first recommendation to our community for next steps in studying the tightly coupled interaction between a human and a prosthesis, we suggest detailed investigation is required to determine in what ways real-time machine learning can be applied to prosthetic limbs so that a limbs is in fact able learn something of lasting value to the human user. There are three aspects of a tightly-coupled sensorimotor interaction wherein an agent might learn something valuable: the task, the machine itself, and the human (Pilarski, Dawson, Degris, Carey, et al., 2013). Preliminary work by

A. S. R. Parker et al. (2019a) on learning about the task, user, and machine showed that a machine could learn forecasts about how a participant uses a robot arm to navigate a physical environment which is obscured from them in some way. In this prior work, GVF predictions learned via TD methods were used as feedback signals to a user to help them complete a task they would have difficulty completing on their own or through conventional techniques. This work would benefit from expansion, exploring more algorithms, algorithm features, and types of feedback as well as combinations thereof.

Recommendation 2: After establishing that machine-learning can learn something that is of value to a human user in a direct-action sensory-motor sense, we recommend that a natural next step is to pursue deeper understanding as to a machine's ability to learn to communicate with its user. This communication will ideally be individual to a single user (Castellini et al., 2014; Pilarski, Dick, et al., 2013). We suspect this kind of personalized communication is the key, or at the very least the first step, to establishing strong partnerships between human and machine agents. Previous studies have shown interesting results with machine agents learning to communicate with each other and successfully collaborate to accomplish a task (Cao et al., 2018; Lazaridou et al., 2016). We suggest learning to communicate with a human should first be done with a defined lexicon, and then explore something akin to body language; having the device learn how to use its motors and the features of its control algorithm in order to communicate with the human user.

Recommendation 3: We suggest that a final key component to establishing strong partnerships between humans and closely-connected devices is pairing the quantitative data collection with qualitative studies. Our preference is to use discourse analysis (interviews). This is to explore the relationship that is developing between the human and the machine, and how differing machine-learning approaches can impact the relationship. Quantitative metrics of success alone will not do here; if the user is performing better but strongly dislikes the device then we have not established a favourable relationship that is likely to continue. Qualitative results will be used to inform algorithm design and implementation in subsequent studies. Iterating between qualitative and quantitative findings in this way should bring us closer to ideal interactions between humans and their devices.

2.3 Impact and Future Potential

We believe the outcomes impact of our recommended line of research extends beyond artificial limbs into any partnership between a human and a machine that unfolds in real time through ongoing interaction. Further, we believe the RL problem setting is an ideal lens by which to study tightly coupled interactions of all kinds. Lessons learned on the highly specific and intimate pairing of human and artificial limb will be relevant to any interaction between a human and a device that is capable of the required levels of computation for continual learning. There may even be a natural decrease in the complexity required of the machine agent when the interaction with a human is less intimate, which further suggests prosthetic limbs as a sound place to begin such studies. The potential exists for research on RL as applied to tightly coupled human machine interaction to revolutionize how humans and machines interact across a broad range of fields, creating strong partnerships that dramatically increase human ability in collaboration with increasingly powerful computing devices.

Chapter 3

Communicative Capital: A Key Resource for Human-Machine Shared Agency and Collaborative Capacity

In this work, we present a perspective on the role machine intelligence can play in supporting human abilities. In particular, we consider research in rehabilitation technologies such as prosthetic devices, as this domain requires tight coupling between human and machine. Taking an agent-based view of such devices, we propose that human-machine collaborations have a capacity to perform tasks which is a result of the combined agency of the human and the machine. We introduce *communicative capital* as a resource developed by a human and a machine working together in ongoing interactions. Development of this resource enables the partnership to eventually perform tasks at a capacity greater than either individual could achieve alone. We then examine the benefits and challenges of increasing the agency of prostheses by surveying literature which demonstrates that building communicative resources enables more complex, task-directed interactions. The viewpoint developed in this chapter extends current thinking on how best to support the functional use of increasingly complex prostheses, and establishes insight toward creating more fruitful interactions between humans and supportive, assistive, and augmentative technologies.

3.1 Introduction

Technology can be used to amplify natural human abilities, provide access to new abilities, and supplement abilities changed due to injury or illness (Belfiore, 2010; Brooks, 2002; Dewdney, 1998; Doidge, 2007; Geary, 2002; Moss, 2011). Various tools and technological interventions are well known to support humans in physically interacting with their world, improving perceptual abilities, and supporting decision-making and memory (Clark, 2008; Geary, 2002; Osiurak & Badets, 2016; Risko & Gilbert, 2016). Interventions to provide people with the functions they require for daily life are a core area of interest in rehabilitation, as outlined by the International Classification of Functioning, Disability and Health (ICF) (Jette, 2006; World Health Organization, 2007). For example, Geary (2002) describes ways that technology is used to enhance sight, touch, hearing, taste, smell, and mental processes. Millán et al. (2010), Castellini et al. (2014), and Carmena (2012) further present views on the use of technology to supplement and enhance motor and sensory abilities for people who have lost body parts or body functions. Of interest to this work are technological advances in assistive or augmentative technology involving tight coupling (Licklider, 1960) between a person and a machine with the capacity to learn. This coupling affects the ability of the combined human-machine partnership to have, seek, and achieve goals.

We present the perspective that a human's ability to have, seek, and achieve goals can be supported using machine intelligence, specifically by combining human ability with reinforcement learning agents (Sutton & Barto, 2018). We term this *human-machine shared agency*. This perspective suggests that a human and their machine counterpart should be viewed as partners attempting to accomplish a shared task, where the agency of each partner combines to allow for greater potential capacity to accomplish tasks.

As a main contribution, we introduce **communicative capital: a resource that is built up over time in a human-machine partnership that allows the partners to eventually perform tasks at a capacity greater than either individual could achieve alone**. The resource can consist of accumulated propositional or procedural knowledge, conventions, beliefs, models, and predictions of the other agent. Communicative capital is represented within each agent and is stored within the individual memory of both agents. Communicative capital directly affects the behavioural collaborative capacity of the human-machine partnership.

In this chapter, we specifically consider the case where the resource is in the form of predictions learned over time from interaction between human and prosthetic devices. While our setting of interest is human-machine interaction, a helpful motivating example is a human-guide dog partnership that allows both independent agents—human and canine—to accomplish a greater range and complexity of shared tasks (discussed in Section 3.6.1).

3.2 Robotic Upper-Limb Prostheses

Robotic prostheses and other examples from the field of rehabilitation technology help us focus our thinking on direct human-machine interactions that can be well supported by machine intelligence. The rehabilitation technology setting is appealing in that it involves a direct, immediate, tightly coupled collaboration between a human and their technology to achieve a goal (Licklider, 1960; A. S. R. Parker & Pilarski, 2021). Examples of assistive rehabilitation devices include semi-autonomous wheelchairs (Millán et al., 2010; Viswanathan et al., 2014), robotic manipulators and locomotors (Castellini et al., 2014; Ortiz-Catalan et al., 2020), exoskeletons (Herr, 2009), smart living environments (Rashidi & Mihailidis, 2013), and socially assistive robotic coaches (Feil-Seifer & Matarić, 2011). The representative example of assistive rehabilitation technology we focus on in the present work is *robotic upper-limb prostheses*: assistive electromechanical devices attached to the body of individuals with amputations (Pilarski & Hebert, 2017) (Fig. 3.1). Despite the evolution of prosthetic devices from iron hands to more dexterous mechanical manipulators, and improvements in quality of life for some users, state-of-the-art devices have yet to create a satisfactory solution for many individuals (Castellini et al., 2014; Peerdeman et al., 2011; Williams III, 2011; Ziegler-

(a) (b) (c)

Figure 3.1: Prostheses examples: (a) robotic upper-limb prosthetic, (b) human using a research prosthesis (Dawson et al., 2014), (c) human using a supernumerary limb (A. S. R. Parker et al., 2019a).

Graham et al., 2008; Zuo & Olson, 2014).

In the prosthetic setting, movement control contributions from both human and machine must combine effectively in order for the device to benefit the human user. In this setting challenges result from the limited number of degrees of human control and the lack of feedback from the device (Castellini et al., 2014; Schofield et al., 2014). The coupling of human and device is further complicated by the dynamic, non-stationary nature of human environments (Saridis & Stephanou, 1977). This coupling has been improved by muscular, neural, and osseointegration allowing for a more direct, high-bandwidth connection between human and machine (Castellini et al., 2014; Hochberg et al., 2006; Ortiz-Catalan et al., 2020; Ortiz-Catalán et al., 2014; Zuo & Olson, 2014). To provide a bidirectional flow of information between prostheses and their users, cameras have been used to augment perception (Marković et al., 2014), microphones and speakers have been used to facilitate natural language interactions (Kollar et al., 2010), and both surgical practices and prosthetic feedback approaches have evolved (Hebert et al., 2014; Schofield et al., 2014). Prosthetic devices of the future will receive an unprecedented density of data about human users and their environment, and they should be well equipped to translate such data into actions which support the goals of the users.

Despite the potential of advanced prostheses to support human abilities, current neuroprosthetic literature describes that one remaining limitation on the interaction between human and machine is the number of independent signals flowing between human and machine partners (Castellini et al., 2014). This constrains control strategy design of upper-limb prostheses to a small number of degrees of freedom, actuated by classification or regression algorithms for real-time control. Giving the upper-limb prostheses some autonomy in their control mechanism has been shown to allow for simultaneous control of multiple degrees of freedom while still using the same number of independent human generated control signals (Castellini et al., 2014). For example, pattern recognition-based controllers have provided an improvement over conventional controllers in standardized tasks in randomized clinical trials in part because of their ability to learn to interpret and act upon diverse collections of signals provided by a human user (Hargrove et al., 2017; Vu et al., 2020). Importantly, these systems therefore require upfront investment on the part of both the device and the user in the form of initial training and subsequent adjustments in order to see the autonomy-related improvements they offer. Increasing the autonomy of a prosthetic device has been shown in many specific cases to significantly increase the capacity of the human-prosthesis partnership to efficiently and effectively accomplish functional tasks (Castellini et al., 2014). Perhaps surprisingly then, given the diverse data streams and automation capabilities noted above, the specific consequences of prostheses themselves being considered to have and share in agency during human prosthesis interaction has remained relatively under-explored. We now examine the relationship between agency and capabilities in human-prosthesis partnerships.

3.3 Prostheses as Agents

In this section, we consider the implications of treating a prosthetic device as an *agent*—an autonomous goal-seeking system. This is not a common perspective—it suggests both sides of a tightly coupled human-machine interface should be thought of as agents with goals. Drawing insight from relationships found in human-human joint action and interaction (Knoblich et al., 2011; Pesquita et al., 2018; Pezzulo et al., 2013; Sebanz et al., 2006), treating a human-prosthesis interaction in this way is in fact not as unfamiliar as it might first seem; with an agent-centric view, each agent would be expected, within its capability, to grow to understand the capabilities of the other and predict how to act accordingly. That is, each agent would naturally and, to the best of its ability, explicitly model the agency of the other to increase the capacity of the partnership in a continual and incrementally increasing fashion. This form of model building and adaptation is present in rather constrained ways in existing state-of-the-art upper-limb prostheses, and something the community hopes to enhance within future prosthetic systems (Castellini et al., 2014).

We first delineate degrees of agency and the resulting capabilities that each side of the prosthetic human-machine partnership may obtain. Here, the human and the machine are considered analogous to co-actors in a joint action task (Knoblich et al., 2011; Pesquita et al., 2018; Pezzulo et al., 2013; Sebanz et al., 2006) or the leader and follower in a two-agent partnership (Candidi et al., 2015); this collective shared agency is cooperation between a natural and an artificial system (Misselhorn, 2015). We define agency as the degree to which an autonomous system has the ability to have, seek, and achieve goals. This definition is inspired by the Belmont Report (Brady & Jonsen, 2014), wherein a system assumes agency if it is "capable of deliberation about personal goals and of acting under the direction of such deliberation". Hallmarks of agency include the ability to take actions, have sensation, persist over time, and improve with respect to a goal. These hallmarks give rise to an agent's ability to predict, control, and model its environment and other agents. By taking prior perspectives on agency into consideration (Tosic & Agha, 2004), along with the nuances of the prosthetic setting of interest, we focus on five attributes of agency that may be present in the human or machine agent.

1. Be a mechanism: The agent acts in a predetermined way in response to stimulus. For example, a myoelectric controller that processes electromyographic (EMG) signals via a fixed linear proportional mapping to create control commands for prosthetic actuators (P. A. Parker et al., 2006).

2. Adapt over time: In addition to being a mechanism, the agent has the capacity to adapt in response to the signals perceived. Through adaptation, the agent may acquire knowledge about its situation (e.g. by modelling and adapting to perceived signals). Adaptation can occur during training, as in the supervised learning of a pattern recognition classifier, or during ongoing experience (Castellini et al., 2014; Pilarski, Dawson, Degris, Carey, et al., 2013).

3. Pursue a goal: The agent has defined goals and an intent to optimize some measure of its own situation. One example of the pursuit of a goal is the maximization of a scalar reward signal, as in computational and biological reinforcement learning (Sutton & Barto, 2018).

4. Model the other agent as adapting: The agent views the other agent as adapting during ongoing interaction. This can alter the way one agent presents signals to the other. For example, a human user trains a pattern recognizing prosthetic with knowledge that the device is adapting to their signals.

5. Model the other agent as pursuing a goal: The agent views the other agent as not only changing in response to received signals, but also as pursuing its own objectives. This preliminary theory of mind further alters the way that the one agent presents signals to the other agent.

We present this list of attributes with the caveat that it is likely not exhaustive. We can imagine that there may be higher order attributes of agency which mirror the recursive theory of mind. Additional attributes may parallel high order intentionality and reasoning, as in research in animal ethology, machine theory of mind, and cultural intelligence (Cultural General Intelligence Team et al., 2022; Heyes, 1998; Rabinowitz et al., 2018; Zhu et al., 2021). This line of thinking is discussed further in Section 3.5.3.

We now outline a schema (Fig. 3.2) for considering degrees of agency and relate agency to the combined capacity of a human-machine partnership. Capacity and agency in this schema are agnostic to the units of measurement and the exact attributes of agency, so as to be compatible with, and still helpful across, multiple definitions of agency.

Capacity is a measure of task performance accomplished by the human-machine partnership as quantified by some metric. *Maximum capacity* is the

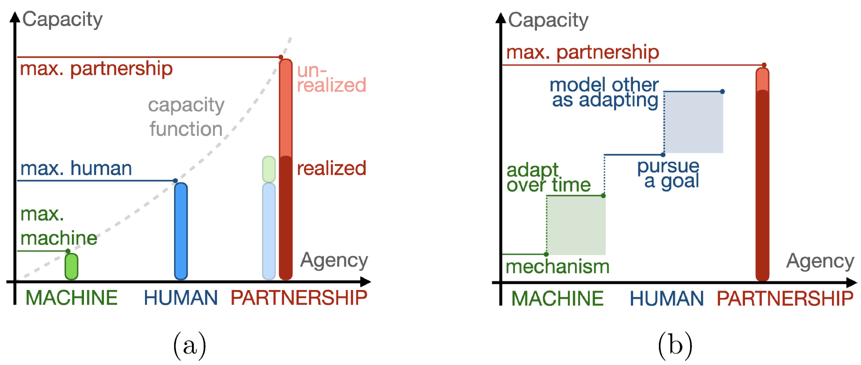


Figure 3.2: The capacity function (dashed grey line) is the relationship between *capacity* and *agency*. (a) The capacity of the partnership (red) is a function of the contributions from the machine agent (green) and the human agent (blue). (b) Illustrative example of how attributes of human and machine agency can relate to maximum partnership capacity. The light green shaded rectangle represents the capacity increase when a machine agent adapts over time versus when it acts only as a mechanism. The light blue shaded rectangle represents the increase in capacity when a human pursues a goal versus when it also models a machine partner as adapting.

optimal performance that could be achieved by the partnership, illustrated and labelled 'max partnership capacity' in Fig. 3.2(a). This maximum capacity can be realized or unrealized. *Realized capacity* is the actual achieved capacity of the partnership, shown as a solid red bar in Fig. 3.2(a).

Agency: Agency is the summation of contributions from individual degrees of agency, either discrete or continuous in nature. Multiple degrees combine to increase agency of the agent and shared agency of the partnership (Fig. 3.2).

Capacity function: Agency is related to capacity by a *capacity function*. By finding the point on a capacity function corresponding to a given level of agency, we can visualize the maximum capacity of partnership. A system that is a mechanism has less agency and less capacity than a system that is a mechanism, adapts over time, and pursues a goal. A partnership may result in greater capacity than the sum of the two individual systems if both partners model each other and how to effectively utilize the capabilities of both agents. A partnership can also result in a capacity less than the sum of the two individual systems if, for example, the partners interfere with each other.

As an illustrative example, Fig. 3.2 uses this agency-capacity schema to compare a human-mechanism partnership (without shaded rectangles) to a partnership where the machine is able to adapt (with shaded rectangles). Note how the maximum capacity of the partnership is greater than either could achieve on their own. That capacity may be initially unrealized and change over time, or it might only be realized if both agents can model the other as pursuing a goal.

The way that the goals of the human and machine align is a problem related to team formation in human-human and human-animal partnerships (Fishman, 2003). Such alignment can occur during normal sensorimotor interactions between agents (Pezzulo & Dindo, 2011; Pezzulo et al., 2013; Sebanz et al., 2006; Sebanz & Knoblich, 2009). To examine the process by which such alignment might occur during human-machine interaction, we now introduce the idea of communicative capital. Communicative capital is a resource built up through ongoing interactions between a human and their machine counterpart that correspond to how well both agents understand each other and the partnership (Pilarski et al., 2015).

3.4 Communicative Capital

As depicted in Figs. 3.2 and 3.3, the agency of the human and the machine contribute to the capacity of the partnership. *Communicative capital* is a resource built through interaction between both sides of the partnership. It enables a partnership to eventually perform a task at a capacity greater than either individual could achieve alone. Accumulating communicative capital requires investment to establish and maintain (see the 'cost of signalling' described by Pezzulo and Dindo (Pezzulo et al., 2013)). The cost of investing in communicative capital may be incurred passively during the interactions of a partnership, or, in many cases, through dedicated effort tangentially related to the ultimate goals of the partnership. For example, users of prosthetic devices learn about the use of their prosthesis before they take it home for use in activities of daily living. In advanced devices that use pattern recognition, teaching both sides of a partnership to engage in a system of meaning-by-convention (Santoro et al., 2021) (e.g., a series of commands to a prosthesis phrased in terms of patterns of myoelectric signals) may require significant additional time and energy but lead to increased future efficiency.

Building communicative capital can also be viewed as a process of compression and decompression, or via the lens of Scott-Phillips et al. (Scott-Phillips, 2014; Scott-Phillips et al., 2009), one related to ostension and inference. One agent takes an action and thereby encodes information into a signal. The other agent must decode the signal as it arrives, and thereby recover the associated information. To begin to form communicative capital, at least one of the two agents must be able to adapt. Further, we expect the greatest opportunities to build communicative capital will exist when both the human and the machine exhibit the highest possible degrees of agency. We now discuss how communicative capital can be built and used to progressively realize more capacity in prosthetic human-machine partnerships.

3.5 Building Capital through Interaction

So far we have considered settings where a communication channel exists between the human and the machine. While this channel can be either unidirectional or bidirectional, two-way communication is often beneficial for interactions between multiple goal-seeking agents. If the agent's goals are not furthered by the information received, then it may ignore the received information. If one agent's goals are not furthered by what the other agent does with received information, it will choose to not send such information in the future. The agent can send many possible things, and can therefore choose how to balance the cost of sending information with the expected outcomes for itself and the partnership (Pezzulo et al., 2013). It follows that both agents should vary their communication to send information that results in both improving with respect to their goals. The variation of communication could be independent, or guided by other parties—e.g., the work of clinical staff to train a patient for prosthesis use, or an instructor helping someone collaborate with

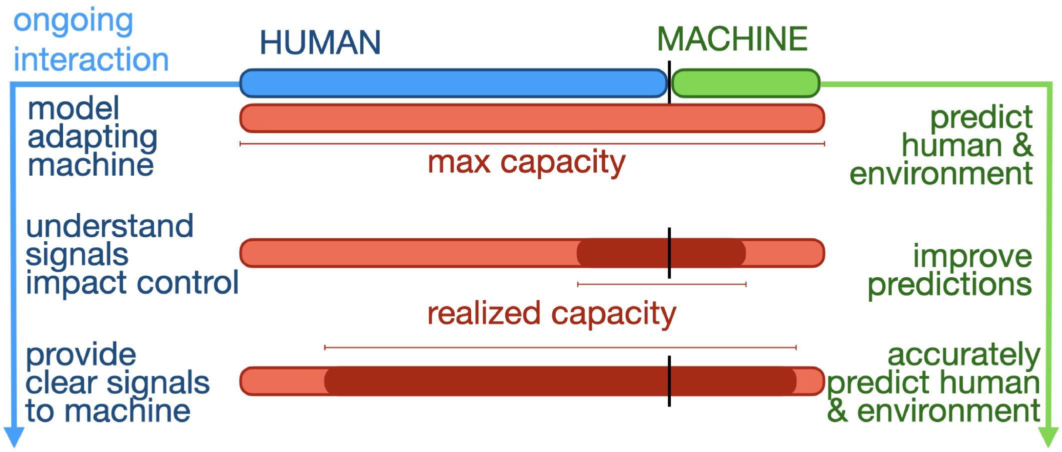


Figure 3.3: *Communicative capital* is acquired by the partnership over the course of ongoing interaction. Prior to the partnership interacting, the human (blue) and machine (green) have acquired no communicative capital and thus have no realized capacity. Then, over the course of ongoing interaction (from top to bottom) through modelling, improved predictions, and understanding of the signals of one another, the partnership acquires communicative capital which leads to increased realized capacity (dark red).

a guide dog (Pfaffenberger, 1976).

In effect, the processes of building communicative capital toward the attainment of goals is about the specification and identification of things each agent cares about, as in "when I do *this* it means *this*". There can then be a natural progression in the interaction as the two sides get to know each other better. For example, in the progression shown in Fig. 3.3, improved predictions represent one form of communicative capital. Beneficial collaboration often requires that at least one agent model and predict information about the other. This modelling of the other enables the partnership to achieve tasks with less effort and less explicit communication. This viewpoint is compatible with perspectives on human-human motor coordination (Candidi et al., 2015) and with prosthetic control approaches like pattern recognition (Micera et al., 2010; Scheme & Englehart, 2011) as discussed below.

In the following sections, we use the idea of communicative capital and the agency-capacity schema defined in Sec. 3 to examine experimental work where prosthetic control has been improved by ongoing interactions between the device and the user. First, we explore human interactions with *adaptive* *mechanisms* like pattern recognition systems in commercially available prostheses, and then we detail interactions with *goal-seeking prosthetic agents*.

3.5.1 Adaptation: Prediction Enhanced Control

First, we consider communicative capital in adaptive control paradigms—specifically, machine learning based prosthetic controllers. There are multiple examples where the human views the machine as adapting and where the machine models and predicts information about the human to better fulfill the human's intentions (Castellini et al., 2014; Edwards, 2016; Edwards, Dawson, et al., 2016; Pilarski, Dawson, Degris, Carey, et al., 2013).

In commercial prostheses with *pattern recognition*, the human engages in a training phase to inform the device about the preferred motions to perform in response to complex patterns of myoelectric activity recorded from the human's body (Castellini et al., 2014; Scheme & Englehart, 2011). The use of pattern recognition can provide users with more intuitive control of their prosthesis (Castellini et al., 2014). The human becomes more skilled at providing clear training commands, in part because of their knowledge that the machine is learning and adapting from the ongoing interaction. The result is improved capacity due to an increase in communicative capital: the number of human controllable functions can now exceed the number of available degrees of control available in conventional myoelectric control which depends on antagonistic muscle pairs for each degree of freedom (Smith et al., 2016).

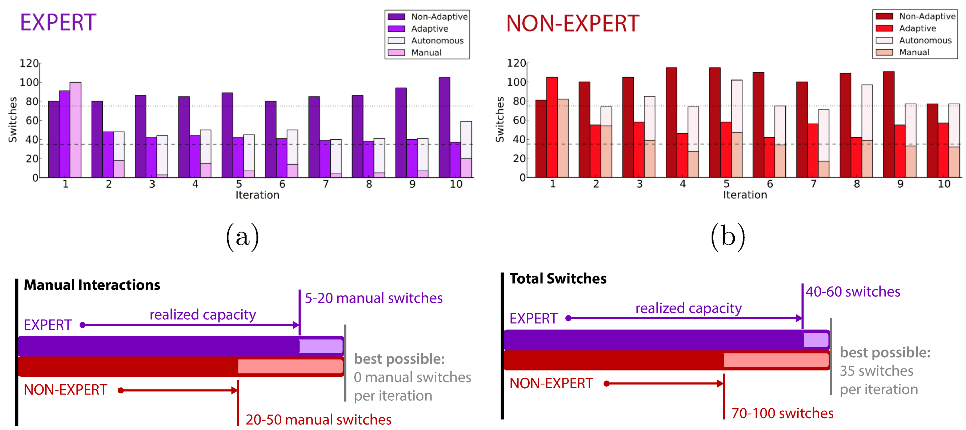
A second example is *adaptive and autonomous switching* (Edwards, 2016; Edwards, Dawson, et al., 2016; Edwards, Hebert, et al., 2016). In this setting, a machine learns to make ongoing predictions about how and when a human will decide to switch between controlling one functional joint of a prosthetic device (e.g. the wrist, elbow, or shoulder) and another (Fig. 3.4). In manual switching, the human uses a separate biophysical control interface to send a 'change currently controlled joint to the next in a fixed list' signal to the device. In adaptive switching, the device adapts to the human by suggesting which joint it predicts the user might want to control next. The human's ability to quickly perform tasks is improved by these suggestions. The device improves

(a) (b)

Figure 3.4: (a) An illustrative example of how adaptive switching enables a prosthetic device to model the way a human uses the functions of a prosthesis and thereby increase agency when compared to the manual switching condition. The increase in shared agency from the manual to adaptive mode of interaction corresponds to increased capacity in terms of (red) time to complete a task, and (purple) the number of switches required to complete a task. This plot shows data approximated from Edwards, Hebert, et al. (2016) for illustrative purposes, and (b) their participant using the device (Edwards, 2016; Edwards, Dawson, et al., 2016; Edwards, Hebert, et al., 2016).

its suggestions based on ongoing observations about the human's actions and preferences. The adaptive nature of the machine, and the increased agency of expert humans to model the machine, lead to increased capacity to successfully complete the task efficiently in terms of reducing both total task time and total switches needed by a human user to complete a task (Fig. 3.4(a,b)). In autonomous switching, the device automatically switches which joint is currently controlled. This is done by making and using predictions to automatically switch between the functional control of different prosthetic device joints (see Fig. 3.5) (Edwards, 2016; Edwards, Hebert, et al., 2016). Predictions are an acquirable form of communicative capital built up by a machine learning agent during its interactions with a human and the environment.

Observations from both adaptive and autonomous switching suggest that the human begins to model the device as an agent that makes predictions (Edwards, Hebert, et al., 2016). As human subjects became more familiar, both with their execution of a task and with the role of machine learning as it adapted to a task, they reported greater trust in the autonomy of the device. In these experiments, certain regions of task spaces were observed where the



(a) (b)

(c) (d)

Figure 3.5: Measured capacity for autonomous and adaptive switching for (a) expert and (b) non-expert humans (plots adapted from Edwards, Hebert, et al. (2016)), summarized in capacity functions relating to (c) the number of manual interactions and (d) the total number of switches required to complete a control task, expert humans realized more capacity than non-experts.

learning system performed with close to 100% prediction accuracy. In these regions, subjects' behaviour suggested they needed to monitor the prosthetic arm less (e.g., the reduced number of manual switches in Fig. 3.5).

In the autonomous switching experiments of Edwards, Hebert, et al. (2016), users began to predict autonomous switches, often moving the next functional prosthetic actuator prior to hearing a cue alerting them to the machine's automatic switching behaviour. Increased capacity in terms of reduced manual switching, and the communicative capital that supports it, is evident in users who have extensive prior experience operating adaptive prosthetic devices (see Fig. 3.5(c,d)). Users who had a greater understanding of the prosthetic learning system tended to perform actions that benefited learning, allowing the prosthetic arm to build up expectations about their behaviour more swiftly.

Another related example is the work of Sherstan et al. (2015). In this work, a human and a machine learning system share agency in controlling the movement of a robotic arm. The user is only able to control a single joint of the arm at a time and must switch between joints as needed in order to complete a task. The machine agent observes the human's behaviour and learns to predict the expected joint angles of the robot arm. These predictions are then used to move the arm in collaboration with the human's own actions (Sebanz et al., 2006).

As a final example of adaptive assistive technology related to the upper-limb prosthetic setting, Xu et al. (2013) describe a walking-aid robot designed to autonomously adapt to different users. The robot uses reinforcement learning to adjust the relative control of the human in real-time for smoother, faster movement. Smoothness of motion, system safety, and intuitive control can all be viewed as different capacity functions that are improved by the adaptive nature of the machine.

3.5.2 Goals: Reward-Based Control

Goal-seeking behaviour on the part of both the human and the machine—behaviour driven by processes of reinforcement learning—enables a more detailed progression of interactions than is possible with an adaptive, but not goal-seeking, machine. What follows is one hypothetical progression of the training of an assistive machine, where both the human and the machine are goal-seeking agents, and where the human starts to model the device as a goal-seeking agent. This modelling and adaptation can be observed behaviourally as in the previous section.

1. At the outset, the human can only provide positive feedback (i.e. reward) signals indicating their approval; no other signals have any agreed upon meaning.

2. Using these rewards, the machine can learn a function that maps signals from the human, or other environmental cues, to a valuation that is grounded in cumulative reward (a *value function*, as detailed by Sutton and Barto (2018), and used in face valuing by Veeriah et al. (2016)).

3. Using this value function, the human teaches the machine a convention that may be used to interact at a low level—e.g., simple commands, body language, cues like pointing, and the basics of shifting between different functions of a system. The human begins to model how their behaviour affects the learning and adaptation of the machine.

4. Using these developed conventions, higher-level abstractions can be established between the human and the machine. These built-up conventions are one component of communicative capital which enable the realization of additional partnership capacity.

With this progression in mind, there are a variety of compatible ways to incorporate human knowledge into a learning system (Amershi et al., 2014; Chernova & Thomaz, 2014; Pilarski & Sutton, 2012; Thomaz & Breazeal, 2008). Starting with the idea of training based on primary reward, as in the progression described above, Knox and Stone (2012) introduced the *Interactive Shaping Problem*, wherein an agent is acting in an environment and a human is observing the agent's performance and providing feedback to the agent such that the agent must learn the best possible way to act based on that feedback (Knox & Stone, 2012). The interactive shaping problem is related to communicative capital, as it is a readily observable case of information sharing between two goal-seeking systems with a limited channel of communication.

Goal-seeking behaviour in a machine, and developing communicative capital through the human's modelling of the machine as goal-seeking agent, increases the maximum capacity of a partnership. A human's interactions with a machine are supported by a channel of communication with defined semantics (e.g., the reward channel in reinforcement learning (Sutton & Barto, 2018)) that allows the human to shape the machine's behaviour in ways that are not possible for an adaptive, non-goal-seeking machine. This communication channel is integral to realizing the goal-seeking agent's capacity to deal with non-stationary tasks, changing problem domains, and novel environments, in a way that aligns with the human's goals. Providing the means by which to shape behaviour can also reduce the amount of pretraining for the system, as interactions are now accompanied by online, real-time human feedback. Reward allows the human to shape the machine learning agent to perform the task in a personalized, and situation-specific way—an adaptive goal-seeking agent has the ability to incorporate engineered knowledge, but also move beyond it.

Previous work has demonstrated how both predefined and human-delivered reward could be provided to a goal-seeking agent to gradually improve the control capabilities of a myoelectric control interface (Pilarski, Dawson, Degris, Carey, et al., 2013; Pilarski et al., 2011). By using a goal-seeking reinforcement learning agent to control the joints of a prosthesis, informed by predictions about future movement, the human-machine partnership was found to be able to progressively refine the simultaneous multi-joint myoelectric control of a robotic arm. In these studies, human approval and disapproval was delivered to the machine with full knowledge of the machine's learning capabilities. These initial results have been extended to more complex settings which informs how mutual, goal-seeking behaviour supports myoelectric control (Mathewson & Pilarski, 2016). These results demonstrate the value of developing communicative capital through the explicit incorporation of human feedback signals. In this representative work communicative capital led to an increased partnership capacity.

3.5.3 Models, Shared Agency, and Feedback

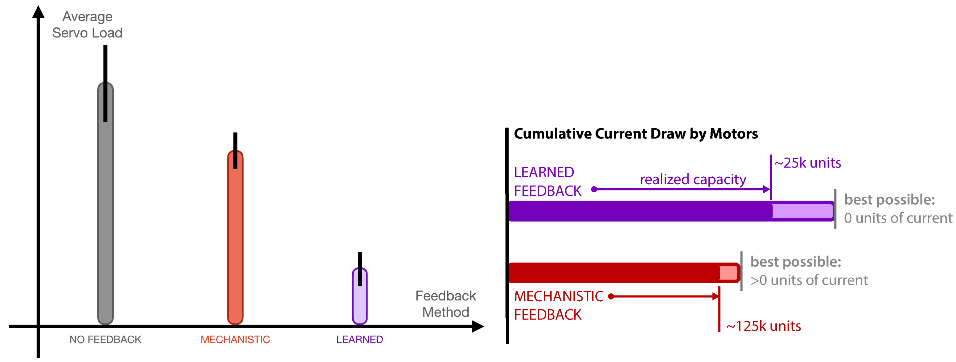
Beliefs about the nature of internal and external signals are a kind of knowledge that we broadly denote as *models*. Models are required for the higher level attributes of agency; it is useful for a machine to represent, or construct a model, of its partner and the world, in order to achieve more effective interaction. Agent models, as they apply to a human-prosthetic partnerships, may take many forms. They may include, for instance, a collection of learned, temporally extended predictions about the dynamics of the world and the behaviour of the human (Pilarski & Sherstan, 2016; Sutton & Barto, 2018; Sutton et al., 2011).

As described by Pezzulo and Dindo (Pezzulo & Dindo, 2011) shared representations may be a critical part of communication during human-machine interaction, and central to the formation of more effective models in terms of beliefs, actions, and intentions. This moves us towards developing a theory of mind—an agent predicting the internal beliefs, motivations, and thoughts of another especially as applied to observable sensorimotor interactions (Candidi et al., 2015; Pezzulo & Dindo, 2011; Pezzulo et al., 2013). Recursive theory of mind might imply higher levels of agency, as presented in Section 3.3, and parallel higher order intentionality (Cultural General Intelligence Team et al., 2022; Heyes, 1998; Rabinowitz et al., 2018; Zhu et al., 2021). Future work may explore this other-modelling and how it can be leveraged to build shared knowledge.

As one example of how models can impact a human-machine partnership, Bicho et al. (2011) describe a shared construction task in which a robot and a human must work together to assemble a toy. Completion of the assembly task required actions from both agents. The robot infers the goal of the human from contextual clues and acts accordingly, communicating its intention at each point during the task using a speech synthesizer. This allows the human to further model the internal processes of the machine. Another example of a joint task in which a robot infers the goal of the human comes from Liu and Hedrick (Liu & Hedrick, 2016). In their work, participants and virtual robots collaborate to accomplish a task, and the robot infers the human's goal based on motion. This research suggests that goal inference (i.e., the modelling goals) decreased the time required to finish tasks and improved other measures of performance, including human-machine trust.

The impact of feedback from an adaptive prosthetic is quantified in work by A. S. R. Parker et al. (2019a). In their work, three different kinds of feedback were used to supply a human with information about how best to control the movements of a wearable robot in the form of a supernumerary limb (see Fig. 3.1(c))—no feedback, mechanistic feedback, and adaptive feedback in the form of predictions. The human needed to move the robot in a confined work space, coming as close as possible to the work space's walls without making physical contact. The human was blindfolded and was acoustically isolated by way of noise-cancelling headphones, so that they only received information about the world via the machine's feedback.

The two capacity functions of interest in A. S. R. Parker et al. (2019a) measured: the current drawn by the motors due to impacts with the work space walls, and the number of times the human was able to use the arm to fully



(a) (b)

Figure 3.6: The difference between learned and mechanistic feedback during control of a supernumerary limb (Fig. 3.1(c)). (a) Adaptive machine significantly reduced current drawn by the motors of the robot arm (adapted from A. S. R. Parker et al. (2019a)). (b) Increased machine agency increase realized capacity of the partnership through an investment of communicative capital.

traverse the work space in the given time. On different trials, feedback from the device was either absent, delivered mechanistically upon contact with the walls, or delivered proportional to learned predictions about impacts with the walls. Realized capacity in terms of current draw was found to increase for the case where the human was paired with the adaptive machine, but was found to approach a reduced maximum capacity for the case of mechanistic feedback from the device (see Fig. 3.6[1]). This work provides insight into how developing communicative capital, specifically through explicitly modelling and increased agency in the delivery of feedback, can influence the maximum capacity possible for a human-prosthetic partnership.

3.6 Discussion

This chapter has discussed the setting of human-machine interaction, specifically the interactions between a human and their prosthetic technology. However, the ideas presented above regarding agency and communicative capital can be identified and analyzed in the interactions between any two or more intelligent systems. In this section, we provide supporting context from both biological and non-biological examples of how agency plays a role in the inter-

[1]The material of A. S. R. Parker et al. (2019a) is in Chapter 4 of this dissertation.

actions of multiple agents to achieve a goal.

3.6.1 Guide Dogs and Intelligent Assistants

A guide dog could be the oldest documented example of an assistive technology with agency, with an early depiction on the wall of a house excavated in Pompeii dated from c. 79 CE (Fishman, 2003; Pfaffenberger, 1976). A guide dog needs to be part of an active partnership—it must have the capability to willingly disobey an instruction when it perceives a danger. The agent in charge of the interaction, human or dog, needs to be able to change from moment-to-moment in order for the partnership to be effective. Because of these desired and atypical behaviours, both the dog and the future owner must be explicitly trained. The human must be taught not only the precise vocabulary understood by the guide dog, but what to expect in response. This requires both parties, human and dog, to invest in communicative capital and learn each others' idiosyncrasies in order to approach an effective partnership (Serpell & Hsu, 2001).

Computers, whether desktops, tablets, or smartphones, all augment our cognitive abilities. At present, there is significant effort to develop virtual assistants on such devices. Such assistants may have some level of agency; these assistants may be adaptive, changing their behaviour and suggestions to meet the user's needs (Markoff, 2015). To date, existing computer interfaces have largely remained fixed and unadaptive. However, thanks in part to increases in available computation, computers are now improving in their ability to predict user needs and to provide users with the information and interfaces that are most needed at any given moment (Langley, 1997; Markoff, 2015). With increased agency, these systems now begin to demonstrate some of the hallmarks of human-human joint action established by the related literature (Knoblich et al., 2011; Pesquita et al., 2018; Vesper et al., 2010).

3.6.2 Interactive Approaches to Instruction, Communication and Control

There are multiple ways that a human and a machine—e.g., an assistive robot like a prosthesis—can beneficially interact to achieve the human's objectives (Argall et al., 2009; Pilarski & Sutton, 2012; Thomaz & Breazeal, 2008). A pertinent family of methods, broadly classified as *interactive machine learning* (IML), has demonstrated the potential to increase the capabilities of decision making systems in complex, dynamic, and novel environments.[2] In much of the existing IML literature, feedback channels are used as a means by which a non-expert can train, teach, and interact with a system without explicitly programming it. Shaping allows for the human to learn how the system accepts and interprets feedback and for the system to learn the goals of the human (Thomaz & Breazeal, 2008).

IML has produced a number of important milestones. With respect to goal-driven systems, trial-and-error machine learning has been shown to be accelerated through the presentation of human-delivered reward and forms of intermediate reinforcement. Examples include the use of shaping signals (Kaplan et al., 2002), the delivery of reward from both a human and the environment (Knox & Stone, 2012), multi-signal reinforcement (Thomaz & Breazeal, 2008), and combinations of both direct control and reward-based feedback (Mathewson & Pilarski, 2016; Pilarski, Dawson, Degris, Carey, et al., 2013; Pilarski et al., 2011). As described in Sec. 3.5.2 above, an agent's learning can be facilitated by a human host through interactive reinforcement learning (Knox & Stone, 2009, 2012; Knox et al., 2013). Griffith et al. (2013) built on the earlier work of Knox and Stone (Knox & Stone, 2009) with a framework to maximize the information gained from human feedback. Loftin et al. (2014) expanded the space of human interaction through detailed investigation of human teaching strategies and developed systems which model the human feedback. Their systems have been shown to learn faster and with less

[2]Though, some argue that **all machine learning is interactive machine learning** because humans interact with machines through every step of the design, development, deployment, and dissemination of such systems (Mathewson & Pilarski, 2022).

feedback than other approaches. Interactive learning from demonstrations and instructions have also been shown to help teach different ways of behaving to a learning machine (Abramson et al., 2020; Argall et al., 2009; Chao et al., 2010; Judah et al., 2010; Kaplan et al., 2002; Lin, 1991, 1993).

Humans can utilize a number of different approaches to effectively communicate their goals to machine learning agents. Through interactive learning, information from a human can help a machine learner to achieve arbitrary user-centric goals, can improve a system's learning speed, and can increase the overall performance of a learning system. Advances in IML provide a basis for increasing the rate with which a human-prosthetic partnership may develop communicative capital and thereby realize capacity, and, in certain cases, can also be expected to increase the maximum capacity of a partnership.

3.6.3 Limitations

There are challenges and limitations in creating machine agents that can build up communicative capital to collaborate more effectively with their human partners. In this section, we highlight several critical areas of focus that should be addressed in future work. Of particular note are challenges related to safely deploying machine learning algorithms in the real-world, especially when deployed on robots tightly coupled to human users. Future work on these algorithms is needed to empirically demonstrate how they are provably robust to a wide variety of environmental factors. As well, mechanisms to align the goals of the human and the machine are critical in shared agency settings. It has been shown in previous research how increasing agency of the machine increases the cognitive demands placed upon the human (Mathewson & Pilarski, 2016). Human's often expect machines to function as mechanisms, unaffected by adaptation. There can be significant implications on their cognitive load once they are required to carry out their own actions as well as model the learning agent (Hedlund et al., 2021). Finally, algorithms deployed in human-machine partnerships will need to adapt quickly to information and signals from the human. Both for reasons of safety, but also because a lack of quick adaptation could lead to human disengagement if the human doesn't perceive the machine as learning fast enough. Future work on safety, alignment, rapid adaptation, understanding human expectations, and making connections between these systems and modern theories of agency is needed as human-machine partnerships move from the laboratory and into the world. This is true for both prosthetic devices and for collaborative machines more generally.

3.7 Paradigms for Evaluation

We expect that increasing the agency of a prosthetic device and investing in communicative capital will allow a collaborative partnership to accomplish tasks faster, easier, more safely, and more efficiently. Work is now needed to test this hypothesis and identify the contributions and practical utility of agency and goal-seeking behaviour on the part of machine learning partner agents. It is our recommendation that researchers design experiments varying the level of agency of both human and the machine in a controlled fashion to assess the contributions from each component of agency. As described in Sec. 3.5.2, increased agency on the part of the machine enables increased shared agency. This increase is depicted as relative changes in the agency and capacity of both agents.

One means by which to test agent contributions is through the conventional outcome measures used to assess the impact of rehabilitation interventions (Hebert et al., 2009; Light et al., 2002; Resnik, 2011; Resnik et al., 2012). Outcome measures provide a clearly defined notion of capacity. Further, prosthetic outcome measures are already used to study the benefits of pairing patients to systems with different mechanistic levels of agency (e.g., during prosthetic fitting and patient assessment). In the majority of clinically deployed prostheses, the control approach and system design of the device is fixed. The communicative capital of the mechanism—how it interprets body signals and maps them to actuators—provides immediate realized capacity at a level determined by the mechanism's designers. Measures like the Southampton Hand Assessment Procedure, the Box-and-Blocks Task, and others are used to provide a quantitative assessment of the impact of these prosthetic mechanisms (Light et al., 2002; Mathiowetz et al., 1985). Recent developments in the assessment of gaze and movement have further shown concrete, capacity-related metrics that evaluate user-prosthesis abilities via changes in the relationship between biomechanics and visual attention, as well as other measurable correlates of perceived control and agency (Hebert et al., 2019; Marasco et al., 2021; Williams et al., 2019, 2021). Some of these measures have been shown to serve as proxies for the state of human predictive models of their machine partner, and thus may provide a way to quantify communicative capital as it is built by the human side of a human-machine partnership (Marasco et al., 2021). Rigorous, incremental testing of agency is therefore highly compatible with existing approaches, and will be significantly extended as more comprehensive motor, sensory, and cognitive outcome measures are developed.

One fruitful avenue for experimentation, as explored in A. S. R. Parker et al. (2019a), is to deliberately reduce the agency of the human by removing control options and/or sensory inputs as they complete a task. In this way, the authors were able to elucidate how different levels of agency in the machine contribute to the performance of the partnership. A second, complementary paradigm is to dramatically increase the agency of the machine beyond what is technically possible, so as to study the outcomes and conditions that support shared agency. One way to do this is a type of sham trial known as a *Wizard-of-Oz* experiment (e.g. Viswanathan et al. (2014)). Paradigms for evaluating human-machine partnerships will continue to develop as technology supporting shared agency evolves. We now conclude with several brief reflections.

3.8 Conclusions

We argue that tightly coupled human-machine partnerships, such as humans and prostheses, should be thought of as adaptive multi-agent systems where the agency of human and machine combine to achieve more capacity than either could independently. We present an agency-capacity schema that relates shared agency to the capacity of human-machine partnerships, and we show

how communicative capital is the key resource that a partnership needs to invest in to access the full capacity of the combined agency of the pairing. Using examples from the literature, we illustrate how increases in the agency of a prosthesis can tangibly improve the capabilities of its human user. We highlight three main conclusions from this work as novel contributions supporting human-prosthesis interaction: 1) we propose that designing assistive devices as goal-seeking agents improves the range of possibilities for robust and flexible interaction, 2) we argue that an agent-based viewpoint of human-machine interaction enables a structured progression toward more capable partnerships between people and devices, and 3) we describe how communicative capital is a resource built through ongoing human-machine interaction which enables a partnership to eventually perform tasks at a capacity greater than either could individually. Machine intelligence enables the acquisition and use of communicative capital in human-prosthesis partnerships to more effectively and more efficiently accomplish tasks. We believe the agency-based viewpoint on assistive technology proposed in this work contributes unique and complementary ideas to the development of highly functional human-machine partnerships. Designers and developers should construct systems which actively invest in communicative capital as such investment will lead to increases in shared agency to achieve more capacity than they would be able to otherwise.

Chapter 4

Exploring the Impact of Machine-Learned Predictions on Feedback from an Artificial Limb

Learning to get by without an arm or hand can be very challenging, and existing prostheses do not yet fill the needs of individuals with amputations. One promising solution is to improve the feedback from the device to the user. Towards this end, we present a simple machine learning interface to supplement the control of a robotic limb with feedback to the user about what the limb will be experiencing in the near future. A real-time prediction learner was implemented to predict impact-related electrical load experienced by a robot limb; the learning system's predictions were then communicated to the device's user to aid in their interactions with a workspace. We tested this system with five able-bodied subjects. Each subject manipulated the robot arm while receiving different forms of vibrotactile feedback regarding the arm's contact with its workspace. Our trials showed that using machine-learned predictions as a basis for feedback led to a statistically significant improvement in task performance when compared to purely reactive feedback from the device. Our study therefore contributes initial evidence that prediction learning and machine intelligence can benefit not just control, but also feedback from an artificial limb. We expect that a greater level of acceptance and ownership can be achieved if the prosthesis itself takes an active role in transmitting learned knowledge about its state and its situation of use.

4.1 Introduction

The loss of a limb, especially an upper limb, can have a significant impact on an individual. A person may be missing a limb from birth, or it could be the result of illness or injuries sustained over the course of one's life. Artificial limbs, also called prosthetic limbs, are often seen as a means of mitigating the absence of a biological limb. In all cases, but particularly when a limb is lost later in life, it can be very difficult to adapt to interacting with the world through a mechanical or electronic device (Antfolk et al., 2013; Hebert et al., 2014; Micera et al., 2010; P. A. Parker et al., 2006; Peerdeman et al., 2011; Resnik et al., 2012; Scheme & Englehart, 2011; Williams III, 2011). There are many prostheses on the market that attempt to fill the needs of individuals with amputations, and many of these have tremendous potential to restore lost functionality and independence to the user; however, even the best prostheses currently available have limitations (Peerdeman et al., 2011; Resnik et al., 2012; Williams III, 2011). There are two major areas where current prostheses begin to show the strain of insufficient technology to properly support them. The first area is a lack of feedback(Antfolk et al., 2013; Hebert et al., 2014; Micera et al., 2010; Peerdeman et al., 2011)—e.g., the sense of touch—and more important to this work, lack of proprioception when using a prosthesis (Antfolk et al., 2013; Williams III, 2011). The second area is insufficient control (Micera et al., 2010; P. A. Parker et al., 2006; Peerdeman et al., 2011; Scheme & Englehart, 2011; Williams III, 2011). Under most current techniques, the person who needs to control the limb has fewer control channels available to them than their device has functions (Micera et al., 2010; P. A. Parker et al., 2006; Scheme & Englehart, 2011). This leads to some clever, but non-natural, control solutions such as routing some of the control channels to alternate locations on the user's body. A final challenge which results from the first two is acceptance of the prosthesis by the user (Peerdeman et al., 2011; Resnik et al., 2012; Williams III, 2011). Despite the great clinical potential of

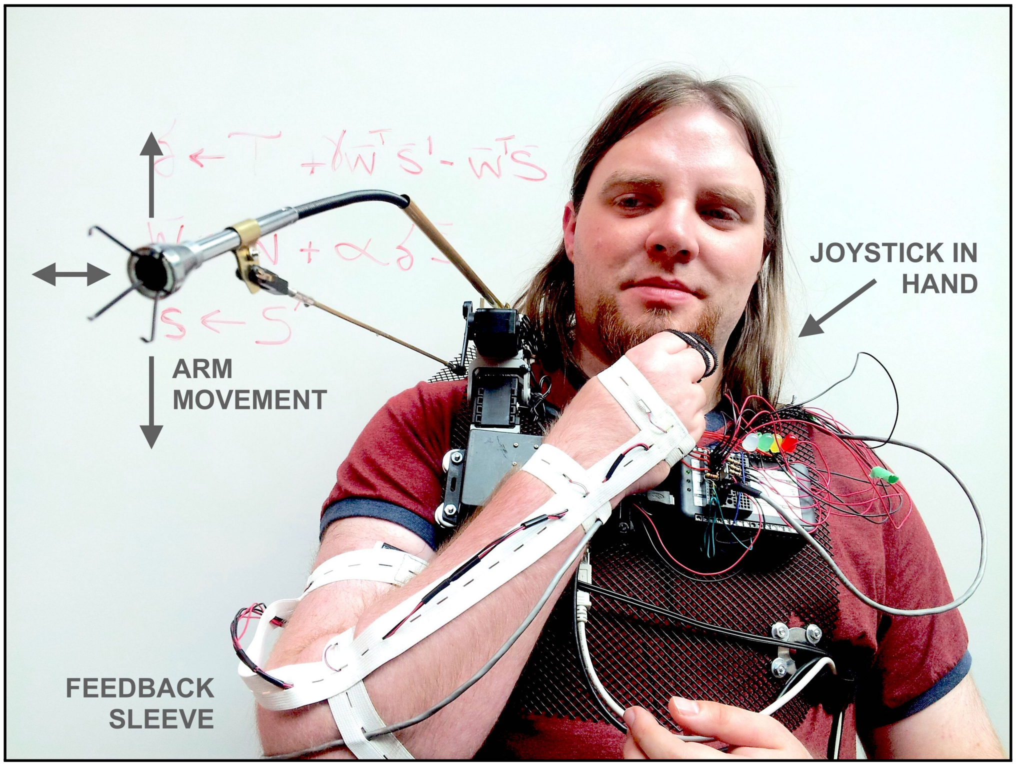


Figure 4.1: Wearable robot limb system used in these experiments. The four degree-of-freedom arm is controlled by a joystick in the user's hand which sends signals to an ADC and then to a laptop, which in turn commands the servos. A vibrotactile feedback sleeve can provide feedback to the user.

many modern prostheses, as a result of the first two limitations a prosthetic can be perceived by the user as insufficient or as a reminder of the functionality that they lost and that the device simply cannot restore (Peerdeman et al., 2011; Resnik et al., 2012; Williams III, 2011). Lack of acceptance is especially prominent in the newer myoelectric (EMG) prostheses, i.e., electrically driven robot limbs, versus the older mechanical types despite the increased potential that myoeletric prostheses have in overcoming the other challenges (Scheme & Englehart, 2011; Williams III, 2011).

Operating a device that interacts with the world is a learned motor function. As infants, we learn the way our limbs interact with our environment through general motion and play (Flanagan et al., 2003; Wolpert et al., 2001; Zacks et al., 2011). This develops the control channels and models required for us to use our bodies to sense and manipulate the world we live in (Wolpert et al., 2001). This interaction involves two parts (Flanagan et al., 2003; Wolpert et al., 2001). The first is the internal forward copy of the action—in effect, knowledge that moving specific muscles will cause a motion which results in the desired sensory feedback. There is also a reverse copy that is processed at the same time. The reverse copy starts at the desired interaction with the environment and links the required muscle action to it. In order to skillfully interact with the environment, both the forward and reverse models must be present (Flanagan et al., 2003; Wolpert et al., 2001).

Artificial intelligence offers a promising solution to the control problems encountered by the users of electromechanical prostheses (Pilarski, Dawson, Degris, Carey, et al., 2013). Offline machine learning in the form of pattern recognition is for the first time seeing use in commercial prostheses, and is considered to be the state-of-the-art in controlling multiple prosthetic joints (Micera et al., 2010; Scheme & Englehart, 2011). Real-time machine learning has also recently been used to ease the control burden on a user by learning joint activation sequences as a limb is being used (Pilarski et al., 2012; Pilarski, Dick, et al., 2013); as one example, predictions about a user's control choices have been learned so as to minimize the number of switches between joints, and consequently the time required to perform a task (Pilarski et al., 2012).

A lack of feedback is frequently responsible for abandonment of prosthetic devices, especially upper-limb prostheses (Biddiss & Chau, 2007b). Feedback is an important aspect of control, and how to provide feedback from upper-limb prostheses to individuals with amputations is an active area of research (Antfolk et al., 2013; Schofield et al., 2014). There are many modalities and means of feedback that are being explored currently. Some examples are substitution, where a signal that is not meant to imitate the lost physiological system is used, and modality matched, where an attempt is made to imitate the physiological sensations (Schofield et al., 2014). Providing feedback in these ways has been shown on subjects without amputation to improve performance on grasping tasks, as outlined by Schofield et al. (2014)

The primary contribution of the present work is to suggest that machine intelligence can be used to enhance not just control—the focus of most prosthesis-related machine intelligence research to date—but also *feedback from a prosthesis*. This feedback was part of a user's intact biological system, and contained information used in operation of their natural limb. In the case of a prosthetic limb, motor awareness and forecasting are now at least partly encoded in the hardware of the prosthesis rather than in a user's biology. Therefore, we may need to provide assistance to the natural system in interfacing with its electronic components. We suggest that machine intelligence can be used to take the internal state of the assistive device and interpret it in ways the biological system cannot do naturally; the results of this interpretation can be communicated to the user in a variety of ways to improve their control over the device. Thus, using machine intelligence, we can help create a forward prediction of an action electrically and communicate it to the user, similar to the operation of the intact biological system.

This work therefore contributes a preliminary exploration of the application of machine-learned predictions, expanding upon the work started by A. S. R. Parker et al. (2014). A simple system for communicating machine-learned predictions via vibro-tactile feedback is used to assist a user in refining their own forward model of motor actions while using a prosthetic limb analog. Specifically, temporal-difference learning is used to generate a prediction about the electrical load the servos of a human controlled robot arm will experience as they near a potentially dangerous collision with objects in the user's environment. This prediction is communicated to the user through a vibration motor. In this way, we emulate the forward predictive model present in a biological limb's motor function. We expect that, similar to the way that the biological operation of a limb is dependent on its forward copy, the addition of an electronic/computational equivalent during human-robot collaboration will yield control improvements over purely reactive feedback. In this study, the amount of load experienced by a servo over the course of an experimental run when the user receives this predictive feedback is compared to the same user receiving the same indication when the servo is actively experiencing high load (reactive feedback).

4.2 Methods

4.2.1 Robot and Experimental Platform

The experimental platform used in this work was a custom-designed robotic arm called the ExArm (Fig.4.1), which was wearable by individuals without amputation. The arm was designed to model the gross motor functionality of joints in a human arm. It had four controllable actuators: shoulder, elbow, wrist flexion, and hand open/close (AX-12/18+ Dynamixel servo motors). Subjects used a 2-axis thumb joystick (SparkFun) to control the motion of the ExArm's joints, and pressing the joystick could change the active joint. The joystick was connected to an ADC (DI-149 data acquisition starter kit, DATAQ Instruments), which digitized the 3.3 V signal modified by the user's control of the joystick. The resulting output signal was sent via USB to a computer, which interpreted the signals and sent commands to the robot's servos. The control software only utilized information from a single axis of the joystick for motion, as well as the joystick button press to indicate a joint switch, to emulate EMG control of a prosthetic limb. The velocity of motion was fixed for all participants in all trials; speed of arm motion was a constant value.

AX-12/18+ servos used in the design of the ExArm provided several useful output signals, including their angular position, angular velocity, motor temperature, voltage, and load. To communicate feedback about these sensors to the user, we designed a custom sleeve embedded with four vibration motors (termed *tactors*) similar to those used in a cellphone or pager. With the sleeve donned, one tactor each was located over the user's shoulder, elbow, wrist, and hand, as shown in Figs. 4.1 and 4.2. The platform therefore emulated the capacity for actuation in many commmon prosthetic devices while adding vibrotactile feedback.

4.2.2 Experimental Procedure

Five subjects were asked to participate in experiments with the ExArm, and gave informed consent in accordance with the study's institutional review

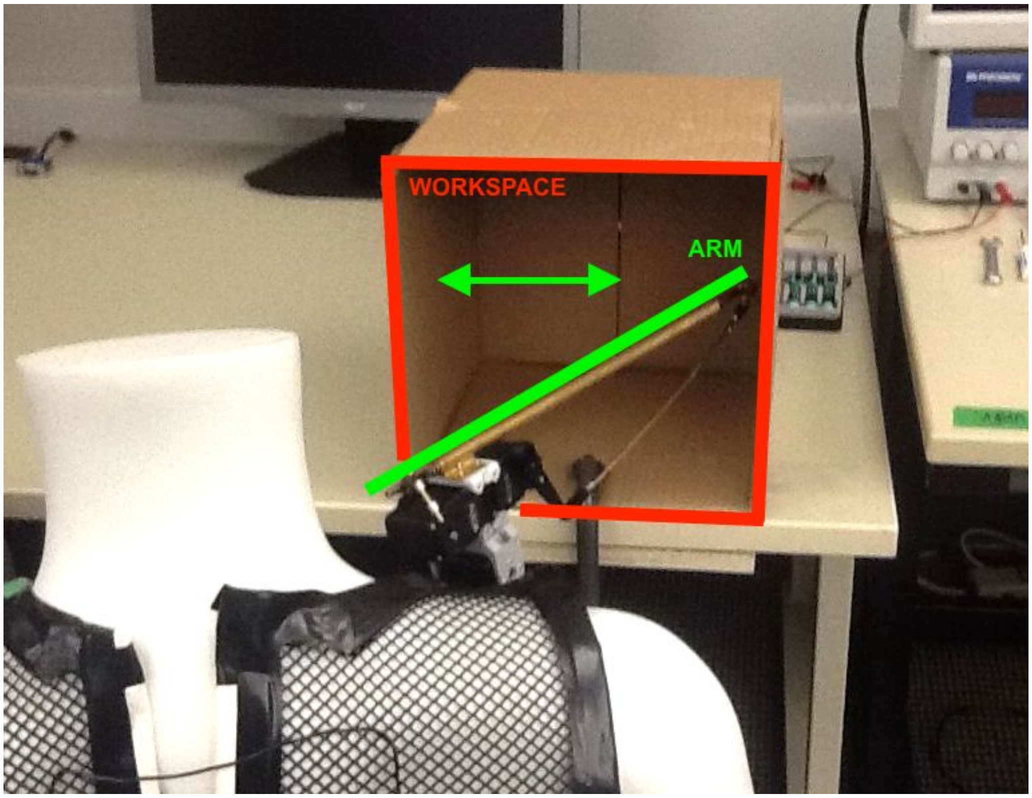


Figure 4.2: The experimental setup: a confined workspace (red), the robotic arm (green), and as in Fig. 1, an experimental subject with attached vibrotactile feedback sleeve (seated to the left of the workspace, not shown).

board approval. Each user wore the sleeve containing the vibration tactors and controlled the back-and-forth motion of the robotic arm's shoulder joint using the thumb joystick. The other joints of the arm as well as the joint switching functionality were not used for this experiment to restrict the motion of the arm to a single path. The ExArm was affixed to a stationary mannequin as shown in Fig. 4.2 to ensure each experiment began with the robotic arm at a constant position and to mitigate the effect of a user's trunk movement. Thus, for this initial work, the position and movement of the user was unrelated to the outcome of the experiment. The workspace was a subspace of the shoulder joint's total range of motion, bounded by a 27 cm square box that was fastened in place. Prior to each experiment, the end effector was centered with respect to the workspace, perpendicular to the rear wall of the box and equidistant from the left and right walls. Each subject was asked to perform four separate five-minute tasks, structured as follows:

Training Task

The first task was designed to provide users with practice controlling the ExArm. For this training task, the user was asked to move the arm repetitively from one side of the box to the other using the joystick, pausing briefly ($\leq$1 second) upon reaching the center of the workspace. At the left and right walls of the box, the user was tasked with pushing the robotic arm against the wall until the arm was fully flexed, causing a temporary increase in the load reported by the servo motor. The user's shoulder vibration tactor was programmed to vibrate at a load threshold of 650 out of a maximum reportable load reading of 1024. This vibration communicated to the user that the load had exceeded the maximum threshold considered safe for the robotic arm, and the arm should be moved away from the wall. In addition to providing each user with practice manipulating the arm, this task produced the source data for prediction learning (described below).

No-Feedback Task

Each subject performed the second task without any knowledge of the position of the arm within the workspace other than its starting location. In order to establish a baseline with no visual, auditory, or tactile feedback, subjects were given a blindfold and listened to music through earphones throughout the task. The volume of the music was increased to a comfortable level at which they could not hear the arm tapping the walls of the box. During this task, vibratory feedback about load was also turned off. The instruction given to the user for this task and those that follow was to avoid excessive load on the servos by not colliding with the barriers too harshly while approaching the left and right walls closely in an alternating fashion.

Reactive-Feedback Task

The next task was identical to the no-feedback task; participants were blindfolded and sound isolated and asked to navigate from wall to wall without stressing the servos with collisions. For this trial the participant was provided with reactive vibration feedback when the current load experienced by the robot arm's shoulder servo reached a threshold of more than 420, determined experimentally. The maximum value of the load recorded during the trials was 827.87, which means the threshold was 50.7% of the maximum experienced. Thus, the tactor triggered every time the user hit a wall but not during travel in between. This task provided an indication of the effectiveness of having reactive tactile feedback only, and specifically examined how well the user could approach each wall without incurring a forceful impact when feedback was delivered at the moment the arm first contacted the wall.

Predictive-Feedback Task

For the final task given to participants, users were again blindfolded and sound-isolated, and given the same task as the previous two trials. In this case, they were provided with tactile feedback from predictions of the electrical load on the robot's arm servo motor. Predictions were provided by a real-time machine learning system trained while the participant was performing task 1. This prediction learning system is described in the following section. When the load prediction rose above 900, determined experimentally, the shoulder tactor was programmed to vibrate. The maximum prediction during the trial was 3857.5, which means the threshold was 23.3% of the maximum prediction value. This task was designed to determine how communicating the learned prediction of load changed the user's ability to approach the wall without incurring a forceful impact.

All load and prediction thresholds used were determined from the analysis of data prior to experiments. We determined the noise level of the load signal while traversing the workspace and set the thresholds so they would not trigger during travel. The prediction threshold and learning parameters were also set so as not to signal an impending high load event too early in travel.

4.2.3 Machine Intelligence and Prediction Learning

The main component of this study is an incremental prediction learner to generate expectations about future impact given learned knowledge about the user's previous motion choices, their outcomes, and the current state of the robot arm. To make predictions about the world, intelligent systems require sensory inputs. These inputs can then be divided into discrete states for increased or decreased resolution. The shoulder joint of the ExArm has a rotation range of 300°. In our protocol, we used the servo encoders value to determine the position of the shoulder joint as a sensory input, divided into 32 distinct states (termed *bins*). These states were motion-dependent; as such, each of the 32 states was further expanded into three: one set of 32 position bins used to represent the state when the servo is moving clockwise, a second set to represent the position while the servo is moving counter-clockwise, and a third set that represent the position when the servo is not moving. The immediate state of the arm was noted in a feature vector (denoted x, of length 96) as a single active bit indicating the current position and direction; this feature vector also contained a single active baseline unit. A weight vector of corresponding length, denoted w, was used to store the learned predictions about the interactions between the robot arm and the walls of the workspace.

The weight vector w was learned from data using standard techniques from temporal-difference learning and recent generalized value function methods, as outlined for the prosthetic setting in Pilarski, Dawson, Degris, Carey, et al. (2013) and more generally in Modayil et al. (2014). Weights w were updated on each time step according to the temporal difference between the instantaneous load being reported by the servo (denoted τ) and predictions about the immediate and next load readings (the inner products $w_t^\top x_t$ and $\gamma w_t^\top x_{t+1}$, respectively, where γ is the timescale or level of temporal abstraction for the prediction of interest). The update to the weight vector on each timestep t was done according to:

$$w_{t+1} = w_t + \alpha(\tau_{t+1} + \gamma w_t^\top x_{t+1} - w_t^\top x_t)x_t,$$

where α represents a step-size (learning rate, set to $\alpha = 0.1$ in these experiments). The temporal abstraction for predicting the load signal of interest was set to $\gamma = 0.92$; this means the prediction learner was acquiring knowledge about the exponentially discounted expectation of the electrical load experienced by the robot's shoulder servo motor over the next ~12 time steps, or 0.6 seconds; the system learned and operated within a control cycle of roughly 20 Hz (50 ms time steps). This knowledge could then be retrieved and used in predictive feedback by reporting the prediction as the inner product $w_t^\intercal x_t$. As noted above, in the predictive feedback task, vibratory feedback to the user was triggered when the prediction's value exceeded a fixed threshold, indicating an impending collision with the walls of the workspace.

Learning was only enabled during the training task, such that the system acquired and updated user-specific predictions about servo motor load while each subject was performing their first task. Learning weights were then frozen (i.e., $\alpha = 0$) during all remaining tasks, including the predictive feedback task. Learning could in principle continue during all tasks; however, for clear assessment of the principles of interest, our experimental protocol featured defined training and testing periods.

4.3 Results

When compared to the case where purely reactive feedback was given to the user, giving learned predictive information as feedback was found to reduce the load experienced by the shoulder actuator of the robot limb. One way repeated measures ANOVA was used to analyze the difference in load when the feedback system was triggered differently. As shown in Fig. 4.3, the average load across all participants (N = 5) for each entire trial was significantly less with predictive feedback than with reactive feedback. For these comparisons, Maulchy's W indicated sphericity could be assumed (0.602). The uncorrected F statistic was $F(2, 8) = 16.385$, $p = 0.001$. The difference is specifically between the predictive feedback case and the other two (no feedback to predictive feedback $p = 0.031$, reactive feedback to predictive feedback $p = 0.038$). No significance was found between the no feedback and reactive feedback cases.

There was a notable increase in visits to more central positions when using predictive feedback. Figure 4.4(a–c) shows the frequency of visits to each position as seen by the system (bin). Results shown are the average for each

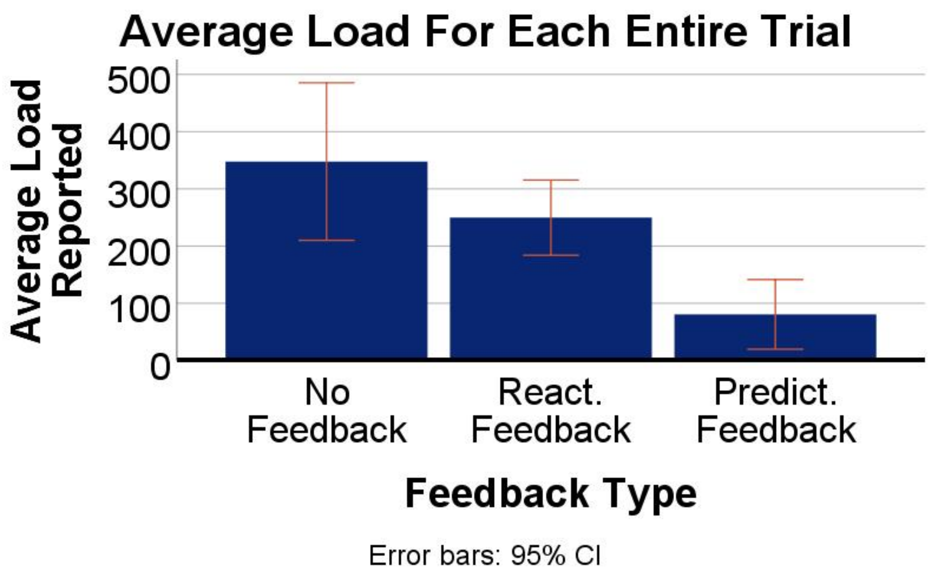


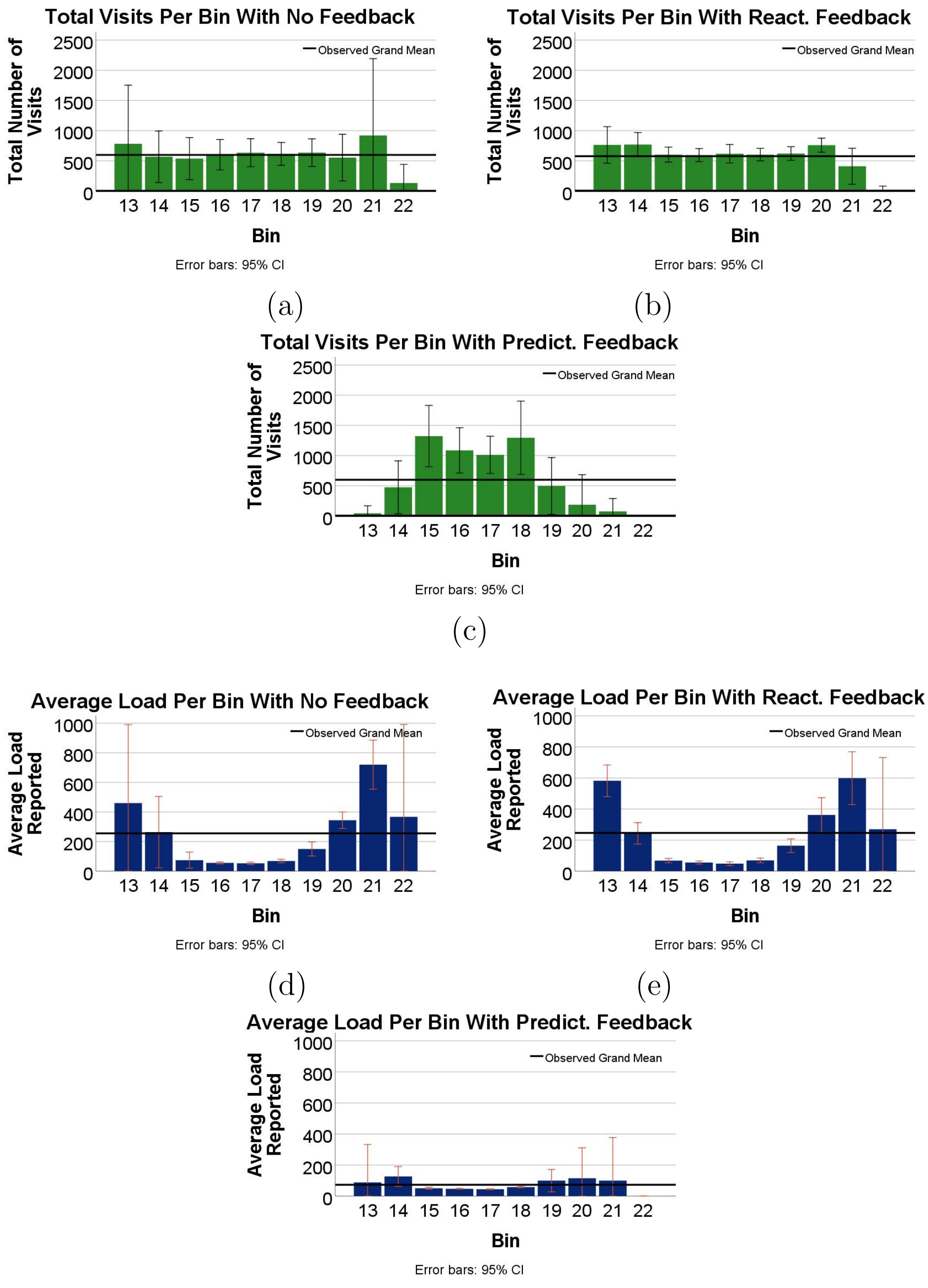
Figure 4.3: Key finding: the use of predictive feedback reduced the load (a measure of impact intensity) experienced by the system during use. The load shown is averaged across all five participants and the entire duration of each trial.

trail across all five participants, where again a one way repeated measures ANOVA was used for statistical analysis. Maulchy's W indicated sphericity could not be assumed (< 0.050). The Greenhouse-Geisser epsilon was < 0.75 (0.081), so this correction was used. The corrected F statistic was $F(2, 9) = 4.994$, which is above the critical value of 4.26, and significance was detected ($p = 0.030$). The $N = 5$ potentially interferes with some bins being significant, as the variance is high. Despite this however, significance is noted in bins 15–19. Significance was not found on the extreme end bins, 13 and 22, due to the high variance. When using reactive feedback, subjects were observed to contact both walls of the workspace with approximately even frequency as can be seen in Fig. 4.4(b), with the robot arm deflecting noticeably on both sides due to the contact. When predictive feedback was provided to the user, the robot arm was also observed to approach the two sides of the workspace symmetrically, but with much less or no visible deflection to the arm upon contact. The figures show increased visits in central regions of the workspace under predictive feedback compared to the other feedback modes.

The edges of the workspace show less average load for the predictive feedback case than the other feedback cases examined here. The relationship between the feedback type, position, and load can be seen in Fig. 4.4(d–f). The load is shown as an average across all 5 participants, using the visits per bin to average the load in each bin individually. As with previous analyses, one way repeated measures ANOVA was used to determine any statistical significance. Again, Maulchy's W indicated sphericity could not be assumed (< 0.050), but the Greenhouse-Geisser epsilon was again < 0.75 (0.088) so was again used to correct the results. The Greenhouse-Geisser correction returns an F statistic of $F(2,10) = 6.805$ compared to a critical value of 4.10, with $p = 0.010$. Examined on a bin by bin basis, significance was found in bins 13, 20, and 21. Bin 22 was not found have statistically different results between the feedback types. The figures show greater load on the extreme ends of travel for the no feedback and reactive feedback cases. This is the region where collisions with the barrier of the workspace would occur. The load that is incurred in this region is not seen in the predictive feedback case. The visitation frequencies in Fig. 4.4(a–c) appear to coincide with the lower load experienced by the system in the end region.

4.4 Discussion

Feedback is an important aspect of skilled control. As noted above, we defined the control of the robotic device to be successful and skilled if the load experienced by the device while moving near the border of the work area is low—the task objective given to our subjects during testing was to closely approach but not impact the walls of the workspace. With different forms of feedback or different settings, we expected a subject might never get near the wall (overly sensitive predictions, thresholds, or too much temporal extension), that they might do so with high variability (as when operating with minimal feedback), or that they might impact the wall consistently but forcefully if feedback comes too late (e.g., with overly delayed or reactive feedback). Our observations support these expectations.



(f)

Figure 4.4: Aggregate results for all five subjects showing (a–c) the frequency of visiting any given servo motor positional bin and (d–f) the average load in each bin as reported by the servo, using the frequency of visits to average.

When the machine-learned predictions about collisions were used in providing feedback to the user, the user was able to reduce the overall strain on the system. Figure 4.3 demonstrates the effect that different types of feedback had on skilled control of the robotic device. In the no feedback case, the load experienced by the device was large and variable (the maximum load that can be reported by the servos is 1024). The variance improved in the reactive feedback case, but while the overall load decreased, it was not enough improvement for a study with N=5 to find statistical significance. No matter how sensitive the threshold is to initiate the reactive feedback, the user must still perceive the feedback and act in an appropriate, timely way; load is inevitable since it is already occurring. The strain on the system can potentially be reduced via fast human reaction time—subject-specific reaction time is one possible source of the variance in Fig. 4.3.

The similarity in the means of the no feedback and reactive feedback cases has interesting implications. It seems that if we only cared to limit the load experienced by the system during operation that there is little reason to use reactive feedback, which would be a typical first solution, over no feedback. This has major implications as to the importance of feedback in the operation of prosthetic devices. The significance of the predictive feedback case is a little surprising for the small sample size. Increasing the sample size may begin to differentiate the no feedback and reactive feedback cases, but that there is significance in the predictive feedback case with such a small sample size suggests it's power. *The simple machine learning agent is capable of learning something that when communicated to the human user improved their performance according to at least one outcome measure.*

The source of the overall reduced load can be seen in how participants moved the arm differently using the different feedback sources. A more detailed indication of the motion of the device is illustrated in Fig. 4.4(a–c). In particular, the figures highlight differences in the feedback modes in the area of travel between the borders of the workspace (between bins 14 and 21). Specifically in bins 15 to 18 the bin-by-bin cumulative visits are shown to be higher for predictive feedback than the same measures using reactive or no feedback. Despite this, all three modes have similar grand means, as would be expected for the constant fixed velocity motion the participants used. The higher frequency of visits in the central region is an indication of successful operation, as it shows that the device spent a greater portion of the moving time during the trial transitioning from border to border rather than impacting the walls, under which condition the servo can move a small amount while the arm flexes. This observation is further supported by the differences that can be seen between the types of feedback at and beyond bins 13 and 21—the borders of the workspace. With predictive feedback, the user moved into the borders of the workspace less frequently (Fig. 4.4c)). The impact this had on the load can be seen in the bin by bin comparisons of Fig. 4.4(d–f). As a result of the system visiting the border cases less frequently the system experienced significantly less (or no) load in bins 13 and 21–22 (Fig. 4.4(c–f)), indicating less time spent under impact conditions or flexing of the physical device. The similarities in average load reported by the servo for the central bins, bins 15 to 18 is expected as these bins are where the arm would be moving steadily with no perturbations to cause changes in the load experienced. This also aligns with the area more frequently visited with predictive feedback. There is some discontinuity in the load and position values for bins 20-22 when compared to bins 13 and 14. This may be the result of the discretization of the encoder positions not being related to the physical workspace—the difference in load reported on the left and right sides may be because the physical workspace barrier fell in between two bins, causing the load from contact with the wall to divide between two locations as perceived by the system.

The lack of notable difference between the no feedback and reactive feedback cases is also seen in Fig. 4.4. Despite having no awareness of the robot arm during the no feedback trail, participants still navigated the full space between the boundaries of the workspace, although with much greater variance. For this low sample size, the way participants moved under the no feedback and reactive feedback cases can be seen to be similar. It took the addition of machine-learned predictions to minimize the load *and* improve participants ability to travel inside the workspace. This reinforces the importance good feedback solutions for device users; machine-learned predictions, we suggest, are one such solution.

4.4.1 Tuning, Training, and Adaptability

With predictive feedback and the settings described above, we qualitatively observed subjects stopping the robot arm's motion such that it made only light, unloaded contact with the wall. This level of contact could be modified by varying the learning parameters of the artificial intelligence, and parameters could be adjusted in a number of ways to achieve a number of outcomes. There is no one "correct" setting for sensitivity; instead, there are a number of possibilities for how the device can assist the user in achieving their objectives. Learning parameters could be tuned to provide feedback behaviour that duplicates that of the reactive feedback case. The converse is not true—reactive feedback is not capable of providing preemptive feedback about future events. Also, when using predictive feedback, we observed that the threshold for indicating an impending load could be made more sensitive than the equivalent reactive load, a lower threshold, without incurring false positives. The reactive threshold was set to the lowest value that would not be triggered by the normal noise of motion. As the learned predictions are mathematical expectations conditioned on servo position, they are not affected by spurious load variance or noise due to motion, as would be a purely reactive approach, and allowed us to set a much lower threshold for triggering feedback in the predictive case.

For clarity of assessment, the artificial intelligence system in this preliminary study only acquired and updated its predictive knowledge during a defined training period. In any machine learning setting with a fixed training period, variability in training can noticeably affect learning system performance, but should not affect fixed or reactive approaches. Differences to the training of the learning system or a slight shift in the experimental setup may have resulted in an earlier feedback prompt to this user in terms of absolute servo-motor position on one side of the workspace. Omissions during training or changes to the domain of use may be corrected or updated through the use of continuing or ongoing machine learning. This has been suggested in prior work (Pilarski, Dawson, Degris, Carey, et al., 2013), and is a natural way to robustly extend the present study. As learning is already done in a per-time-step, incremental way during training, there are no technical or algorithmic barriers to continuing the learning of feedback-related predictions during operational use. Specifically, off policy algorithms would allow the system to learn in real time during the trials. The issue with the current learning approach is that when the system is successful in avoiding collisions with the workspace, it “forgets” the workspace is there. Off-policy learning offers a solution to this issue, and will be tested in a subsequent study. While many offline or batch prediction learning methods could potentially be used to generate expectations for use in feedback (e.g., the work of Pulliam et al. (2011)), the continuing and computationally inexpensive nature of our chosen learning approach makes it well suited for use in a prosthetic environment (Pilarski, Dawson, Degris, Carey, et al., 2013). Our prediction learning approach is suitable for subject-specific, task-specific learning with no requirement for a priori domain knowledge; it is also well suited for adapting to ongoing changes in a task or a user’s behaviour during persistent, real-time use.

4.4.2 Feedback Modalities

As noted above, much work is being done to restore missing feedback to prosthesis users (Antfolk et al., 2013; Hebert et al., 2014; Schofield et al., 2014). Focus has been placed on restoring touch, including sensations such as pressure, texture, temperature, and even pain. A large body of this research has explored feedback using sensory substitution, wherein one sensation is replaced with another different sensation that the user must be trained to skillfully interpret; use of this approach is largely due to the physiological constraints of prosthetic human-machine interaction (Antfolk et al., 2013). Modality-matched feedback is also receiving growing attention; in matched feedback, sensations are restored either invasively or non-invasively to the natural or proxy locations that convey sensations of the lost or damaged biological system as closely as possible (Antfolk et al., 2013; Hebert et al., 2014; Schofield et al., 2014).

Our present study can be thought of as a form of substitution feedback—predictions about the electrical current drawn by the device during operation (perhaps thought of as the device's "pain" or motor fatigue) are communicated to the user via a vibratory buzzing sensation in order to prompt the user to take action to prevent it. This buzzing is not a natural sensation, and it is not communicated at an equivalent natural location on the user's body. What separates this choice from the usual form of sensory substitution is that fact that the information being transmitted from the user to the device is not a biological sensation—it is specific to the internal hardware of the device and encodes a prediction about future changes to that hardware. While communicating these anticipations is helpful to the successful operation of the device, it is not a natural thing for the user to feel; as with most substitution feedback, it takes training to interpret such a sensation (as noted in Hebert et al. (2014)). This training need was perhaps minimized for our participants because of the precedent in modern society to interpret the vibration of personal device as a prompt to act (e.g., cellphone vibration in response to a new text message).

However, our work should not be thought of solely in terms of sensory substitution. Our study is intended to be a small window into a larger area for research: the use of machine intelligence as a method for filtering, selecting, and *communicating* salient information about the internal state of a complex device. This communication can be thought of as a form of *transparency*, as used by Thomaz and Breazeal (2008). Communication of such non-biological knowledge to the device's user—e.g., prompts regarding a device's internal state, decisions, and anticipatory knowledge—promises to streamline human-machine interaction in many domains, and should be equally suited to feedback via both sensory substitution and modality-matched percepts.

4.4.3 Future Work

The results presented in this work are preliminary, and there is much room for further study in this area. The incremental learning algorithm used in this experiment was effective but monolithic. If a control-learning system were used in conjunction with the present prediction-learning algorithm, it may be possible for a device to adapt the timing and magnitude of its feedback to better suit its domain of use. For instance, the feedback threshold or level of temporal abstraction γ could be tuned on the basis of reward-like signals of approval or disapproval delivered by the user, using techniques from related work on the human training of machine learners (Knox & Stone, 2013; Pilarski et al., 2011; Thomaz & Breazeal, 2008); predictive load information could be communicated at distances from the collision which have been learned to be appropriate for a specific user and their task preferences. Exploration could also be done into how effective the predictive feedback is when it is learned in real time while the user is doing the task, rather than freezing learning during the trials. Further, as artificial intelligence use in artificial limbs becomes more prevalent, finding ways of communicating the actions that the system has learned to the user, rather than solely a predefined environmental signal, may help allow more control to pass to the prosthetic—the case of shared control and sliding-scale autonomy. Transparent communication between the operator and their device could be the keystone which allows an intelligent prosthetic and a human user to co-operate, combine processing power, and more effectively restore lost function.

4.5 Conclusions

Feedback is important to prosthetic limb control. While machine intelligence has been used to improve the interpretation of control signals given to a limb from the user, its use in modulating feedback is often overlooked. This article contributed a look at the potential value of predictions and machine learning in feedback to close the loop between a human and their artificial limb. To our knowledge, this is the first study investigating the use of real-time prediction learning in the feedback path of a human controlled robotic limb, and suggested the potential value of continuing this line of exploration.

When compared to strictly communicating momentary electrical load to the user, communicating a machine-learned forecast of the same load was found to decrease the load experienced by a robotic limb as a result of impacts with a workspace, and to increase the ability of our subjects to navigate the limb despite the absence of all other feedback. The increase in precision in terms of both position and load for the predictive feedback case over the no feedback case was dramatic, especially given the low subject pool. Additionally, the improvement in load minimization over purely reactive feedback was significant. Though preliminary, these results promise two related outcomes for the user of a prosthetic limb. First, we expect that increased communication from the device about its internal state and setting of use may allow the user more personalized and more trustworthy options for control. Over the long term, predictive feedback could therefore lead to greater acceptance and assimilation of the device as part of the user. Further, by creating a computational predictive forward copy of an action and communicating it to the user, operating an assistive device may become more precise. These expectations remain to be verified during the use of predictive feedback in real-life functional tasks.

Chapter 5

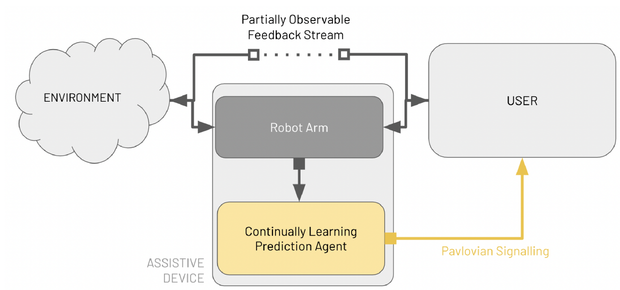
Continually Learned Pavlovian Signalling Without Forgetting for Human-in-the-Loop Robotic Control

Artificial limbs are sophisticated devices to assist people with tasks of daily living. Despite advanced robotic prostheses demonstrating similar motion capabilities to biological limbs, users report them difficult and non-intuitive to use. Providing more effective feedback from the device to the user has therefore become a topic of increased interest. In particular, prediction learning methods from the field of reinforcement learning—specifically, an approach termed Pavlovian signalling—have been proposed as one approach for better modulating feedback in prostheses since they can adapt during continuous use. One challenge identified in these learning methods is that they can forget previously learned predictions when a user begins to successfully act upon delivered feedback. The present work directly addresses this challenge, contributing new evidence on the impact of algorithmic choices, such as on- or off-policy methods and representation choices, on the Pavlovian signalling from a machine to a user during their control of a robotic arm. Two conditions of algorithmic differences were studied using different scenarios of controlling a robotic arm: an automated motion system and human participant piloting. Contrary to expectations, off-policy learning did not provide the expected solution to the forgetting problem. We instead identified beneficial properties of a look-ahead state representation that made existing approaches able to learn (and not forget) predictions in support of Pavlovian signalling. This work therefore contributes new insight into the challenges of providing learned predictive feedback from a prosthetic device, and demonstrates avenues for more dynamic signalling in future human-machine interactions.

5.1 Introduction

There are many sophisticated devices designed for people who have lost an upper limb to assist them in their daily lives. These devices, prosthetic arms and hands, can take the form of advanced robotic limbs capable of mimicking many, if not all, of the degrees of freedom of a biological limb. Despite the potential of these technologies because of difficulties in control and general use they are sometimes abandoned and therefore unable to fulfill their function (Biddiss & Chau, 2007a; Smail et al., 2021). A prosthetic limb is intended for frequent use and close collaboration with the human user; such devices are attached to the body and are intended to be considered a part of the user. The intended closeness of the connection is what makes this particular case of human-machine interaction interesting, and challenging.

Machine learning has been actively pursued for some time as a way of improving the control of prosthetic upper limbs in various ways, many of which are outlined by Castellini et al. (2014). Pattern recognition, for example, learns to associate patterns of muscle activation, often read by surface electromyography (EMG) electrodes, with motions of the prosthesis (Castellini et al., 2014; Scheme & Englehart, 2011). Because the connection between the control signal coming from the user and the motions of the device are learned, pattern recognition allows individually tailored solutions for users, which is a highly desirable trait in rehabilitation medicine (Castellini et al., 2014). These methods require training offline however, which means that if there is a change to the user's body or the way they generate control signals the system needs to be retrained to resume proper functionality. A more experimental technique is adaptive switching (Edwards, Dawson, et al., 2016), which learns in real-time



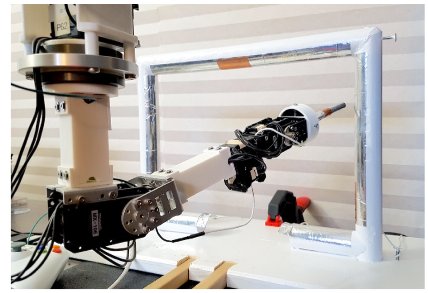


Figure 5.1: Schematic of the human-machine-environment interaction loop and photo of the physical system: a person interacts with a partially observable environment by way of an assistive machine, wherein their ability to perceive relevant decision-making information is limited in both time and space; as part of their interactions with the environment via the device, a continual prediction learning machine provides signals that allow timely action based on impending changes in the environment, with enough time for the person to process and act on these signals and their current state.

to re-order a list of joints that the user has to sequentially switch through to access different movements on a multi-joint prosthesis. This significantly reduces the amount of switching effort required compared to the more common approach of using a fixed list.

Machine learning methods like those used for adaptive switching have the potential to improve upper limb prostheses and rehabilitation in general. Specifically, temporal-difference (TD) learning approaches (Sutton, 1988) learn in real-time as they are used, and are able to approximate general value functions (GVFs (Sutton et al., 2011))—predictions of signals of interest from a robot, human, or environment (Pilarski, Dawson, Degris, Carey, et al., 2013). GVFs ability to be learned via TD in real-time gives them the potential to adapt and grow as the user does (Pilarski, Dawson, Degris, Carey, et al., 2013). This could allow users to develop their skills and abilities as they more naturally would, rather than having to learn and compensate around preset capabilities of the device they are using. The ability of reinforcement learning techniques to adapt also unlocks the potential for personalized user solutions to be found by the system, rather than having designers customize every device for every user (Castellini et al., 2014). While machine learning has been studied extensively for control improvement in prostheses, *its applications to* *providing feedback from the prosthesis to the user remain under explored.*

Feedback from a limb to the user, such as vibro-tactile or audible signals, can be an important aspect of control as well as user experience. Sensory feedback has been shown in some studies to improve user outcomes, as it closes the loop between the user's control and the device's execution. The comprehensive review by Schofield et al. (2014) covers many methods and modalities of providing feedback. Non-biological feedback, such as sound relating to direction of motion, can also help with the development of internal models, which are hypothesized as being what the central nervous system uses feedback to develop and are potentially responsible for the performance gains when feedback is available (Shehata et al., 2018a). Feedback of this type can be thought of less as feedback in the sense one might have from a biological limb and more along the lines of signalling between two agents. To date most of the explorations into providing feedback have used a fixed mapping between the sensor and the feedback device that does not change or adapt over time.

There are other domains where interactions between a human and a device have shown adaptive feedback to be helpful to the human (Crandall et al., 2018). These domains are beginning to include machine learning, and explore how it might signal a human user to assist in accomplishing a task the human would otherwise struggle with (Brenneis et al., 2022; Crandall et al., 2018). Early explorations of this kind used what has been recently termed as *Pavlovian signalling*[1]: a fixed signalling response (a token) emitted by one agent—or part of an agent—to another in response to the magnitude of a learned prediction about the environment or a task (Pilarski et al., 2022) (Fig. 5.1). Studies have been done where a machine agent is learning about a task and providing the human user with a signal they then use to successfully complete the task (Brenneis et al., 2022; A. S. R. Parker et al., 2019a; Pilarski et al., 2022). In this way the human and the machine can be seen as partners in accomplishing the task, and effective, reliable, communication is key.

An earlier study by A. S. R. Parker et al. (2019a) (Chapter 4 of this dis-

[1]This is outlined in greater detail as a contribition of Chapter 6, as well as Pilarski et al. (2022) and Butcher et al. (2022)

sertation) took a very simple approach to Pavlovian signalling in human machine interaction, specifically the task of navigating a robot arm in a confined workspace without touching the walls (A. S. R. Parker et al., 2019a). The paper used TD learning—an on-policy method, i.e., one that learns about the actual policy or behaviour being followed—to acquire GVF predictions, but found that learning had to be done during a training session and frozen during a trial. This was because when the machine agent was succeeding at helping the human avoid contacts with the work space, the machine agent was also avoiding those contacts and so unlearning them or "forgetting" they were there. While the results showed that the machine agent could learn something that assisted the human users in accomplishing the task, being unable to learn continually in real-time undermines the benefits of using continual learning techniques like TD learning. It was theorized that off-policy learning would overcome this limitation, i.e., that forgetting during continual learning could be mitigated by the use of learning algorithms that learn about a target policy different from the policy being followed, capturing "what if" style predictions (Sutton et al., 2011) about user behaviour even when the user is performing well in a task or setting. In this study, we extend prior work to contribute a detailed examination of hypothesized differences between on- and off-policy TD learning in the context of forgetting during Pavlovian signalling in ongoing robotic control. As this work is interested in exploring the application of machine reinforcement learning techniques in a rehabilitation setting in order to investigate solutions that adapt with the user over time to become individual, personalized, user experiences, the findings on the single human participant are not meant to be generalized. They are, however, worth bearing in mind as if something doesn't work for even one person, then this impacts the ability of the technique to become a personal solution for at least that individual (Mook, 1983).

5.2 Methods

5.2.1 Experimental Setup

The robot arm used for these experiments was the Bento Arm (Dawson et al., 2014). This arm is a 5-degree-of-freedom open-source 3D printed device that uses servo motors for actuation at the shoulder, elbow, wrist, and hand as shown in Fig. 5.1. For this work, the Bento Arm hand was replaced with a rod wrapped in conductive tape. The work space for the arm was a square interaction region constructed of wood and lined with conductive tape, which is also shown in Fig. 5.1, with these conductive elements all connected to analog inputs of an Arduino Leonardo. For human-participant interaction, the thumb-stick of an Xbox 360 controller was used to control the shoulder joint of the arm; built-in vibration of the controller was used to signal the participant that a trial was complete. Contact, or the prediction of contact, between the conductive rod and the workspace was signalled to the participant using a sound effect delivered over a set of noise-cancelling headphones with ambient white noise playing in them. The Bento Arm, Xbox 360 controller, and Arduino were connected to a laptop via USB. The laptop ran a modified version of the control software brachI/Oplexus (Dawson et al., 2020). This software package designed for use with the Bento Arm with built-in functionality for multiple control sources, as well as the ability to map control inputs to the Bento Arm.

5.2.2 Procedure

Two different settings were used to study the impacts of the algorithmic decisions made in this work: machine-machine control interactions and single case of human-in-the-loop control. For machine-machine interactions which provided a clear baseline for algorithmic comparison, we used a version of brachI/Oplexus that included a motion sequencer feature that allowed the researcher to program step-by-step arm motions, where upon reaching a set position the arm automatically transitions to the next movement in a list. This was modified for the autonomous trials to use the signal from the Arduino's analog inputs to move on to the next motion on the list. After a contact or prediction were observed to be above a threshold, the next motion was to back-off a pre-determined amount, and then move in the opposite direction. One complete motion involved moving to the right, backing off the pre-determined amount, moving to the left, then backing off again. For human-in-the-loop comparisons, a single human participant was recruited, who provided informed consent and the study was approved by the Research Ethics Board of the University of Alberta (Pro00085727).

The goal in both settings was to control the robot arm to move from side-to-side without making contact with the workspace edge, following the protocol established in A. S. R. Parker et al. (2019a). Both the human participant and automotion sequence moved the robot's shoulder actuator from left to right and back again within the workspace. Signalling tokens relating to contact with the workspace and predicted contact (described below) were provided as feedback in both the human and machine case (Fig. 5.1). For the human participant, the task was made partially observable for them by having them face away from the workspace while being sound isolated in noise-cancelling headphones with background white noise—they needed to rely completely on the feedback system to perceive the outcomes of their robot control commands. Similarly, the automotion controller was given no information about the workspace other than the tokens provided through feedback signalling. One trial consisted of 50 back-and-forth motions. Each learning technique was run for five trials by automotion and the human participant did three trials of each of the algorithms they tested.

5.2.3 Prediction Learning Algorithms

Algorithm 1 presents the main loop of the Pavlovian signalling process that occurred during user-robot interaction: observations from the robot and workspace were sampled and used to generate predictions; should these predictions rise above a threshold (or actual contact be made), a token was generated and sent to the user of the robot (automotion or human) that informed and/or changed their control of the system (c.f., Pilarski et al. (2022)). For the hu-

Algorithm 1 Pavlovian Signalling

```
set thresh for Pavlovian signalling
loop
    observe contact
    observe GVF prediction value
    if contact or prediction > thresh then
        generate signalling token
    if signalling token then
        if human trial then
            signal with audible feedback
        if automated motion then
            back-off motion
            change direction
        clear signalling token
    update all predictions
```

man participant, this took the form of an audio cue proactively alerting them to contact with the workspace; for the automotion case, this took the form of the robot arm backing slightly away and changing direction. All predictions were then updated according to the information obtained during the time step. The threshold empirically chosen (400) and constant across all conditions.

Predictions themselves were learned via temporal-difference (TD) reinforcement learning techniques (Sutton, 1988; Sutton & Barto, 2018; Sutton et al., 2011) to generate predictions about the expectation of future contact (A. S. R. Parker et al., 2019a). These learning algorithms learned from experienced data in real time during each trial; there was no training period before any of the trials. Three different implementations of GVF learning approaches were examined and compared for their ability to avoid contact events, and continue to avoid contact events over repeated motions: TD(λ), GTD(λ) (Sutton & Barto, 2018), and a variant we term look-ahead TD(λ). All three approaches used the same representation which consisted of the shoulder position, shoulder velocity, elbow position, and elbow velocity as reported by the servos responsible for motion of the Bento Arm. To construct the representation the allowed range of motion and velocities of the shoulder and elbow were normalized to produce a value between zero and one with zero being the left-most limit of

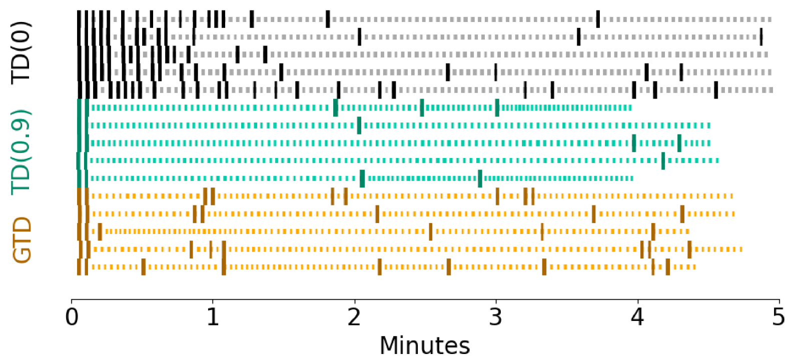
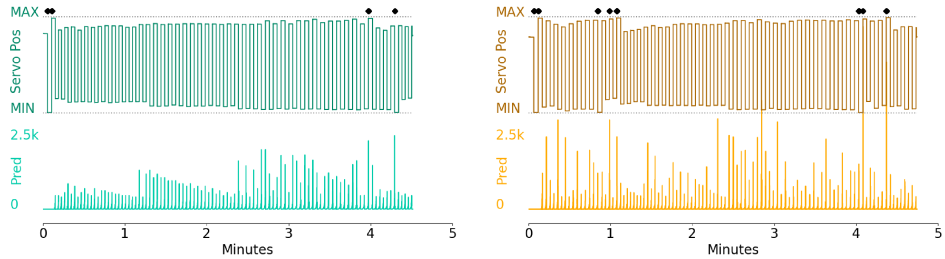


Figure 5.2: Contacts in the automatic motion case over time for all 5 trials for different algorithm settings: (black top) TD(0) (teal, middle) TD(0.9) (orange, bottom) GTD. Large ticks indicate contact, small ticks indicate generation of signalling tokens based on learned predictions.

the shoulder and bottom-most limit of the elbow, and one being the right-most limit of the shoulder and upper-most limit of the elbow. This was then divided into 32 evenly spaced sections, or bins, on two axes, one for position and one for velocity—an implementation of the function approximation method know as tile coding (Sutton & Barto, 2018).

Following on work by A. S. R. Parker et al. (2019a), the first GVF-learning used was the TD(λ) algorithm, which was tested using two different λ values, 0 and 0.9. In this algorithm, a temporal-difference error δ is computed based on an observed cumulant C, and the difference between the learned expected value for the present state (the inner product of a weight vector $\vec{w}$ and the function approximated state $\vec{x}(S)$ based on robot arm measurements S as noted above, discounted by γ) and the value of the last state $\vec{w}^{\intercal}\vec{x}(S_{Last})$. This error is used alongside a step size α and replacing eligibility traces $\vec{e}$ with decay rate $\lambda\gamma$ to update $\vec{w}$. A. S. R. Parker et al. (2019a) used a TD algorithm with $\lambda = 0$, so this was re-examined here to compare to using a λ value above zero.

As our second approach, we introduced a comparison with gradient temporal-difference learning (GTD) methods (Sutton & Barto, 2018). GTD(λ) is an off-policy algorithm; an off-policy algorithm learns about a target policy π_t from samples generated by a possibly unrelated behavior policy. The observed overlap between target and behaviour policies is specified by a parameter ρ,



(a) (b)

Figure 5.3: Contacts (black diamonds), motion (top) and prediction (bottom) for exemplar trials of (a) TD-lambda and (b) GTD to show the motion getting closer to the contact positions, marked by the dashed lines around the position, and the predictions diminishing, and spiking when a new position extreme is moved into.

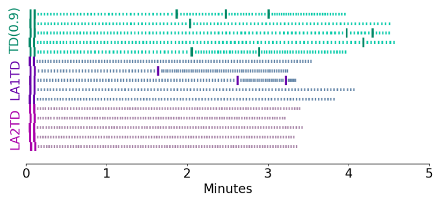
and is able to learn stably in this setting thanks to a separate set of weights $\vec{u}$ and learning rate β. For the off-policy comparison the system had two possible policies: one for each direction of motion of the shoulder actuator. For this study, ρ was always either zero or one as it is assumed the user was always acting on policy; ρ was one for the GVF responsible for learning the direction the servo was currently moving in, and zero for the other direction.

The final approach we explored was TD(λ) but utilizing a fixed look-ahead in the computation of the observation S, and is therefore termed "look-ahead TD". Rather than learning and acting on the same time step, this method learned on the current time step but used the prediction from a fixed number of position bins in the state space ahead in the same direction of motion to signal either the human participant or the automated motion controller. So the system should not be entering the states used trigger the warning. Such states could still be intentionally entered by a user, and the system would then adapt what it has learned about those states when they are entered.Algorithms for both TD(λ) and GTD(λ) are according to the standard implementation and can be found in the supplementary materials.

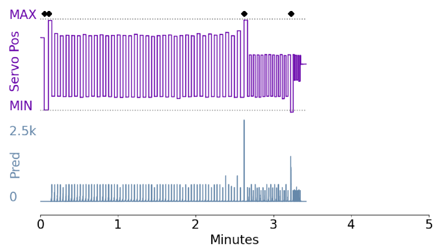
5.3 Results

Figure 5.2 shows the contact events, the dark lines, and the prediction going above the threshold value, the light lines. The three initial algorithms for the study, TD(0), TD(0.9) and GTD. TD(0) requires the most contacts to learn the boundaries of the workspace in the first place, and continues to make contact frequently. GTD initially learns fast, but continues to make frequent, though less frequent, contact. TD(0.9) also learns to avoid contact quickly, but makes contact much less frequently over the duration of each trial. In Fig. 5.3, we can see the contacts as they relate to the position of the shoulder for TD(0.9) and GTD. The position, shown in the top part of each plot is the location where the servos reported stopping, which typically occurred during changes of direction. Primarily of note is how the extremes of position can be seen expanding by bin sizes of position until contact, the location of which is shown by the dashed envelope around the position. After a contact is made the motion extremes then retract and the prediction, shown on the bottom part of each plot spikes. This spike is also visible when the position extremes expand, and is likely the result of visiting a new bin that has already learned about contact and not been visited in recent motions. This is more shown in more detail the supplementary material Fig. 5.6, where the prediction can be seen diminishing over repeated motions until it drops below the threshold and contact occurs.

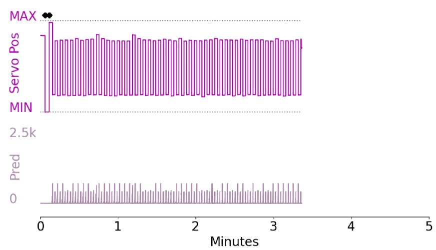
TD(0.9) compared with look-ahead TD is shown in Fig. 5.4. Figure 5.4a specifically compares contacts over the duration of trials between TD(0.9) alone and TD(0.9) with a look-ahead TD of one and two bins. The one bin look-ahead TD did occasionally make contact, which should not be possible since the predictions on the timesteps being used to signal are not being updated on the same timestep they are used. It can be seen in Fig. 5.4b that the one bin look-ahead TD prediction does not diminish, but a contact happens anyway. This causes the prediction to increase and motion to restrict further. This happens again on the other extreme of motion. Figure 5.4c shows the prediction and motion for two bin look-ahead TD. We can still see the



(a)

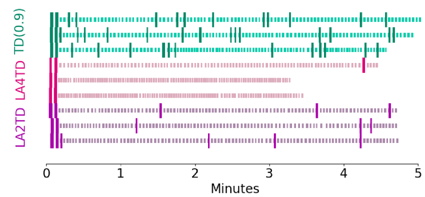


(b)

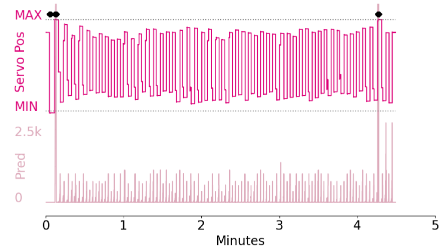


(c)

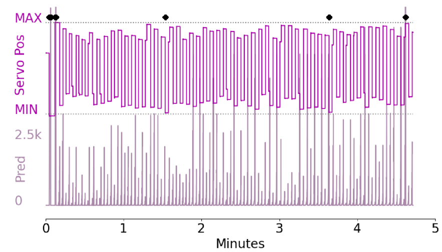
Figure 5.4: (a) Contacts as in Fig. 5.2, but for TD(0.9), 1 bin look-ahead (LA1TD), and 2 bin look-ahead (LA2TD). Contacts, position, and predictions for (b) LA1TD and (c) LA2TD.



(a)



(b)



(c)

Figure 5.5: Participant data for (a) contacts as Fig. 5.2 all three cases tested with the human participant, and contacts, motion, and prediction(b) LA4TD and (c) LA2TD.

shoulder position at reported motion stops not being perfectly steady as we would expect from robotic motion, but drifting slightly in both extremes. The additional bin prevents these drifts from making contact.

Human participant data is show in Fig. 5.5. The contacts over the trials for TD(0.9) shown in Fig. 5.5a show that TD(0.9) made repeated contacts after over the course of the trials. TD(0.9) with a two bin look ahead also struggled. This is likely a result of additional delays in the human-machine system from the sound used to signal the human having a delay in its onset and the reaction time of the participant. Figure 5.5b shows the predictions for TD(0.9) with the look ahead increased to four bins remains high over the course of the trial, but a single contact was still made in the 50 motions of the trial despite this. The motion that this happened on is one straight shot; it doesn't have any of the pauses before the extreme that some of the motions exhibit. The two bin look-ahead TD shown in Fig. 5.5c shows more passing through the bin where the signal is initially triggered by the prediction and into bins beyond—despite the prediction at the extremes of motion being high, contacts are made.

5.4 Discussion

TD Compared to GTD: In a previous study it was found that a machine learning agent using reinforcement learning techniques can learn something that is of value to the more sophisticated sensorimotor decision making of a human user (A. S. R. Parker et al., 2019a). In that study, however, one of the primary advantages of reinforcement learning techniques was not utilized in the trials—the system was not learning while the task was being done. This was because there is a particular challenge involved when success on the task involves no longer encountering the event in the signal space that is trying to be avoided. It was proposed in A. S. R. Parker et al. (2019a) that off-policy learning would overcome this. That proposal was the first algorithm examined in our present study. Surprisingly, Fig. 5.2 clearly shows that the off-policy learning algorithm, GTD in particular, did not overcome the forgetting seen in A. S. R. Parker et al. (2019a) The closest algorithm to that of the original study used here was TD(0), and Fig. 5.2 does show how this method struggles. However it is also clear that GTD did *worse* than TD(0.9). This was with robotic auto-motion, which is relatively consistent when compared to a human interacting with the device. This indicates that a human user would not find the expected success using an off-policy algorithm while the agent is learning online in real time as is the goal.

In order to determine what is happening further exploration of the data was required. Figure 5.3 suggests some of the challenges faced by GTD. It seems that part of what was supposed to be the benefit of the off-policy algorithm becomes a challenge; when the system motion stops or changes direction the observations do not impact learning about motion in the direction it was previously travelling. In this case, the agent gets less experience on each motion and learns more slowly. The pattern of diminishing predictions and increasing extremes of position until contact is visible here as well but occur more frequently. This strongly suggests that the benefit we were expecting to see from off-policy learning, not forgetting about where contacts were made when it no longer encountered them, is being at least in part overridden by the lower number of samples seen by each GVF as they are learned. more detail of the contact events can be seen in supplementary material Fig. 5.6. In particular closer examination of the predictions shows the prediction gradually diminishing over repeated motions in both directions until it is below the threshold and contact is made.

It is important that the machine learning agent tasked with assisting the human does not "forget" about a feature when success involves avoiding that feature. This places the machine agent in a position where it is at best not providing anything useful to the human other than perhaps some initial priming of the human's expectations of a task, after which the human must ignore the signals from the agent. At worst the agent is damaging the interaction between the human and the device by being unreliable and untrustworthy (Brenneis et al., 2022; Schofield et al., 2021). At the same time, the machine agent should be able to forget when it is appropriate to do so, or it risks becoming specialized to a very specific set of circumstances and is not fulfilling a primary advantage of reinforcement learning methods in continual human-machine interaction of real-time adaptability. If the human chooses to ignore a signal from the machine agent, and the workspace shape had changed so that contact was further out, the system should be able to adapt to this change in how the human is doing the task. For these reasons it is worthwhile to pursue reinforcement learning methods to improve human-machine interaction.

TD Compared to Look-Ahead TD: There are several challenges to this task that necessitated the use of a simple domain to attempt to isolate the effects algorithmic differences on learning to avoid some part of a workspace or task, and doing so to reliably generate predictions to aid this avoidance. The avoidance itself poses a challenge for the TD learners. Since these algorithms learn from experience, when they no longer experience something they not only stop learning about it but the experience that is avoided is replaced with more recent experiences. GTD, an off-policy algorithm, was thought to be the solution to this problem, however it encounters challenges as well. It is very possible that some of these challenges are based in the nature of physical systems; the play in the servos, inertia from the motion, and time synchronicity challenges that arise in part as a result of running each part of the system as fast as it is able in order to not miss events of interest. It should be noted that the servos used in the Bento Arm are durable, high precision servos, so while there may be hardware that would assist in overcoming the physical challenges the expense becomes significant. Therefore, instead of physical modifications to the arm, we changed how the predictions were being accessed. Until now, the timestep that was being learned on was also being used to generate the prediction token to instruct the system to change directions in order to avoid contact. This was the purpose of the look-ahead TD method; by learning separately from acting in the state space, this method should make it impossible for the predictions being used to signal to diminish since they are not physically visited by the system.

TD(0.9) was used as the learning algorithm with look-ahead TD, as it learned quickly. Of particular note in Fig. 5.4a, the two bin look-ahead made no contacts after an initial one per side. The one bin look-ahead did, on occasion make contact after initial contacts. It is clear in Fig. 5.4c that the motion is stable; the extremes of position don't change substantially for the most part, except in a few places where motion stopped in a further bin, but no contact was made. The predictions remain undiminished, at least in their peaks. Interestingly, upon close inspection what happens to the predictions is they decay in the bins being visited, but not the ones being used. This gives them the appearance of a spike, rather than a gradual increase. The appearance of the motion and prediction in Fig. 5.4b and Fig. 5.4c is very similar until the contacts, at which point Fig. 5.4b subsequent motions in that direction stay further from the previous extreme as the prediction value of all relevant bins, as configured in the GVFs, increases. It is interesting that look-ahead of any number made contact at all. This strengthens the implication that there are more issues facing the system than unlearning the events it stops observing. It is likely the contacts made by the one bin look-ahead can be attributed to previously mentioned physical challenges of the system. That there seems to be no regularity to the contacts caused by this is an extra challenge, that seems as though it can be avoided by increasing the number of bins of look-ahead.

Human Participant Outcome: In order to begin to explore the impact that these algorithmic choices have on the ability of the system to adapt and become an individualized solution in human-machine interaction, a case study was done with a single participant using TD(0.9) and look ahead. The originally pursued off-policy GTD was not included in the human controlled setting as it's problems would only be exacerbated by human use. Figure 5.5a shows that TD(0.9) does not fare as well at reducing the number of contacts with the human participant as it did with automated motion. The two bin look-ahead had fewer repeated contacts, but with the ideal goal being zero contacts after the initial motions it is not performing adequately. Even increasing the look-ahead to four bins resulted in a single contact. This contact happened on a trial that took longer to complete than the other two, which implies that it was covering more distance between extremes and thus was closer to contact the entire trial. The motion of the human shown in Figs. 5.5b and 5.5c is, understandably, less regular than that of automated motion. This, coupled with delays in reaction caused by the participants reaction time and a software delay in the onset of the sound signalling the participant of an event, increased the difficulty of the task for the learning agent. Even with four bins of look-ahead there was a contact late into one of the trials. The look-ahead technique, if nothing else, adds an extra parameter that can be adjusted to assist the success of the learning agent. While look-ahead TD should completely prevent forgetting, even using on-policy algorithms such as TD(λ), the physical factors of the system make it not so simple. If the number of look-ahead bins is not correctly set, over extended periods of time the slippage into adjacent, closer, bins would still lead to forgetting. It is possible, in fact likely, that different representations and even reward functions could be used to accomplish this specific task. However, the choice of representation and cumulant here were made because of they do not use designer knowledge of the task, but rather are grounded in the construction of the robot, specifically the information available from the servos, and signals from the environment alone. By grounding the representation and cumulants in the construction of the robot in this way, the system must learn about the environment and task by way of it's own motions and signals. This is important because, if successful, it allows the system to learn about the world in terms that will always be available to it even with different users, tasks, and environments. The use of a single participant for the human portion of these trials is a significant limitation to statements about the generality of this work. However, in this domain, the driving rational of using RL techniques is that they can adapt to an individual user, therefore any single participant where this is found to be not true (as in the present experiment) is worth bearing in mind for future work. For the interested reader, a more advanced proposal for learning predictive look-ahead can be found in the supplementary material (Fig. 5.7).

5.5 Conclusions

In this work we investigated a clear prosthetics-motivated example of one of many settings where agents interact with other agents in uncertain via signals that they adapt in real time and through ongoing experience. In addition to its specific contributions to improving feedback from prosthetic limbs by demonstrating the use of Pavlovian signalling in a human-robot arm interaction, this paper provides concrete evidence on how algorithmic differences impact continual temporal-difference learning of approximate general value functions used in feedback and signalling; this work contributed new insight into the importance of on- and off-policy learning choices, predictive representations of state, and function approximation, and how these factors act differently on real-world platforms with and without a human in the loop. As this work is intended to further methods for reinforcement learning techniques to adapt to individual users and provide specific solutions, the findings on the inability of the selected off-policy method to reliably provide signals the human could use is worth keeping in mind moving forward. We therefore expect these findings to support the development of next-generation neuroprostheses and other assistive technology, and more broadly a range of diverse applications where multiagent interaction occurs in complex domains in concert with or as mediated by prediction learning machines that continually learn during ongoing human-in-the-loop interaction.

5.6 Additional Materials

5.6.1 Temporal-Difference Learning Algorithms

Algorithm 2 TD(λ) Update

set $\alpha \leftarrow 0.1, \gamma \leftarrow 0.9, \lambda \leftarrow 0$ or 0.9
init $\vec{w}, S, S_{Last}, \vec{e} \leftarrow 0$
on update call
  observe C, S
  $\delta \leftarrow C + \gamma \vec{w}^{\mathsf{T}} \vec{x}(S) - \vec{w}^{\mathsf{T}} \vec{x}(S_{Last})$
  $\vec{e} \leftarrow \mathsf{min}(1, \vec{x}(S_{Last}) + \gamma \lambda \vec{e})$
  $\vec{w} \leftarrow \vec{w} + \alpha \delta \vec{e}$
  $S_{Last} \leftarrow S$

Algorithm 3 GTD(λ) Update

set $\alpha \leftarrow 0.2, \gamma \leftarrow 0.9, \lambda \leftarrow 0.9, \beta \leftarrow 0.01$
init $\vec{w}, \vec{u}, S, S_{Last}, \vec{e} \leftarrow 0$
on update call
  observe C, S
  $\rho \leftarrow 0$
  if behaviour aligns with π_t **then**
    $\rho \leftarrow 1$
  $\delta \leftarrow C + \gamma \vec{w}^{\mathsf{T}} \vec{x}(S) - \vec{w}^{\mathsf{T}} \vec{x}(S_{Last})$
  $\vec{e} \leftarrow \rho \cdot \mathsf{min}(1, \vec{x}(S_{Last}) + \gamma \lambda \vec{e})$
  $\vec{w} \leftarrow \vec{w} + \alpha[\delta \vec{e} - \gamma(1-\lambda)(\vec{e}^{\mathsf{T}} \vec{u}) \vec{x}(S)]$
  $\vec{u} \leftarrow \vec{u} + \beta[\delta \vec{e} - (\vec{x}(S_{Last})^{\mathsf{T}} \vec{u}) \vec{x}(S_{Last})]$
  $S_{Last} \leftarrow S$

5.7 Extended Results

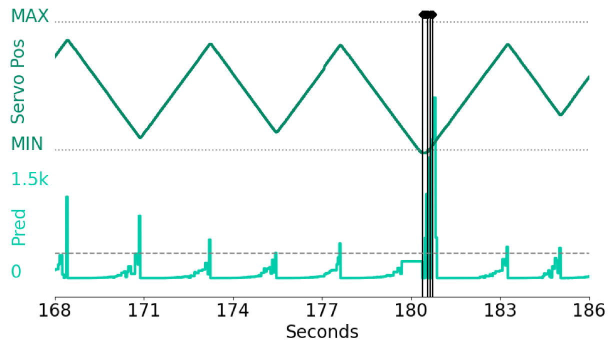


Figure 5.6: Contacts (black), motion (top), and prediction (bottom) for TD-lambda zoomed in to show the point where contact occurs. Dashed line across the prediction is the threshold for token generation. Here the position is as reported by the servo on each timestep.

5.7.1 Future Directions: Learning the Predictive Look-ahead

Three different look-ahead values were used in this study, which were manually chosen. One of those values worked with automated motion, but in the human trials further look-ahead, or perhaps adjusting the learning parameters, is required. It is not out of the question to think that different participants, or methods of delivery of the signal to the participant, would necessitate different

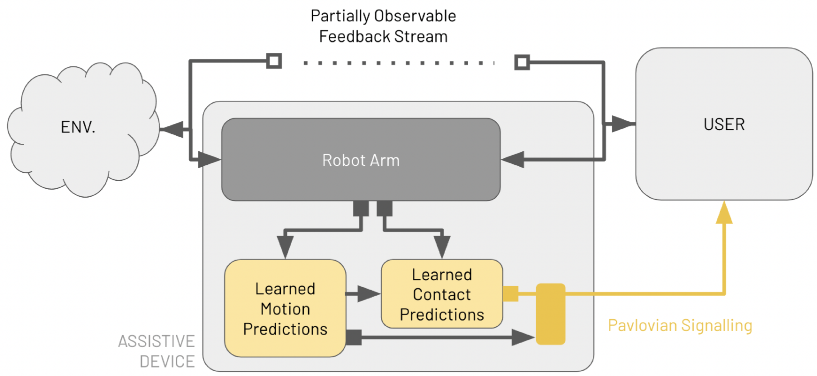


Figure 5.7: Schematic of more advanced learned predictive architecture. Multiple prediction learners responsible for different parts of the task and the controller combines them for use in signalling the user.

amounts of look-ahead. In the vein of having the system learn about the world in terms referential to itself rather than having the designer impart specific task, or partner, knowledge upon it, a system that also learns the amount of look-ahead required may perform better and more generally. This could be done with multiple or layered predictions. An example of an architecture for such an interaction of prediction learners can be found in Figure n the supplementary material. In this proposed architecture, the robot interacting with the environment to generate predictions about contact remains the same. Following methods from Sherstan et al. (2015), a second set of GFVs would learn about the way the arm is moving, and make predictions about what position the arm is expected to be in. These predictions can be combined in the controller to determine if the predicted position is in a predicted contact area, and this is in turn used to signal the user. This could enable a learned value in place of what here is a designer-specified look ahead, and increase the success of the learning agent in assisting the human to avoid all contacts in future.

Chapter 6

Assessing Human Interaction in Virtual Reality With Continually Learning Prediction Agents Based on Reinforcement Learning Algorithms: A Pilot Study

Artificial intelligence systems increasingly involve continual learning to enable flexibility in general situations that are not encountered during system training. Human interaction with autonomous systems is broadly studied, but research has hitherto under-explored interactions that occur while the system is actively learning, and can noticeably change its behaviour in minutes. In this pilot study, we investigate how the interaction between a human and a continually learning prediction agent develops as the agent develops competency. Additionally, we compare two different agent architectures to assess how representational choices in agent design affect the human-agent interaction. We develop a virtual reality environment and a time-based prediction task wherein learned predictions from a reinforcement learning (RL) algorithm augment human predictions. We assess how a participant's performance and behaviour in this task differs across agent types, using both quantitative and qualitative analyses. Our findings suggest that human trust of the system may be influenced by early interactions with the agent, and that trust in turn affects strategic behaviour, but limitations of the pilot study rule out any conclusive statement. We identify trust as a key feature of interaction to focus on when considering RL-based technologies, and make several recommendations for modification to this study in preparation for a larger-scale investigation. A video summary of this paper can be found at https://youtu.be/oVYJdnBqTwQ.

6.1 Introduction

Technology increasingly relies on learning to improve performance. Autonomous systems that continually support human users are expected to soon need to learn continually even during use in order to perform well in their general and changing settings of interest (e.g., assistive technologies; (Dalrymple et al., 2020; Edwards, Hebert, et al., 2016; Pilarski, Dawson, Degris, Carey, et al., 2013; Sherstan, 2015)). Humans that use these systems will interact with technology that has a constantly changing level of competency and reliability, but the ramifications of a system's continual learning on human behaviour and human-machine interaction are not well understood. Here, we begin to investigate this interaction by considering a human involved in a timekeeping task, partnered with a machine agent that learns from a blank slate to help the human. In general terms, an intelligent machine of this sort is able to make predictions about the dynamics of the world that a human partner either cannot or does not want to compute on their own (possibly due to the difficulty or time-consuming nature of the computation, (Risko & Gilbert, 2016), or the human's inability to sense relevant information). In order to convey the benefit of these predictions, it is natural that machine agents must be able to communicate information to the human (Crandall et al., 2018; Lazaridou & Baroni, 2020); learned communication is built upon relationships, and relationships can be built up through interaction (Scott-Phillips, 2014; Scott-Phillips et al., 2009). Take for example your interactions with a wristwatch: if up to now it conveyed accurate time-information to you, you would have every reason to continue trusting its information the next time you consulted it. If its degree of competency degraded for some reason, and the information communicated were incorrect, you would quickly lose trust in the device and look to other sources for the information you need. Now suppose instead that your wristwatch was not designed to convey regular time intervals, but instead predict the onset of stochastically reoccurring events. How would your interactions with your wristwatch be affected by the fact that the device must continually learn, update, and change its behaviour *while you are using it*?

In this paper we describe a pilot human-agent interaction study, investigating how time-based prediction agents can augment human predictions, and how the relationship between the human and agent develops as the agent develops competency. Specifically, we describe and compare two simple agents that learn to predict future stimuli using *general value functions* ((Sutton et al., 2011); from the field of reinforcement learning), and communicate those predictions to a human participant using *Pavlovian control* (which maps predictions to a small set of actions, (Modayil & Sutton, 2014)). We introduce a virtual reality (VR) task designed to assess human-agent interaction in a time-interval prediction task. VR is a compelling tool for human-computer interaction (HCI) research because it is immersive, allows flexibility and control for experiment parameters, and enables measurement of human movements which provide a window into decision-making (Gallivan & Chapman, 2014). VR also requires physical participation—due to COVID-19, we were unable to recruit external participants. We took this as an opportunity to engage in the present work: a thorough preliminary investigation in search of interesting trends and themes that might deserve careful investigation with respect to continual learning during human-machine interaction.

6.2 Background

6.2.1 Prior Work on Human Interaction with Learning Systems

Human interaction research regarding autonomous systems spans from early software interfaces for email and calendar applications (Maes, 1995) to more complex and personal domains such as the control of prosthetic limbs (Pilarski, Dawson, Degris, Carey, et al., 2013; Schofield et al., 2021), and has included a wide variety of automation techniques. Automation has traditionally been hand engineered to provide reliable performance, and therefore reliable human interaction. More recent machine learning systems are typically pre-trained before deployment, after which their parameters remain fixed. Research specifically involving interaction with continually learning algorithms has hitherto mainly focused on investigating agent learning dynamics using human interaction as part of the learning signal (Li et al., 2019). Autonomous systems that learn from human signals are important technologies, but system learning dynamics are inherently intertwined with interaction dynamics. Amershi et al. (2014) convincingly argue the case for separating human interaction from agent learning in order to study "how people actually interact—and want to interact—with learning systems". They describe case studies involving people interacting with machine learning systems, and by specifically focusing on the human component of the interaction, they are able to discover novel modes of interaction, unforeseen obstacles, and unspoken assumptions about machine learners. A meta-review of factors that affect trust in human-robot interaction (Hancock et al., 2011) suggests that system-specific factors such as behaviour, predictability, and failure rates greatly affect human trust in autonomous systems, justifying a system-specific investigation of human interaction with RL-based systems as distinct from machine learning systems. The particular feature of the RL algorithm that we study that distinguishes it from other autonomous systems and warrants direct investigation is continual learning during the course of a task, and the effect that will have on human interaction.

6.2.2 General Value Functions

Reinforcement learning (Sutton & Barto, 2018) is a class of machine learning methods wherein an agent learns to predict future values through a process of trial-and-error. The value of a state (a prediction of how much reward can be expected in the future from that state) is learned by incremental updates to a value function $v_\gamma(s)$ for state s. The discounting factor γ corresponds to the horizon of the prediction, and is typically between 0 (for next-step predictions) and 1 (for an infinite horizon). By substituting any signal of interest (called a *cumulant*, C) in place of the reward, the value function becomes a *general value function* (GVF) $v_{\gamma,C}(s)$ which predicts the discounted sum of the future cumulant (Modayil et al., 2014; Sutton et al., 2011). Informally, a GVF represents a prediction question: *what will be the total accumulated value of some signal of interest over the next specified time window?* Equation 6.1 gives the formal GVF definition, for a simple fixed-γ on-policy prediction formulation.

$$v_{\gamma,C}(s) = \mathbb{E}\left\{ \sum_{k=0}^{\infty} \gamma^k C_{t+k+1} \,\middle|\, S_t = s \right\} \tag{6.1}$$

In practice, an agent learns to approximate the above value by interacting with a stream of states and corresponding cumulants. Let $x(s) \in \mathbb{R}^d$ be a feature vector summarizing the state s. We approximate the value by $v_{\gamma,C}(s) \approx w_t^\intercal x(s)$, where $w_t \in \mathbb{R}^d$ is the weight vector at time t. We use the TD(λ) algorithm to update w_t on each time step:

$$\begin{aligned}
e_t &= e_{t-1} + x(S_t) \\
\delta_t &= C_{t+1} + \gamma w_t^\intercal x(S_{t+1}) - w_t^\intercal x(S_t) \\
w_{t+1} &= w_t + \alpha \delta_t e_t \\
e_t &= \gamma \lambda e_t,
\end{aligned}$$

where α is a scalar learning rate and $e \in \mathbb{R}^d$ is an exponentially decaying memory of previous feature activations.

6.2.3 Pavlovian Control

Inspired by prediction learning for reflexive control in animals (Kehoe & Macrae, 2002), the term *Pavlovian control* as used here refers to the use of a GVF to predict an external stimulus, coupled with a fixed reflexive control policy dependent on that prediction (c.f., (Dalrymple et al., 2020; Modayil & Sutton, 2014)). A simple Pavlovian control policy emits a discrete action a_1 when the GVF prediction of the external stimulus is below a certain threshold τ, and emits a discrete action a_2 otherwise (where, importantly, that action may be a communication signal (A. S. R. Parker et al., 2019a; Pilarski et al., 2022)). The precise setting of τ for a useful policy depends on the timescale and stochasticity of the prediction, as well as the amount of advance notice needed before the external stimulus in order to take action.

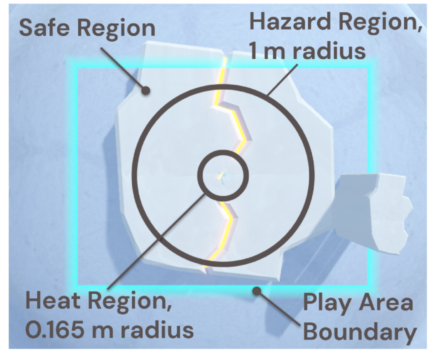
6.3 Experimental Methods

Our experiment situates a human participant in a virtual reality (VR) environment we call *the Frost Hollow*, wherein they must keep track of an external event that occurs on a roughly periodic schedule (c.f., Rafiee et al. (2023)). They are paired with a machine agent that uses a GVF to predict the onset of this event, and cues the human when its prediction exceeds a threshold. We look at task performance, behavioural differences, and qualitative notes to compare two agent architectures against the control condition where the participant completes the task with no agent assistance.

6.3.1 Virtual Reality Environment

The premise of the Frost Hollow task is that the participant stands in a "warm" center region of the environment (radius 0.165 m, participant position reported by the headset) to slowly gain heat, and must periodically dodge out of a hazard region (radius 1 m) when the wind blows to avoid losing heat. When standing in the center region, a heat gauge visible to the participant fills from 0.0 to 5.0 at a rate of 0.1875 heat/second (26.67 seconds to fill the gauge); when the gauge is full, the participant can raise one of their VR controllers above their headset to cache the heat gained as a point (one unit of game reward). When hit by the hazard, the participant loses 25 heat/second, so any hit longer than 200 ms will deplete the gauge. Cached points are not lost. Our VR environment (depicted in Figure 6.1) was implemented in Unity 2019.2.17f1 (Unity Technologies, USA) with Steam VR (Valve Corporation, USA) and presented to the participant via a Valve Index headset and two handheld controllers (Valve Corporation, USA; headset max render rate of

(a) Participant view

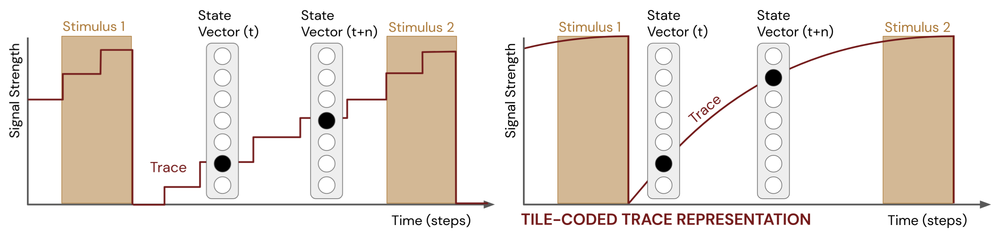


(b) Top-down view, annotated

Figure 6.1: Depiction of the virtual reality environment.

Figure 6.2: A virtual-reality environment made to emulate a beautifully snowy frost hollow. Snow covers the ground and surrounding trees. A ice-sculpture heat-gauge is visible left of center. In the right figure, a top-town view is presented, annotating a center circle with radius 0.165 m corresponding to the heat generation region, and a larger 1 m radius circle corresponding to the hazard region.

144 Hz) at a base Unity time step length of approximately 8ms (VR protocol follows from prior work (Pilarski et al., 2019)). Detailed descriptions of the audiovisual presentation of the environment are available from Pilarski et al. (2022). We studied three inter-stimulus-interval (ISI) conditions between the hazard pulses: fixed, drifting, and random. The base ISI was set to 20s (as measured from the falling edge of the pulse to the next pulse's falling edge); the hazard pulse duration was 4s in all conditions. For the random condition, the inactive portion of each ISI was varied uniformly by [-4s, 6s] between 12s and 22s in length. For the drift condition, the inactive portion of each ISI was shifted by a uniform random amount between [-2s, 2s] from the previous duration, with all ISI durations outside [12s, 22s] cropped to the extremes of the range. When the hazard pulse was active, the participant's left hand-held controller vibrated, and a visual bloom stimulus was presented on hazard contact; communication from the agent was presented as vibration in the right hand-held controller (c.f., (Edwards, Hebert, et al., 2016; A. S. R. Parker et al., 2019a)).



(a) Bit Cascade Representation (b) Tile-Coded Trace Representation

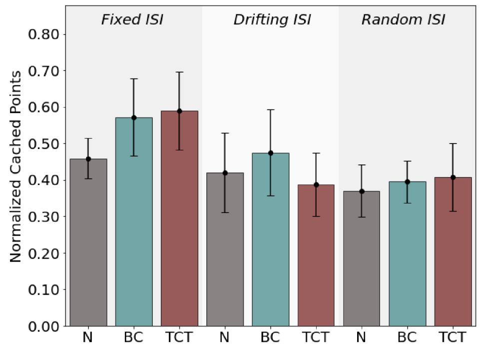
Figure 6.3: Representations of time used in this experiment. Time (state) is represented as a one-hot vector of features which activate according to a trace function which resets at the falling edge of each stimulus pulse.

6.3.2 Agent Architectures

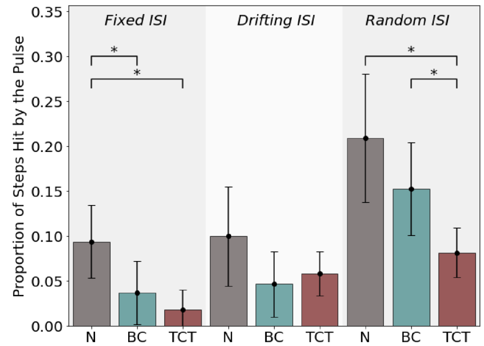
Two agent architectures are compared which differ only in the way that they represent the passage of time between stimuli: a *bit cascade* (BC) representation, and a *tile-coded trace* (TCT) representation (depicted visually in Figure 6.3). The decision to vary agent representations of time rather than other agent parameters is motivated by a larger study of Pavlovian signalling (Pilarski et al., 2022) for which this work plays a supporting role. These representations of time were motivated by and modeled after biological models of time-keeping in animal brains (Paton & Buonomano, 2018). The BC representation is modelled after population clocks (sequentially firing chains of cells), while the TCT representation is informed by ramping models (changes in the tonic firing rate of cells or cell populations). The bit cascade representation involves a one-hot vector of 40 features which activate sequentially, with each feature being active for 0.5s. The tile-coded trace representation also involves a one-hot vector of 40 features which activate sequentially, but the activation timing for each feature is prescribed by an exponential decay trace with a per-step decay rate of 0.998. Both the BC and TCT representations restart their sequence (i.e. the first feature is active) immediately after the hazard pulse deactivates. The timing parameters for both representations were set so that both used roughly the same number of feature bins when presented with an ISI of 20s. Learning parameters were empirically determined for an acceptable learning speed over a 5 minute trial time, resulting in $\alpha = 0.1$, $\lambda = 0.99$, and $\gamma = 0.99$. The Pavlovian control threshold τ was also empirically determined to give adequate lead-time for a human participant in advance of a pulse after learning had converged, resulting in $\tau = 10$ for both agents. The fixed control policy was set such that the agent vibrated the participant's handheld controller when its prediction rose above τ, and did not vibrate when below τ. Agent-learned weights were discarded and re-initialized to zero between trials so the agent learned from a blank slate for each trial.

6.3.3 Experiment and Analysis Protocol

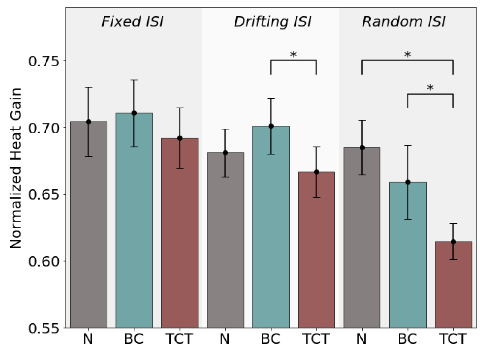
We engaged a single participant (male, age 40, no history of sensorimotor impairments) *who was also a member of the study team* due to COVID-19 constraints (see Section 6.8), and followed our approved human research ethics protocol. This participant had a deep understanding of the task and dynamics, but was not practiced with the particular conditions. The study followed a within-participant 3 (ISI type) x 3 (agent type) design; experimentation took place over the course of ten sessions, each consisting of nine trials that were five minutes long (one for each pairwise combination of [fixed ISI, drifting ISI, random ISI] and [no agent, TCT agent, BC agent]). Trial order was randomized and blinded to the participant, and the initial ISI duration for the fixed and drifting conditions was randomized to further obfuscate the conditions. Each individual session was conducted in roughly one hour, with small breaks between each of the trials for the participant to remove the headset and drink water or write qualitative notes. Sessions were spread over a one month collection period, with one or two sessions per day on data collection days. This protocol was found to be slightly physically fatiguing and moderately cognitively fatiguing, depending on the trial. To avoid injecting biases into the analysis, the team member who acted as participant for the study did not re-engage with the study until both qualitative and quantitative analyses were completed by other team members. Statistical analyses were conducted to determine whether for this participant there were any differences in performance across agent types. Data violated assumptions of normality in nearly every comparison, so non-parametric methods were used. Data were grouped pair-wise by session, so Friedman's tests were conducted followed by Wilcoxon



(a) Overall task performance



(b) Time-Steps hit by the hazard



(c) Total heat gain

Figure 6.4: Performance metrics. Bars represent the mean over trials for each metric, normalized by the maximum possible value of that metric. Error bars represent the 95% confidence interval. N = no agent; BC = bit-cascade agent; TCT = tile-coded trace agent.

Signed-Rank tests with a Holm-Šidák correction for multiple comparisons. Significance is reported in Figure 6.4, in all cases at the family-wise $\alpha = 0.05$ level. Specific results of the statistical analyses are reported in Table 6.1.

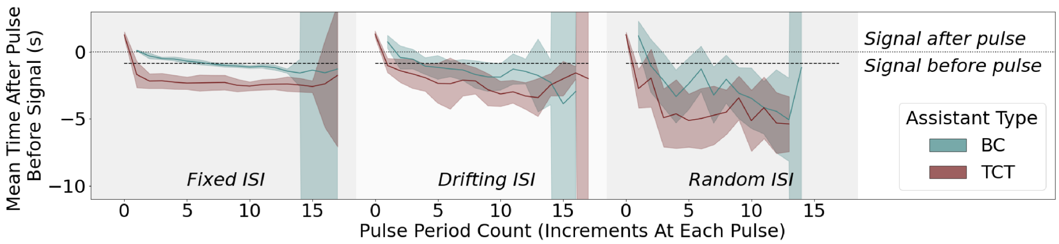
6.4 Quantitative Analysis

Looking first at overall task performance (Figure 6.4a), we see a small (and not statistically significant) increase in performance in the fixed ISI condition when the participant was paired with either agent. For the more difficult conditions where the ISI changes over the course of the trial, there is no clear difference in overall task performance depending on agent pairing. In general, these results suggest that overall task performance is not a clear indicator of any differences between human-agent pairings in this setting. Figure 6.4b

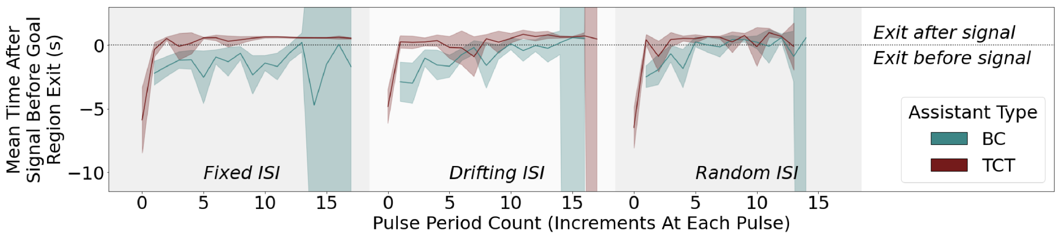
Table 6.1: Statistical analysis results. Comparisons are made across assistant pairings (N = no agent; BC = bit-cascade agent; TCT = tile-coded trace agent) for each ISI condition. Significance ($\alpha = 0.05$) is indicated in bold text. For Friedman's tests, $\chi^2_{critical}(2) = 6.20$.

		Points Cached	Steps Hit by Hazard	Heat Gain
A Priori Tests (Friedman's Chi-Square)				
Fixed	$\chi^2(2)$	5.2	**14.0**	1.4
	p	0.0755	**0.0009**	0.4966
Drifting	$\chi^2(2)$	1.4	1.4	**7.2**
	p	0.4895	0.4966	**0.0273**
Random	$\chi^2(2)$	0.7	**11.4**	**13.4**
	p	0.7165	**0.0033**	**0.0012**
Post Hoc Tests (p-values from Wilcoxon Signed-Rank)				
Fixed	N vs BC	0.1415	**0.0432**	0.9594
	N vs TCT	0.1724	**0.0151**	0.5550
	BC vs TCT	0.7344	0.1763	0.4930
Drifting	N vs BC	0.6831	0.4257	0.3642
	N vs TCT	0.6831	0.4881	0.3642
	BC vs TCT	0.5507	0.6465	**0.0206**
Random	N vs BC	0.7990	0.2026	0.0593
	N vs TCT	0.7990	**0.0278**	**0.0151**
	BC vs TCT	0.7990	**0.0278**	**0.0329**

shows differences in the proportion of time-steps where the participant was hit by the hazard. In the fixed ISI condition, the participant spends less time being hit by the pulse when paired with either agent as compared to none. In the random ISI condition, the participant is hit by the pulse less when paired with the TCT agent than when paired with the BC agent, or no agent. Figure 6.4c displays the participant's heat gain in each condition, which corresponds to the proportion of time spent in the goal region. Differences here appear in the more challenging conditions, where the participant spends less time in the goal region when paired with the TCT agent than when paired with the BC agent. Considering the charts of Figure 6.4 together, it appears that the participant engages in more cautious behaviour when paired with the TCT agent as compared to the BC agent (they gain less heat, and are hit by the hazard less often), while attaining comparable task performance. This result



(a) Mean interval between signal from agent and hazard onset over trial length. The minimum useful lead time (dashed line) before the hazard pulse (dotted line) is 0.89 seconds, and corresponds to the participant's mean exit speed. It does not include reaction time.



(b) Time interval between signal from agent and goal region exit, shown as a trajectory over the length of trials. Negative data indicate the participant leaving the goal region in advance of the agent's signal.

Figure 6.5: Data are shown as the mean (solid line) and 95% confidence interval (shaded region) of the data for each pulse. Due to the randomization of the starting ISI and fixed trial duration, some trials with shorter ISIs presented more pulses than others. This led to the occurrence of one or two trials with high pulse count (>14), resulting in the large confidence intervals at the ends of these plots.

suggests possible differences in participant behaviour across agent pairings.

We are able to examine agent learning directly because the agent's learning of the task does not depend on the participant's actions. Figure 6.5a shows the mean interval between the agent signal and the onset of the hazard pulse, for each pulse over the length of the trials. This interval can be interpreted as "how long before the onset of the pulse did the agent's prediction of the pulse rise above the threshold for signalling?". When the agent's cue is less than 0.89s before the hazard (above the dashed line), the signal doesn't give the participant enough time to react given how long it takes to leave the hazard region. In the simplest prediction task with fixed ISI, the BC agent is unable to reliably give a useful signal (below the dashed line) until after about the sixth or seventh pulse while the TCT agent is able to give a reliably useful signal after only the second pulse. More challenging conditions introduce more variance in these intervals, but the trend remains that the TCT agent provides useful signals earlier, and more reliably. This is because the BC representation has finer feature bins in the region near the pulse compared to the TCT representation, leading to more accurate but slower learning.

The length of time between the agent's signal and the participant's exit from the goal region is plotted in Figure 6.5b. A negative value on this chart indicates that the participant left the goal region before being cued by the agent. Here, we see that the participant exhibits clear behavioural differences when interacting with each agent. When paired with an agent with a TCT representation, the participant nearly always waits for the agent signal before leaving the goal region (data above the dotted line). When paired with an agent with a BC representation, the participant is much more likely to exit the goal region before the agent gives a signal. In the fixed ISI condition, when paired with the TCT agent, the participant seems to move after the agent's cue as early as the second or third pulse of a trial. Under the same conditions, when paired with a BC agent, the participant relies entirely on their own internal timing. For the more difficult conditions, the participant eventually moves after the cue of either agent, but aligns their movements with the TCT agent's cue more readily than with the BC agent's cue. While it is tempting to interpret this feature of the data as the participant relying on the TCT agent's cue more than the BC agent's cue, there is insufficient evidence from these charts alone to conclude how the participant is using either signal, as we will see in Section 6.5.

6.5 Qualitative Analysis

Qualitative data was gathered by the participant after each session in free-form text, prompted but not restricted by the following questions. Experimenter-developed questions were posed by the member of the study team conducting the qualitative analysis at the outset of the trials. Participant-developed questions were generated independently by the member of the study team acting as the participant, and evolved as the study progressed.

Experimenter-developed questions:

- Are you trying to figure out how the agent (and environment) work?
- For the whole trial?
- If not, did you figure it out or just start to trust it?
- After time, or successes?
- How much do you notice or think about the other agent at the beginning? The middle? The end?

Participant-developed questions:

- Changes in when and how I counted: did I count from the start of the trial? Did I shift to just counting from the agent cue and not counting from the beginning? When did I shift between these and under what conditions or observations on timing?
- What agent behaviours did I like and not like?
- Adaptation rates: what were my expectations on response or learning times for agents?
- Thinking of agents as adaptive systems / predictors or not?
- What conditions did I lose confidence in the machine; when did I gain confidence?
- When did trust in the agent occur quickly?

Discourse analysis seeking recurring sentiment and themes indicated that trust was a major component of the participant's interaction with the system, which affected other factors including cognitive load and use of the agent's cue in unexpected ways. The participant noted that they trusted the agent more when it was demonstrably correct earlier in the trial. Once trust was built, they noticed a decrease in cognitive load: "*With trust in my agent, I can let [my] mind* *wander*". Notes such as "*[The] agent helps me feel like I have a lower bound of safety once it is trained, and then can choose my risk based on its feedback*" suggest that the participant engaged with the task actively and strategically, and used the agent's cues as part of that strategy in more complex ways than rote cue-to-movement. In fact, with sufficient experience with the agent, the participant would sometimes engage in risky behaviour: "*I was at times racing the pulse; the agent would cue me but I would see the heat bar almost full and then gamble that it would fill fast enough before [the hazard] came, given what I knew about the relationship between cue and future pulse.*" Even in cases where the agent inadequately predicted the hazard, the participant still used the agent signal as information to inform their strategy, but relied on their mental timekeeping to inform their movements. Regarding these situations the participant notes: "*it was not fast enough to be useful in advance of [the hazard], so I mainly used it as a checksum*", indicating that they verified their mental timekeeping by comparing it against their acquired knowledge of how the agent keeps time. While this particular behaviour is likely unique to participants familiar with GVF agents of this nature, the anecdote provides a clear example of how a user's mental model of an agent will affect their interactions with it. It also indicates that evaluating future participants' understanding of how the agent learns will be key to understanding their interactions.

6.6 Comparing Quantitative and Qualitative Results

In both the quantitative and qualitative analyses we see human trust of the agent emerging as an important theme. The participant's notes suggest that using the sign of the signal-to-exit interval (Figure 6.5b) as an indicator of human trust might miss parts of the picture, since the participant makes use of the agent signal in other ways than as simply a cue to move. Other quantitative measures of trust should be sought, to corroborate this interpretation. One particular notion of intense trust called out in the participant notes (when the participant is "racing" the pulse, caching points after the agent signal but before the hazard) is also visible in the quantitative data. Of the 14 instances where a point caching event is recorded after an agent signal and before a hazard, 13 of these instances occurred when the participant was paired with the TCT agent. This risky behaviour with the TCT agent contrasts with the indications from Figure 6.4 that the participant behaved more cautiously with the TCT agent. Pairing this contrast with the qualitative discussion, we see that with high levels of trust in the agent, the participant is able to more flexibly choose a strategy, behaving boldly or cautiously as the situation warrants. It should however be noted that (as shown in Figure 6.5a) the TCT agent reliably gives more lead-time than necessary before the pulse, leaving time for pulse racing that the BC agent does not, meaning that pulse racing may not be a fair indicator of trust.

6.7 Discussion

Specific quantitative and qualitative measures to assess human trust in the agent would be particularly informative for future studies, especially if such measures could assess changes in levels of trust over the course of a trial or across sessions. One such task modification might involve the introduction of a secondary, voluntary and cognitively demanding task that could be performed simultaneously while gathering heat. While engaged with the secondary task, the participant would need to place trust in the agent to keep track of the timing in the primary task (i.e., effect a form of cognitive offloading (Risko & Gilbert, 2016)), making engagement in the secondary task a good measure of trust.

Supposing that future studies with a direct measure of trust confirm that participants trust the TCT agent more than the BC agent, two points of discussion emerge. First, the apparent differences in levels of trust between the types of agent can only be attributed to the different representations of the agents, as all else is equal. While the BC agent is able to achieve greater prediction accuracy than the TCT agent (because of the BC agent's finer feature bins in the region of the hazard), fast learning appears to be more important than accuracy for the development of human trust. The threshold and representation bin-widths in this experiment were chosen considering late-trial performance, so that once the predictions stabilized both representations would give roughly equal notice before a pulse. A lower threshold or wider feature bins would likely have allowed the BC agent to provide reliably useful signals earlier in the trials. Understanding the relationship between feature representation and threshold levels in both early and late-learning contexts will be important for any future studies or applications making use of Pavlovian signalling for communication. Second, the participant in this experiment displayed more richly varied strategies with the TCT agent than the BC agent, presumably because of a greater degree of trust. Specific assessments regarding how participant strategies are affected by trust in the agent may be illuminating, and should involve specific metrics to assess changing strategies over time. Finally, using a pre-trained agent as a baseline comparison will be necessary to assess the effect of active learning on these measurements.

6.8 Limitations

The generality of this pilot study is limited by our use of a single participant who was also a member of our study team. While blinded from the particular conditions they were interacting with, they were deeply familiar with the agent architectures, task dynamics, and learning machines in general. We expect that the introduction of naïve participants will also involve a co-learning phase at the beginning of sessions where the participant and agent are both learning the task simultaneously. Since we found early interactions to be of great influence in trust-building with our expert participant, we expect that a co-learning phase will affect trust, but make no hypothesis about what that effect might be.

6.9 Conclusions and Future Work

This pilot study examined an approach to agent-human support characterized by real-time machine learning and straightforward ongoing interactions; our results suggest that trust in the system's capabilities is a major component of a human's interaction with a continually learning system. There are also indications that this trust may be dependent mainly on early interactions with the system, while the agent is still developing competency. Future studies should include metrics that specifically measure trust, and should include analyses to determine possible correlations between levels of trust and agent competence. There may also be correlations between levels of trust and strategic behaviour. Finally, a future study should include a greater number of participants, with a diversity of experience in interacting with learning machines. For other future time-based prediction experiments or applications involving human actors with machine agents, we make no particular recommendations about representation or threshold choices, as we understand these to be task-specific. We do however stress the importance of these choices, and recommend that they be made with both early and late learning stages in mind, and considering the interaction between the human and machine's actions.

Chapter 7

Understanding Human Interaction with Real-Time Adaptive Feedback During Simulated Prosthesis Use

Modern myoelectric prostheses have immense potential to provide increased function for people with limb difference. Machine learning is a widely researched approach to improving how people send signals *to* these devices but is not yet widely studied in providing signals *from* the device to the user. There is a lack of understanding about the impact of real-time adaptive feedback signals on users. Using a convergent-core mixed methods design, our present work contributes a quantitative analysis of performance metrics alongside assessment of movement and gaze behaviours of participants performing a task with a wearable robot hand controlled by electromyography while feedback about a task is being provided by their prosthesis. In addition, semi-structured interviews were also conducted and were analyzed using descriptive thematic analysis. The combination of these methods provided a more thorough understanding of the impact of real-time adaptive feedback signals on users than would be available from questionnaires and quantitative data alone. Our study found that despite the learned signals not providing warnings in the early interaction, and being less useful to the user, they were not disregarded or cause the user to shun the device. It also served to highlight a disconnect that exists between the experience of users and the measurable performance of their interactions with devices. This demonstrates the potential of mixed-methods approach's to contribute nuanced understandings of human machine interactions and accelerate advances in the field, especially in the domain or upper limb prostheses.

7.1 Introduction

Prosthetic limbs have the potential to play a vital role in improving function, independence, and quality of life in general for individuals experiencing limb loss. However, significant challenges remain in achieving the full potential of the current state-of-the-art devices. Users have difficulty accessing all the features of modern myoelectric prostheses with the control options available, and feedback is not yet widely commercially provided (Biddiss & Chau, 2007b; Cordella et al., 2016; Yamamoto et al., 2019). A principal reason for users' struggles to access the full capacity of modern devices stems from the high degree of anatomical functionality available in devices, contrasted with the limited ability of users to interact with them. As a result, significant research efforts are required to successfully harness the full capacity of these modern prostheses, instead of relying on fixed, anatomically appropriate gestures that function similarly to traditional cable and pulley systems. Researchers are actively exploring advancements in both *control* and *feedback* to bridge the gap between the capabilities of these devices and their successful application for the users who stand to benefit the most.

Studies into the control of myoelectic prostheses date back as far back as the 1940s (Childress, 1985). In the 1970's, published development of machine learning for multi-joint control began to appear (Childress, 1985; Herberts et al., 1973). Around the same, research into providing feedback to users of myoelectric prostheses also began (Mann & Reimers, 1970). While research into control progressed quickly to include machine learning, the research interest in machine learning applications to upper limb prostheses has not turned to feedback. It is especially under-explored with regard to real-time machine learning and feedback. Our present work seeks to add knowledge to the existing gap in our understanding of humans interacting with a device that uses machine learning to adapt, in real-time, feedback signals for the human user.

7.1.1 Machine Learning in Control

There are several promising avenues of research relating to acquiring and utilizing signals from a prosthesis user to operate their device as seamlessly and naturally as possible (Castellini et al., 2014). The goal of prosthesis control is to allow skillful operation of several of the degrees of freedom of the limb in as natural and effortless a manner as possible (P. A. Parker et al., 2006). Many studies have examined what sensors to use when reading signals from the user, such as force myography, sonomyography, and the widely used electromyography (EMG), as well as techniques to process those signals to ultimately achieve control (Castellini et al., 2014; Fougner et al., 2012; González & Castellini, 2013; P. A. Parker et al., 2006; Radmand et al., 2016; Sartori et al., 2018). All of these attempt to connect the desires of the user to device actuation by reading physiological signals.

A commonly explored technique in research is the application of machine learning to improve control of robotic prostheses. Machine learning is applied to help interpret and use the signals being read from the user in more nuanced ways. This most commonly takes the form of what is termed "pattern recognition" to get more out of EMG signals by learning associations between EMG signals in combination with each other and a desired actuation (Scheme & Englehart, 2011). A weakness in pattern recognition is that varied limb positions can change the way EMG signals are read and therefore render pattern recognition systems unreliable (Scheme et al., 2010). Recent work has seen recurrent convolutional neural networks with regression applied to restore reliability, as well as allow simultaneous use of multiple joints (Williams et al., 2022; Williams, Shehata, Cheng, hebert, et al., 2024). Another technique that has been explored is temporal-difference (TD) learning techniques from the field of reinforcement learning (RL) (Pilarski, Dawson, Degris, Carey, et al., 2013). As with pattern recognition, RL has been demonstrated to be able to allow multiple joints to move simultaneously (Sherstan et al., 2015). Techniques from the field of RL have also been used to allow a fixed switching order of functions to, over time and through use, anticipate the function the user will use next and move it to the top of the list. This method is called "adaptive switching" as introduced by Edwards, Dawson, et al. (2016) and used both in prosthesis and exoskeleton research (Edwards, Dawson, et al., 2016; Faridi et al., 2022).

The machine learning methods used for adaptive switching, techniques from the field of reinforcement learning, are of particular interest. This is because these methods are capable of learning in real-time as a device is used, from the sensory-motor data the robot experiences of the interaction that is actively happening (Sutton, 1988; Sutton & Barto, 2018). Temporal-difference (TD) learning methods can utilize a variety of signals being read from the environment and learn to make predictions about them (Sutton et al., 2011; White, 2015). The capacity for TD methods to learn in real-time from observations being made about a task as a user is doing it adds the potential for assistive robotics to adapt to patient-specific uses and needs, which is of great interest in meeting the injury and life specific needs of patients in rehabilitation science (World Health Organization, 2007).

7.1.2 Machine Learning in Feedback and Coordination

Along with control, there have been other advances, and increasing interest in recent years, in providing signals from the device to users (Schofield et al., 2014; Svensson et al., 2017). This body of work examines both the kinds of feedback to use, such as vibrotactile, mechanotacile, and audible, as well as how to provide it (Clemente et al., 2015; Shehata et al., 2020; Shehata et al., 2018a; Wells et al., 2022). We note that while in this work we focus on non-invasive means of feedback, there are also promising results utilizing advanced surgical procedures or implants to provide sensation to users (Hebert et al., 2013). Invasive feedback methods have even been able to allow the perception of different textures (Svensson et al., 2017). Interestingly, results on the benefits of non-invasive feedback in the control of artificial limbs are mixed (Saunders & Vijayakumar, 2011; Sensinger & Dosen, 2020; Shehata et al., 2018a). This has

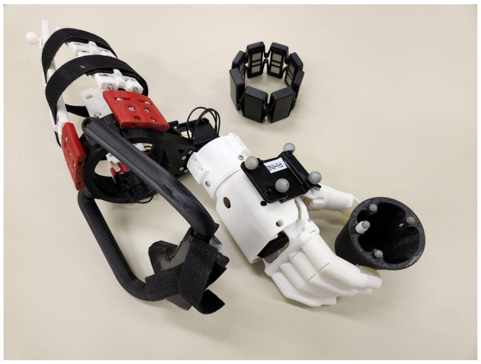


Figure 7.1: The brace restricts motion of the wearer's wrist, which encourages isometric contraction to "simulate" muscles in a residual limb. The brace also allows the robot hand to be placed in front of the user's biological hand. Reusable cups 3D printed from NinjaFlex with the same shape but different stiffnesses. EMG signals were captured using the Myo armband.

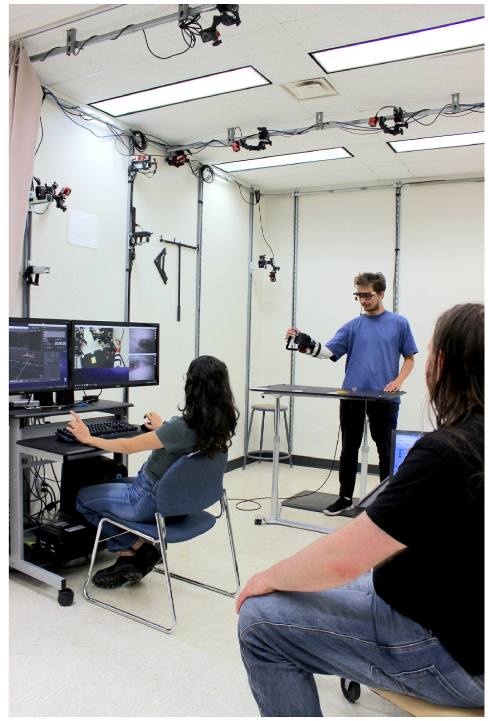


Figure 7.2: Gaze and Movement Analysis (GaMA) combines gaze vector and motion capture data. The participant wears glasses with eye-tracking cameras, and fixed motion capture cameras record the motion capture markers on lab equipment. The laptop is running the robot software. All three data streams are recorded and synchronized using Lab Streaming Layer.

led to a fascinating conversation on the nature of feedback in the interaction with a prosthesis and how it is actually used (Sensinger & Dosen, 2020).

The use of machine learning to provide signals about a task to humans making control desicions is a relatively new area of research. Work from the field of reinforcement learning has demonstrated the ability of machines to learn to signal to other machine learning agents to coordinate to accomplish tasks (Cao et al., 2018; Lazaridou et al., 2016). It has also been shown that machine agents can be designed to consider the impact their actions have on other agents, which leads to communication between agents even when an explicit channel was not specified (Jaques et al., 2019). With regard to machine learning and humans, the current state-of-the-art research is primarily focused on how humans can effectively teach or otherwise assist machine learning agents in learning (Retzlaff et al., 2024; Taylor et al., 2023). This is akin to control in the prosthesis setting, as the focus is on attaining signals from the user in order to have the device interact with the world in some way. When research involves signals going from the machine agent to the human, more aligned with feedback in the prosthesis setting, the focus is on humans understanding what the machine agent has learned, and why it has learned those outcomes (Holzinger, 2018; Longo et al., 2024).

Prosthesis use has an important feature compared to many human-machine interactions. Not only is the user interacting with a device, but a device that is a robot that shares intimate space with the user. As machine learning is added to devices in this setting, the complexity of the interaction beween the user and the device becomes more complicated an nuanced. It therefore becomes increasingly important to understand the interaction between the user and the device. It has been postulated that enabling bi-directional communication between the user and the machine-learning-enabled prosthesis is of vital importance (Kalinowska et al., 2023; A. S. R. Parker & Pilarski, 2021). The ability of the machine agent to signal the human could be the key to promoting effective interaction between humans and devices (Jackson, 2002; Schofield et al., 2021). Although there are concerns around two intelligent agents attempting to co-adapt in real-time, there has been previous work demonstrating that a human and computational agent can co-adapt in real-time and achieve positive outcomes (Couraud et al., 2018; Müller et al., 2017). This has also been demonstrated in recent work that utilized a human participant and a reinforcement learning agent that worked together to avoid a negative environmental interaction (Brenneis et al., 2022; Pilarski et al., 2022). Motivating the work that follows in this manuscript: to use machine learning to adjust the feedback a user receives from a device in real-time, it is important that we understand how users process these interactions. To examine this, sixteen participants were broken into two groups of eight that would use different warning conditions. They wore a simulated prosthesis (Fig. 7.1) that they used to move deform-able cups on a table using either direct or predictive feedback.

7.1.3 Methodological Overview

To explore this question we designed a mixed-methods study to investigate the effects of real-time machine-learned signals on human interaction with wearable robotic devices. The study utilized a convergent-core mixed-methods design (Creswell, 2021). This design seeks to collect quantitative and qualitative data together in an attempt to learn more from the combined data sets than could be learned from either individually. Since the ultimate goal is to explore how real-time machine-learning can be applied to interventions intended for use with people, it is crucial to this study to examine rich qualitative data on the user's experiences and their interpretations of device use alongside the measurable outcomes.

The philosophical paradigm used in the interpretation of this data was pragmatism. It focuses on the idea that actions we take change the world around us, that there is a connection between action and thought, and acknowledges that experience shapes thought (Dewey, 1908; Goldkuhl, 2012; Kaushik & Walsh, 2019). This allows for there to be an objective reality, which can be externally measured, and an internal reality that is interpreted by the individual but is impacted by the shared external reality (Dewey, 1908; Kaushik & Walsh, 2019). This appeared in the examination of the measurable external actions participants took alongside the internal beliefs participants had developed, and how those internal beliefs and consequent actions taken may have been different when interacting with a learning agent. This made pragmatism well suited to a mixed-methods study design as well, as it does not seek to make claims about the nature of truth and reality but allows there to be multiple truths open to scientific enquiry (Kaushik & Walsh, 2019).

Quantitative data collection used an integration of motion capture and gaze vector data called Gaze and Movement Assessment (GaMA). GaMA provided detailed metrics related to how participants were moving and what they were looking at as they did the task (Boser, 2019; Williams et al., 2019; Williams, Shehata, Cheng, Hebert, et al., 2024). Through these metrics, we sought to gain insight into the participant's internalization of the interaction, such as the participant leading the motion of their hand with their eyes as can be seen in biological limb use (Cheng et al., 2022; Lavoie et al., 2018; Williams et al., 2021).

Analysis of the interviews was done using descriptive thematic analysis, based on methods of reflective thematic analysis outlined by Braun and Clarke (Braun & Clarke, 2021a, 2021b). This was adapted for the mixed-methods, pragmatic, approach of this study. Qualitative coding was done inductively, by a single coder, and codes remained close to the data. The coder was mindful of their belief that feedback will help the participant by assisting in the participant's construction of a model of the interaction, and that trust between the human and device is fragile and important.

In keeping with journal formatting requirements, detailed experimental methods and materials information are provided at the end of this dissertation, where it would be in the Appendix of this chapter for submission.

7.2 Experiment and Results

Sixteen (16) participants were asked to move a cup from one side of a table to the other side without crushing it, and then with the next motion move the cup back. The cup was 3D printed in NinjaFlex (NinjaTek, Lititz, USA) to allow it to be distorted, or crushed, without actually breaking. There were three different stiffnesses of cup: heavy, medium and light. The heavy cup visually distorted the least of the three stiffnesses, and the light cup visually distorted the most for the same load reading from the servo in the hand. When the cup was crushed, the load signal from the servo in the hand would pass a threshold, the same threshold for every cup and an audible signal played over PC speakers connected to a laptop to inform the participant. Similarly, when the participant was doing their warning trial (outlined below), a warning would play at an earlier threshold and if the cup was squeezed further the break sound would play. Participants moved the cup using a physical sim-

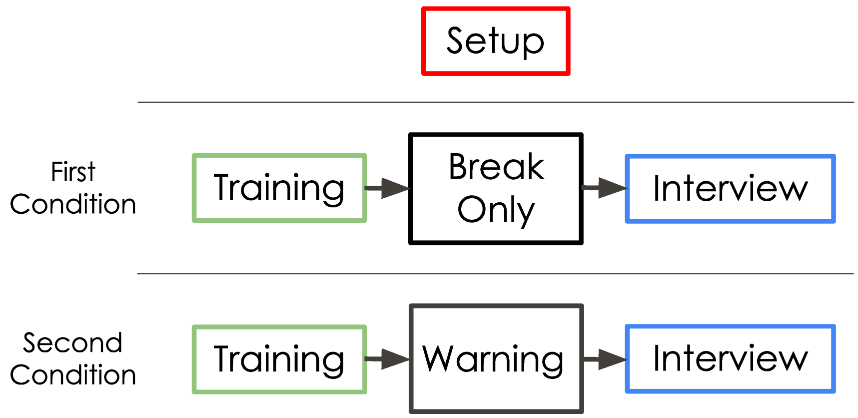


Figure 7.3: Participants were randomly assigned to one of two groups; they used either direct feedback or predictive feedback in the warning condition. After participants were helped into the simulated prosthesis and GaMA calibrations were complete, a training session using a stiff cup was done. Following that, all participants progress through the experiment cups from the stiffest, which visually distorted the least, to the most compliant, which visually disrted the most, with no warning sound. Following this was an interview, another training session identical to the first along with a gaze calibration for GaMA. Participants then again progressed through the experiment cups in the same order, but this time with a warning signal. Finally, another interview was conducted.

ulated prosthesis: a robot hand positioned aligned with and in front of the participants' biological hand by a brace that restricted the motion of their wrist (Hallworth et al., 2022). The robot hand of the simulated prosthesis was controlled via EMG acquired from the user's forearm muscles using a Myo armband device (Thalmic Labs, Kitchener, Canada).

There were three phases to the study which can be seen in Fig. 7.3. The first phase was a training phase where the participant practiced moving a stiff plastic drinking cup using the robot hand and EMG control with no audible feedback. This training phase was also conducted before the second trial condition. In the first trial condition the participants moved all three cups thirty times each, starting with the heavy cup, then the medium, then the light. Each cup used the same value of load signal from the servo for the break point, and later the warning, sound but visibly distorted differently as a result of their different stiffnesses. In the first trial condition the participant only

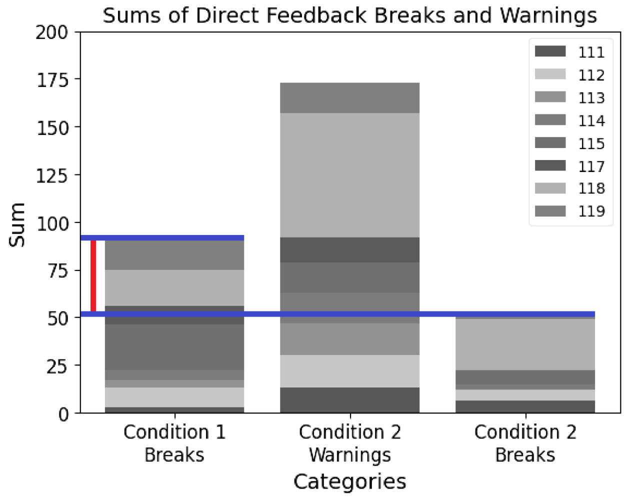


Figure 7.4: The total number of breaks for all participants who would later receive direct feedback in the first condition, break sound only, followed by warnings for the second condition, the trial with warnings, and breaks for the second condition. A large drop in breaks between the trial without and with warnings can be seen.

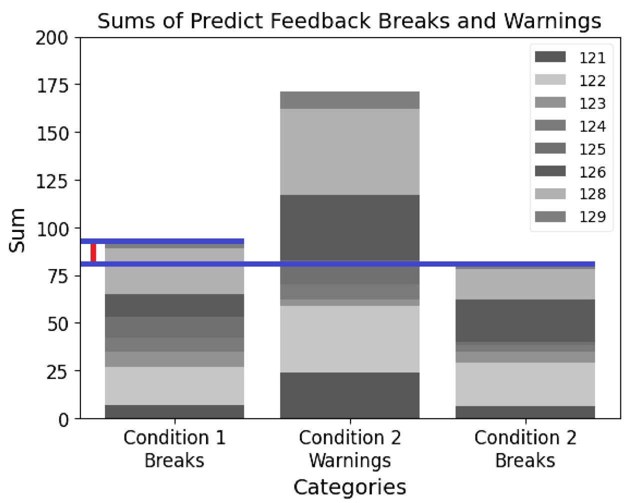


Figure 7.5: The total number of breaks for all participants who would later receive predictive feedback in the first condition, break sound only, followed by warnings for the second condition, the trial with warnings, and breaks for the second condition. It can be seen between this figure and Fig. 7.4 that the breaks in the first condition, with no warnings, and the warnings received in the second condition are similar, but there are more breaks in the second condition when predictive feedback was used.

received an audible signal when the cup had been crushed. The second trial phase introduced a warning sound, distinct from the crush sound, that would play in anticipation of the crush condition. The warning sound was triggered

Feedback Comparison	Cup	Eye-Hand Arrival Latency (s)		Eye-Hand Leaving Latency (s)	
		Pick-Up	Drop-Off	Pick-Up	Drop-Off
None vs Direct	All	2.52 vs 2.41	0.451 vs 0.45	-0.97 vs -0.97	-1.33 vs -1.14
	H	2.65 vs 2.57	0.45 vs 0.46	-0.97 vs -1.04	-1.41 vs -1.21
	M	2.53 vs 2.47	0.46 vs 0.43	-0.98 vs -0.99	-1.36 vs -1.11
	L	2.38 vs 2.19	0.44 vs 0.46	-0.97 vs -0.88	-1.21 vs -1.11
None vs Prediction	All	2.79 vs 2.70	0.47 vs 0.47	-0.96 vs -0.95	**-1.55 vs -1.35** *
	H	**3.03 vs 2.72** *	0.53 vs 0.47	-1.01 vs -1.00	**-1.65 vs -1.33** *
	M	2.70 vs 2.79	0.47 vs 0.49	-0.93 vs -0.89	-1.55 vs -1.40
	L	2.64 vs 2.60	0.43 vs 0.44	-0.96 vs -0.95	-1.46 vs -1.30
Direct vs Prediction	All	**2.41 vs 2.70** *	**0.45 vs 0.47** *	-0.97 vs -0.95	-1.14 vs -1.35
	H	**2.57 vs 2.72** *	**0.46 vs 0.47** *	-1.04 vs -1.00	-1.21 vs -1.33
	M	**2.47 vs 2.79** *	**0.43 vs 0.49** *	-0.99vs -0.89	-1.11 vs -1.40
	L	**2.19 vs 2.60** *	**0.46 vs 0.44** *	-0.88 vs -0.95	-1.11 vs -1.31

Table 7.1: Eye-hand arrival latency (EHAL) and eye-hand leaving latency (EHLL) statistical results. EHAL is the time, in seconds, that the eyes arrive at a target relative to the transport phase, and EHLL is the eyes leaving a target relative to the transport phase. This table shows comparisons between groups, direct vs predictive feedback participants, and within groups, comparing the no-warning vs warning conditions. Results are further broken down by cup stiffness: heavy (H, which visually distorted the least), medium (M), and light (L, which visually distorted the most), as well as all stiffnesses together. Post-hoc significant differences between compared values ($p < 0.05$) are denoted in bold with an asterisk, and underlines mark the improved values. The table shows minor differences that suggest some increased user confidence when they use predictive feedback.

either when the load reading of the servo of the hand went above a threshold (for direct feedback participants) or when a machine-learned prediction of that same load signal went above the same threshold instead (for predictive feedback participants). Randomly for one motion in five, so six times per trial, the participant was asked to crush the cup purposefully. The sixteen participants were split randomly into two groups of eight, half of them had the direct feedback for their second trial, and half had the predictive feedback. Interviews were conducted with each participant after each trial condition, thirty motions with three cups, but not after training. Each participant, then, was interviewed two times for approximately ten minutes each.

Feedback Comparison	Cup	Phase Duration (s)			
		Reach	Grasp	Transport	Release
None vs Direct	All	0.87 vs 0.84	1.91 vs 1.83	1.18 vs 1.15	1.18 vs 1.05
	H	**0.95 vs 0.85 ***	2.04 vs 1.89	1.20 vs 1.20	1.24 vs 1.10
	M	0.86 vs 0.85	1.91 vs 1.86	1.19 vs 1.12	1.18 vs 1.03
	L	0.81 vs 0.83	1.78 vs 1.74	1.14 vs 1.12	1.11 vs 1.04
None vs Prediction	All	0.98 vs 0.96	1.97 vs 1.85	1.25 vs 1.22	**1.31 vs 1.13 ***
	H	1.04 vs 0.99	**2.10 vs 1.85 ***	1.35 vs 1.27	**1.40 vs 1.10 ***
	M	0.97 vs 0.97	1.98 vs 1.93	1.20 vs 1.21	1.27 vs 1.16
	L	0.94 vs 0.92	1.82 vs 1.76	1.19 vs 1.18	1.26 vs 1.12
Direct vs Prediction	All	**0.84 vs 0.96 ***	**1.83 vs 1.85 ***	**1.15 vs 1.22 ***	**1.05 vs 1.13 ***
	H	**0.85 vs 0.99 ***	**1.89 vs 1.85 ***	**1.20 vs 1.27 ***	**1.10 vs 1.10 ***
	M	**0.85 vs 0.97 ***	**1.86 vs 1.93 ***	**1.12 vs 1.21 ***	**1.03 vs 1.16 ***
	L	**0.83 vs 0.92 ***	**1.74 vs 1.76 ***	**1.12 vs 1.18 ***	**1.04 vs 1.12 ***

Table 7.2: Results comparing the total time in seconds for each phase of a motion between groups, direct vs predictive feedback participants, and within groups, the no-warning vs warning conditions. Results are further broken down by cup stiffness: heavy (H, which visually distorted the least), medium (M), and light (L, which visually distorted the most), as well as all stiffnesses together. Post-hoc significant differences between compared values ($p < 0.05$) are denoted in bold with an asterisk, and underlines mark the improved values. This table shows that participants using direct feedback performed the task faster.

7.2.1 Quantitative Findings

Participants using predictive feedback were found to have less time spent in the critical phases, grasp and release, despite direct feedback having overall shorter duration's. Combined with the findings in the gaze data, it seems as though participants had somewhat greater confidence when using the predictions. This confidence is enough to appear in their behaviour during less crucial sections of the interaction, but is not enough to impact the more critical segments, reach and grasp.

Statistical analyses were performed to compare eye-hand arrival latency, eye-hand leaving latency, task times, relative task times, and percent fixation results for different feedback types for each cup stiffness. **For comparisons of no feedback to direct or prediction feedback** (that is,

Feedback Comparison	Cup	Relative Phase Duration (%)			
		Reach	Grasp	Transport	Release
None vs Direct	All	17.68 vs 18.33	36.00 vs 35.58	23.40 vs 23.88	22.91 vs 22.21
	H	18.13 vs 17.98	36.28 vs 35.99	22.57 vs 24.05	23.03 vs 21.99
	M	17.38 vs 18.30	36.23 vs 36.20	23.72 vs 23.48	22.67 vs 22.02
	L	17.53 vs 18.71	35.50 vs 34.55	23.93 vs 24.12	23.04 vs 22.62
None vs Prediction	All	18.63 vs 19.18	33.41 vs 34.07	23.88 vs 24.92	**24.07 vs 21.83***
	H	18.08 vs 19.52	33.65 vs 33.23	23.93 vs 25.62	**24.35 vs 21.63***
	M	18.81 vs 19.12	33.52 vs 34.98	23.50 vs 24.35	**24.16 vs 21.55***
	L	19.01 vs 18.90	33.07 vs 34.00	24.21 vs 24.79	23.71 vs 22.32
Direct vs Prediction	All	**18.33 vs 19.18***	**35.58 vs 34.07***	**23.88 vs 24.92***	**22.21 vs 21.83***
	H	**17.98 vs 19.52***	**35.99 vs 33.23***	**24.05 vs 25.62***	**21.99 vs 21.63***
	M	**18.30 vs 19.12***	**36.20 vs 34.98***	**23.48 vs 24.35***	**22.02 vs 21.55***
	L	**18.71 vs 18.90***	**34.55 vs 34.00***	**24.12 vs 24.79***	**22.62 vs 22.32***

Table 7.3: Results showing the percent of the total time of a motion spent was spent in each phase of a motion, comparing between groups, direct vs predictive feedback participants, and within groups, the no-warning vs warning conditions. Results are further broken down by cup stiffness: heavy (H, which visually distorted the least), medium (M), and light (L, which visually distorted the most), as well as all stiffnesses together. Post-hoc significant differences between compared values ($p < 0.05$) are denoted in bold with an asterisk, and underlines mark the improved values. These results show that for the critical phases, grasp and release, participants using predictive feedback had lower relative durations.

with paired samples): Participants' results were first averaged across cup movements across the table. If results were normally-distributed, a repeated measures analysis of variance (ANOVA) was conducted. If results were not normally-distributed, the Friedman test was conducted. When the resulting p value was less than 0.05, post-hoc comparisons between feedback types (paired-sample t-test/Wilcoxon sign rank test) were conducted and deemed significant if the p value was less than 0.05. **For comparisons of direct to prediction feedback** (that is, with independent samples): Participants' results were first averaged between subjects across cup movements. If results were normally-distributed and population variances were equal, a one-way ANOVA was conducted. If results were normally-distributed and population variances were not equal, Welch's ANOVA was conducted. If results were

Feedback Comparison	Cup	Percent Fixation to Current (%)			
		Reach	Grasp	Transport	Release
None vs Direct	All	73.13 vs 74.65	**96.50 vs 94.48***	38.13 vs 37.31	**94.26 vs 90.49***
	H	72.01 vs 73.31	96.47 vs 94.57	37.25 vs 36.09	94.23 vs 91.18
	M	71.12 vs 74.63	96.93 vs 94.60	38.34 vs 35.61	94.89 vs 90.33
	L	76.25 vs 76.00	96.09 vs 94.27	38.80 vs 40.21	93.67 vs 89.95
None vs Prediction	All	73.44 vs 74.95	92.87 vs 93.59	35.59 vs 35.71	91.48 vs 92.34
	H	74.65 vs 75.53	92.55 vs 94.66	36.33 vs 35.57	88.50 vs 92.83
	M	73.24 vs 76.81	94.04 vs 93.57	36.29 vs 37.44	93.32 vs 93.20
	L	72.43 vs 72.50	92.02 vs 92.53	34.15 vs 34.12	92.61 vs 91.00
Direct vs Prediction	All	**74.65 vs 74.95***	**94.48 vs 93.59***	**37.31 vs 35.71***	**90.49 vs 92.34***
	H	**73.31 vs 75.53***	**94.57 vs 94.66***	**36.09 vs 35.57***	**91.18 vs 92.83***
	M	**74.63 vs 76.81***	**94.60 vs 93.57***	**35.61 vs 37.44***	**90.33 vs 93.20***
	L	**76.00 vs 72.50***	**94.27 vs 92.53***	**40.21 vs 34.12***	**89.95 vs 91.00***

Table 7.4: Percent gaze fixation is the amount of time spend looking at the dropoff or pickup location. This table shows comparisons between groups, direct vs predictive feedback participants, and within groups, comparing the no-warning vs warning conditions. Results are further broken down by cup stiffness: heavy (H, which visually distorted the least), medium (M), and light (L, which visually distorted the most), as well as all stiffnesses together. Post-hoc significant differences between compared values ($p < 0.05$) are denoted in bold with an asterisk, and underlines mark the improved values. This table shows provides little insight directly, but is used to help interpret the other GaMA results.

not normally-distributed, the Kruskal-Wallist test was conducted. When the resulting p value was less than 0.05, post-hoc comparisons between feedback types (two-sample t-test/Wilcoxon rank sum test) were conducted and deemed significant if the p value was less than 0.05.

Eye-hand arrival latency is the time, relative to the start of moving the cup for pick-up or the time of placing the cup for drop-off, that they eyes fixated on the target location relative to the hand reaching it. Eye-hand leaving latency is the time of the eyes leaving, or no longer fixating on, either the start or end location of the cup relative to hand motion. These metrics were selected as it would be expected a user showing more assurance in their motion, which is to say moving more like they would with a biological limb, would show difference in these metrics. Specifically we would expect to see the eyes lead the motion of the object to the target (Parr et al., 2018).

All four comparisons of eye-hand arrival latency at pickup showed significant differences between direct and predictive feedback as can be seen in Table 7.1. The predictive feedback resulted in earlier arrivals of the gaze to the cup and its pick-up location. However, with the heavy cup (which was easiest to crush), the prediction feedback resulted in later eye arrivals to the pick-up location. There were four significant differences found between direct and predictive feedback in eye-hand arrival latency at drop-off, shown in Table 7.1. With the heavy, medium, and all cups, the predictive feedback resulted in earlier arrivals of the gaze to the cup's drop-off location compared to direct feedback. However, with the light cup, the prediction feedback resulted in later arrivals. There were two significant differences found between no feedback and predictive feedback in eye-hand leaving latency at drop-off, which can be seen in Table 7.1. The values under the predictive feedback condition were less negative, greater, than those under no feedback. This indicates that while participants were releasing the cup they looked away from the drop-off location earlier with prediction feedback compared to no feedback but not with direct feedback compared to no feedback.

All direct vs predictive feedback comparison of phase duration show statistically significant difference in favour of direct feedback. These results can be seen in Table 7.2 This finding provides support for the case of the eye-hand arrival latency at dropoff not being functionally interesting. The comparison between direct and predictive feedback in relative phase durations show that predictive feedback has shorter grasp and release phase durations relative to the total task time as shown in Table 7.3. The statistical differences shown in Table 7.4 are very small, and hard to justfy as interesting in this study.

The servo position, load, and velocity from the robot hand were recorded along with the value of the current prediction. The overall results of this can be seen in Figs. 7.4 and 7.5.

The learning algorithm used was simple TD(0) with a selective Kanenerva coded state space, as outlined in Travink et al. 2017, composed of the normalized servo position, velocity, and load (Sutton et al., 2011; Travnik &

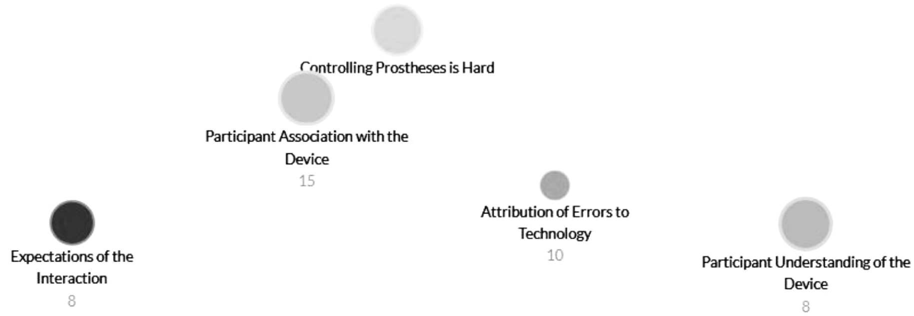


Figure 7.6: A section of the overlap mapping from the Quirkos qualitative analysis tool enabled insight into the "Control is Hard" code. The larger the circle the more quotes there are in the code. The code at the top is the code being checked for overlap. The higher other codes are to this code, the more overlap there is. This suggests the relationships between the struggles participants experienced with control and the tendency to discuss their association with the device, most specifically with "othering" it.

Pilarski, 2017). The learning target for the predictions, which is termed the cumulant, used was the normalized load reported by the servo (Sutton et al., 2011). Further details can be found in the methods section of the Appendix. The learning process was tuned to progress quickly but stably, and for the most part, was beginning to provide warnings after one or two crushes.

7.2.2 Qualitative Findings

The qualitative findings, informed by but not yet merged with the quantitative findings, suggest several phenomenon present in the interaction. Participants regularly brought up control of the device in the context of their interaction when speaking freely. They also used many different signals from the environment to construct their model of the interaction, and participants seemed to be more engaged with the adapting observed in the predictive feedback than with the direct feedback.

Transcripts were autonomously generated by rev.com and doubled checked for accuracy manually. The first round of qualitative coding was done by a member of the research team going over the transcrupts line by line and grouping statements, largely semantically. After the first round of coding, the codes themselves, rather than the transcripts, were checked for consistency and then adjusted to better represent the sentiments expressed in the relevant quotes. Following this, a side-by-side and overlap analysis of the codes was done by member of the research team. This was done using qualitative analysis software to assist with the organization and visualization of the codes and quotes. During this process, the themes were generated and checked to be consistent with the original transcripts. From this, three (3) themes were generated. These are "the difference between being informed (feedback) and being in charge (control)", "constructing the interaction with feedback", and "user understanding and engagement with predictive feedback".

The Difference Between Being Informed (Feedback) and Being in Charge (Control). Participants brought up control when asked to openly discuss their experience, despite it being told to them at several points the study is about the interaction, and the variable being manipulated is the feedback they receive from the device about the task. Errors resulting from the typical difficulties with EMG, or participant inexperience using EMG, were commonly noticed and readily discussed. Expressions of blame towards the device were connected to these control struggles, not feedback signals. This was even true in the predictive feedback case where, as will be mentioned later, participants noticed the signals failed to occur initially. This is suggested in the overlap view shown in Fig. 7.6. Participants' experience of the device occasionally soured during the second trial, and here too it seems to be connected to perceived struggles with control, is seen in the participant referring to the device as an external entity taking actions in Fig. 7.7. These control struggles lead to more statements which distanced, separated, blamed, or otherwise "othered" the hand (Jackson, 2002).

Constructing the Interaction with Feedback. Both direct and predictive feedback was often discussed in the context of participants learning about the task and robot hand as demonstrated in Fig. 7.8. Different people used different sources of feedback signals, seemingly for their own internal predictions of whether the cup was going to be crushed, and usually several signals in combination as can be seen in Fig. 7.9. The feedback signals were discussed as separate from the hand; feedback was discussed as part of the environment

> “Very annoyed. <laugh>, especially at the end when I like set the cup down, I’m trying to get to let go and it just crushed the cup. I’m like, <laugh>, let go, let go the cup.”

Figure 7.7: Participant 124, who would go on to use predictive feedback, explaining their frustration with the hand, expressed here as an entity taking it’s own actions, when the participant struggled to release the cup.

> “I don’t think the sound had played yet when I picked it up. And so I think as I was going, I sort of put more force, um, un, unintentionally at that point. And then when I heard the sound play I was like, oh, I’m putting in too much force then...”

Figure 7.8: Participant 118, who used direct feedback, talking about “the sound” as separate, but related, to the crushing of the cup.

not in association with the device. Feedback was reported to cause mental changes in perception such as shown in Fig. 7.14. Understanding of the device was, occasionally, hampered by the inconsistency in feedback created by the learning agent.

User understanding and engagement with predictive feedback. Participants who had direct feedback tended to express that they had predictive feedback when asked because it felt inconsistent. They were not very confident in this assumption, and tended to mention it was almost entirely a guess. The remarks made by participant 114 in Fig. 7.11 suggest this. The participants who had predictive feedback, on the other hand, tended to more confidently state they had the predictions. There was, however, still some lack of confidence in their assumption. Early failures of the predictive feedback to send a warning signal when a cup was crushed were noticed, but participants typically moved past this error. Later successes were then noticed and discussed as improvements in consistency as Fig. 7.10 suggests, whereas direct tended to continue to be discussed as inconsistent. A few people were confused about or disliked the predictive feedback but expressed “warming up” to their experience in

> “Probably focused on the amount of deformation that caused the sound and then trying to uh, change it for next time. See how much force I put in and you know, why.”

Figure 7.9: Participant 129 describing what they focused on when they accidenatlly crushed the cup during their first, no warnings trial. They mention several signals from the environment that they used together, as well as suggesting they used the signals to update their own internal model.

> “The first trial, it, it seemed like it took a bit to get dialled in. The second one seemed pretty consistent based on force...”

Figure 7.10: participant 128 when talking about thoughts or feelings from the second trial which used predictive feedback warnings. Notably they mention the signals struggling initially, but becoming consistant.

hindsight when they were told at the end of the second interview that the machine learning agent was generating the warnings. Figure 7.12, is a prime example of this. Learning about the device/interaction happened even when participants expressed a mentality of being a mechanism following instructions. One participant in particular expressed dislike of their prediction feedback trial but they expressed agreement that they appreciated the adaptation once they were told for sure that was what the system was doing.

7.2.3 Data Synthesis

Combining the qualitative and quantitative data provides different, more accurate, insight into the interaction. The converged data indicates a disconnect between the quantitative and qualitative data, as well as reinforcing the light improvement in confidence users had with the predictive feedback. The combined data also indicates this confidence occurs despite the predictions performing worse at preventing crushes.

A single person analyzed both the quantitative and qualitative data sets. This researcher thoroughly familiarized themselves with the quantitative data across the participant groupings first. Familiarization with the qualitative

"I think it would've been machine learning just cuz I didn't always know like if there was like a hard baseline like some it felt like it fluctuated a little bit..."

Figure 7.11: participant 114 discussing whether they thought they had predictive or direct feedback after their second trial which used direct feedback warnings. Most participants, even those using direct feedback, thought they were using predictive feedback.

"Interviewer: I really don't wanna put words in your mouth.
129: Sure.
Interviewer: Would it be fair to say you perhaps appreciated that it was adapting?
129: Yes.
Interviewer: Because you did express that in the first trials you were confused in the second trials it felt overbearing and by the third one...
129: Yeah, no, I think that's exactly it."

Figure 7.12: Participant 129 discussing their experience and switching their opinion of the predictive feedback from negative to positive after being told they did in fact have feedback that was adapting in real time. While there is some leading, the participants enthusiastic agreement suggests strong alignment of the sentiment with the participants experience.

data was achieved by correcting the autonomously transcribed transcripts, and the first qualitative coding pass. While the qualitative data was being analyzed, various questions came up that were then compared or confronted with the quantitative data. From that analysis, combinations of the data, such as those seen in Figs. 7.13 and 7.14 where the participant discussed the velocity seeming to change, were found. Ultimately, the Qualitative data was consistently "thought about" with the quantitative findings in mind, and a second look at the quantitative data occurred with the qualitative findings in mind.

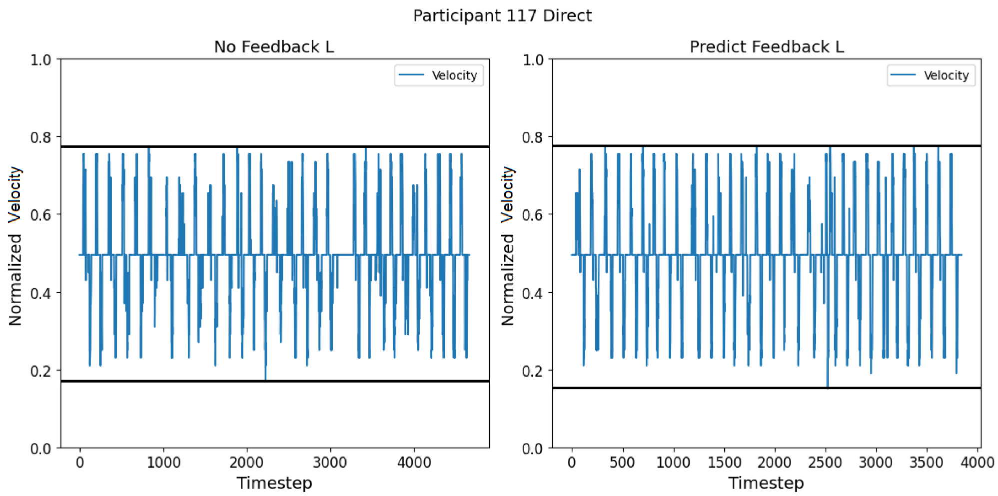
	Total Warnings	Total Breaks	Difference
Triggers	13	6	7

"I don't think that there was a, you can correct me, I don't think there was an instance where I got a warning without getting a break."

Figure 7.13: Contradiction between the recorded warnings and crushes for the second trial, the warning trial, done by Participant 111 alongside a statement they made about the efficacy of the warnings. The participant things the warnings never occured without a crush, despite this being recorded happening 7 times. This participant was displeased with their interaction as a result of control struggles and it seems to have coloured their perception of the interaction to contradict the measurements made.

7.3 Discussion

The integration of the qualitative and quantitative data provides several key insights that would not have been otherwise discovered. It is important to highlight throughout this discussion that, contrary to participant experiences such as shown in Fig. 7.14 there were few differences of import found across the GaMA latency metrics. The statistically significant differences found in the eye-hand arrival latency at pickup, indicating participants using predictive feedback moving theirs eyes to the cup earlier in the reach phase, while curious, are difficult to consider as the action we are modifying does not have impact until the cup is grasped. Therefore, the confident gaze behaviour prior to any intervention on the part of the device should not have been a result of the changes made to the warning sounds. This perhaps indicates some confidence on the part of the participants, but not enough to impact the crucial areas of the interaction, grasp and release. The findings relating to the eye-hand leaving latency at drop-off, indicating participants using predictive feedback looked away from releasing the cup sooner, occurred between the no feedback and predictive feedback case only, and only on all cups and the heavy cup. While other cup stiffnesses did not show significant differences, they did favour prediction. Seeing a benefit on the heavy cup, especially one potentially strong



> "Um, I don't know if I'm reading into it too much, but it kind of felt like the hand closed slower. Okay. Like it didn't go as quickly, but that might have been me. I don't know."

Figure 7.14: An interesting disconnect between the recorded velocity for Participant 117's trials with the lightest cup, which was used last in the trial order, and their perception of the interaction. The blue lines are the velocity reported by the servo for all motions with the light cup, normalized to be between 0 and 1. It is notable that upon inspection, if there is a difference the trial on the left, which was without warnings, could be interpreted as slower motion by the higher number of lower spikes. This suggests the internalization of feedback has effects on user perception.

enough to be detected in the grouping of all cups, is therefore curious as the heavy cup was first in order and therefore the prediction warnings are going to fail on early crush events. The statistically significant findings in favour of the predictions for eye-hand arrival latency (EHAL) on drop-off, indicating participants looked to the drop-off location earlier during transport, while statistically significant, are very small changes. EHAL drop-off might have shown benefit from feedback in general and predictions in particular, since we would expect a user that is more assured in their action while moving the cup to look to the placement location sooner. It is possible that the difference in this metric is accounted for in the longer length of the predictive users' grasp phase.

There were large differences in the number of crushes, however. It was expected that the predictive feedback would be nearly as effective as the direct feedback, with the early-learning misses at providing warnings being balanced by the predictions providing signals about the impending crush more in advance for the same threshold setting. However, direct feedback shows a much larger decrease than predictive feedback in the number of times participants crushed the cups (Fig. 7.4) between the feedback and no-feedback cases. This is despite, in most cases, the predictions learning to provide the warnings very quickly and one participant with direct feedback showing significantly more crushes on their second (with warnings) trial. This can be seen in Fig. 7.4 when looking at participants 118's warnings and crushes. They have noticeably more crushes and warnings than other participants in their group, and more than any other participant expressed that they were trying to figure out if they had predictive or direct feedback. This lead to them intentionally crushing the cup on several occasions to study for themselves when the warning was occurring, and if it was changing.

The qualitative data taken in the context of the above quantitative findings provides several interesting insights that would not be available by examination of either data set independently. First, it is worth highlighting that the discussion around the predictive feedback commonly involved the improvement of the signals over time. This was despite participants having noticed the failure of the system to provide the warning they were told it would before they started the warning condition trials. The expectation of the researcher at the outset of this study was that when the device failed to provide the promised warning it would have broken the expectation of the user and at best been regarded with suspicion, but this is not what seems to have happened. The findings do not have a lot to suggest that participants had more positive feelings towards the device in the predictive feedback case, but there is a sense of positivity in the conversation as the signals are discussed as improving and becoming more consistent. This is supported in part by the statistical favouring of the predictions over direct feedback in the eye-hand leaving latency at drop-off for the heavy cup in particular. This cup would have most strongly demonstrated the adapting of the warnings, as the first few crush events would have had no warnings at all, and the benefit to eye-hand leaving latency could indicate increased comfort with the interaction on the part of the user. The statistical findings of the benefit of the predictions in eye-hand arrival latency at pick are also an indication of the increased comfort the user has with the interaction, especially taken along side the relative phase duration's and the manner in which the predictive feedback was discussed. Participants did not seem to notice the failure of predictive feedback to improve the participant's ability to prevent crushes over not having a warning at all. It seems, then, that the failure of a system to act as expected does not irrevocably damage the user's experience of the interaction. The user's engagement with the adaptation of the system could mean that rather than co-adaptation strictly being a hindrance, it could be of benefit to the user's experience.

A valuable takeaway from this study is the disconnect between the external measurements, few appreciable differences in gaze behaviour alongside direct feedback being more successful at preventing participants from crushing the cup, and the internalization of the interactions held by the participants. This is lightly supported in studies such as Williams et al. where a measurable improvement in control from their intervention was not reflected in the participant questionnaires (Hebert & Shehata, 2022; Marasco et al., 2021; Williams, Shehata, Cheng, hebert, et al., 2024). Figures 7.13 and 7.14 are both strong examples of this. Participant 111 (Fig. 7.13), who used direct feedback, was found to have had some difficulties with the hand closing when they wished to open it. This impacted their perception of the entire interaction. The feedback signals, the warnings in particular, had nothing to do with the control of the device. Despite this, they stated that they never received a warning without also hearing the crush sound, but that occurred seven (7) times in thirteen (13) warnings. If the control difficulties were making it harder for the user to stop, or the warnings and crushes resulted from the hand closing when an open was desired, we would expect to hear the participant mention struggling to stop the hand when they heard a warning, not that they never heard a warning without a crush when it was recorded to be incorrect. The case of

Participant 117 (Fig. 7.14), who also had direct feedback, is more suggestive of the addition of the warning sound drawing the user's focus and thus altering their perception of time. This is interesting in light of known results on how attention, cues, and salience modifies the subjective perception of the duration of events and related decision making activities; for the interested reader, a detailed overview of related phenomena is provided by Buonomano (Buonomano, 2017). The velocity settings were not changed, and there's no evidence that the participant was commanding slower motion through the proportional velocity control they were using.

In trials both with and without warnings the participants made comments about feedback signals, both the audible signals provided and other signals inherent to the task, how they used them to learn about and adapt to the task, and their expectations of what would happen. When there was no warning participants discussed the shape of the cup, the visible grip aperture of the hand, and feeling the flexion of their muscles. Often these were related in combination to the break sound in the trials without warnings, and the warning sound in the trials with warnings. When the warning sound was added, it was something participants used in the same way as the break sound; rather than closing the hand until they heard a warning, they would attempt to stop the hand before the warning sound would play. This is perhaps contrary to the expectation of how feedback would be designed to be used, but could be seen as support for the idea that the feedback signals are being used by participants to develop and adapt their own internal models of the interaction. Although feedback is traditionally implemented to allow users to adapt their actions in real-time even in the absence of vision, it may allow users to develop a strong understanding of the dynamics of the device being used, the task being performed, and predict future interactions with the device and objects in the environment (Clemente et al., 2015; Marasco et al., 2021; Shehata et al., 2018b; Thomas et al., 2021). Predictive feedback, such as that provided in this study, may not only allow for the development of strong internal models, but also may allow for the adaptation of this model; hence improving the overall human-machine interaction and device integration. From this, it is possible to conclude that future applications of real-time machine learning to provide feedback to device users should consider methods of providing signals that can learn to adjust what they are signalling about in order to assist in user model development. For many users in this study, for example, feedback about the EMG signals they are sending to the hand may have been more valuable than a signal about the distortion of the cup. Over time, however, a user would likely gain a thorough understanding of the relationship between their muscle feeling and the hand motion, in which case the feedback about the EMG signals would be less helpful, even possibly annoying, than signals about the pressure being applied to an object.

7.3.1 Limitations and Future Directions

One important direction of future studies is to explore the combination of the above phenomenon, the draw to adaptation and the disconnect between the user's experience and externally measurable outcomes. It is possible, for example, that the short-term novelty of experiencing a device adapting and improving in the user's mind could help bridge the gap between what a user is familiar and practiced with and quantitatively determined improvements they might not register in their experience. Such studies would need to continue using mixed methods such as those used here. The use of mixed methods such as this, combining data captured from device use and participant experience is crucial to furthering our understanding of human and machine-learning-enabled robotic interaction. Such studies would also allow researchers to find ways to modify their interventions so the benefits are more strongly experienced by the people they are intended to help.

While participants had a short time to learn to use the devices, the task was fairly simple and example progressions of total task time, shown in Figs. A.1 and A.2 of the Appendix, suggest that user task learning had stabilized somewhat. While this did not prevent complaints and difficulties around control and participants would undoubtedly benefit from more practice, these comments emerging as top-of-mind for participants could still occur in discussions after more practice given the leveled task times. Further studies would benefit from longer term trials to verify this.

This study provides an interesting start to the examination of the interaction between a human and a device that is adapting the signals it is sending the human in real-time. While there are promising indications of many interesting findings, deeper more focused studies to explore each of them would be of great benefit.

This study also faces some limitations concerning the participant group. While the participants were an interesting mix of backgrounds and ages, they can all be characterized as not having limb difference, all having positive views of technology, all being new to EMG, and the interaction between the participant and device being short-term. If the study were to be conducted over a longer term, we expect the novelty of the adaptation of the predictive feedback would have to lead to at least the same level of preventing crushes as the direct feedback; it would not be able to sustain a users interest in a device that is not performing adequately in the long-term. In the case of a participant with limb difference, they may begin the interaction thinking about using the device in the long term, have a frame of reference for that thinking, and find the adaptation of the feedback less novel. With that said, it is still worth noting how for new users, who could also be thought of as users without prior models of the interaction, a system that can be observed as improving over time adds positivity to the frame of reference even if it does not work perfectly initially. It would also be interesting to explore if the positivity of the observable improvement of the device occurs in cases where the device is working within expectation, then works poorly on a related but different task, and then improves on that task.

7.4 Conclusions

This study presents a unique, mixed methods exploration of human prosthesis interaction where the prosthesis is learning to provide signals about a task in real time. Interestingly, the real-time changes to the signals did not impact the interaction negatively. On the contrary, despite not being as successful at preventing users from crushing the cup as they moved it, the real-time learned signals seemed to provide participants some assurance in their actions. This was despite the machine-learned warnings not being as measureably functional in accomplishing the task of signaling the user to prevent crushes. This diss-connect between the user's excperience of the interaction, and the measureable outcomes of the interaction, is something that has been hinted at in ohter studies, and should be explored further. It is possible that great benefit in using a mixed-methods approach such as this study to further explore the separation between performance metrics and user experience in order to find ways to have users *feel* the positive impacts of research in assistive technologies. The assurance participants seemed to have in the adapting feedback may be the bridge between user perception and externally measurably outcomes. These findings, as well as the mixed methods approach used in this study, are a promising direction in facilitating modern advanced prosthetic devices to more effectively work with their users, and reach further towards their full potential.

7.5 Methods

7.5.1 Participants and Recruitment

All experimental protocols were approved by the Research Ethics Board 2 at the University of Alberta, Pro00123026. The study was carried out in accordance with the approved experimental protocols. Informed consent was attained from all participants using a signed letter of consent approved under the experimental protocols.

There were sixteen (16) participants in the study. They were recruited by posters distributed via email. For increased privacy of the participants, very little personally identifying information was collected. From conversations with the participants, we can state that they represented ages from 18 to post-retirement and a wide range of backgrounds. These backgrounds included, but were not limited to, student, researcher, baker, forklift operator, electrical engineer, librarian, physical therapist, and professor.

7.5.2 Experimental Setup

The robot hand used by participants was a single degree of freedom device designed and 3D printed in the lab on MakerBot (New York City, USA) Replicator 2 printers using polylactic acid (PLA), and actuated by a single MX-64AT servo from Dynamixel (Seoul, South Korea) (Wells et al., 2020). The hand could open and close, and the thumb and all four fingers moved together when they actuated. The hand was attached to a brace that was designed and 3D printed in the lab to position the robot hand in line with and medially to the participant's biological hand, and restrict the participant's wrist motion (Hallworth et al., 2022). Control of the hand was achieved using a Myo armband from Thalmic Labs (Kitchener, Canada), which contains eight (8) EMG sensors and broadcasted via Bluetooth to the associated USB dongle. The hand and Myo armband were both operated through BrachI/Oplexus software, which was developed in the lab to operate a different 3D printed robot arm, the Bento arm, that utilizes the same servos (Dawson et al., 2014; Dawson et al., 2020). This software allowed researchers to configure various settings of the power hand and the Myo armband and facilitated various means of controlling the power hand, Which included the Myo armband. BrachI/Oplexus also had connections to the Python code used to run the temporal-difference learning agent in Edwards et al. 2016 (Edwards, Dawson, et al., 2016).

The objects the participants moved were 3D-printed cups. These cups were printed in NinjaFlex (NinjaTek, Lititz, USA) material to allow them to be distorted without breaking. The cups measure 72mm tall, with a 45mm lower outer diameter, a 61mm upper outer diameter, and a 4mm thick wall. Three (3) different stiffnesses of cups were used. This was achieved by varying the amount of infill used when printing each up. The stiffest cup used a 25% infill, the middle cup was a 10% infill, and the light cup was a 5% infill. Four (4) motion capture markers were placed around the rim of the cups. One of these markers was raised. If the raised marker is taken to be the origin, and placed at the top of the upper ring looking down, the remaining markers were placed at 60° counter-clockwise, 65° clockwise, and 160° clockwise from the master. The cups were tested for how they retained their stiffness over repeated use. Each cup was crushed 200 times to a specified hand-closed position read from the position of the servo. A trend line across the load of the 200 crushes showed little to no degradation.

To determine the threshold settings to use for each of the stiffnesses of cup, first, the normalized load reading of the hand is recorded for five (5) minutes while the hand is open. The most extreme load value in the correct direction recorded over that time is taken to be the $noiseLevel$. It is important to note that the normalized neutral load is around 0.5 since the load signal is directional. In this case, the thresholds for the crushes and warnings are below 0.5 as a result of the orientation of the servo. Each of the three (3) stiffnesses of cup is then recorded being crushed autonomously as far as the torque limits of the hand will allow thirty (30) times. The most extreme normalized load recorded of the thirty (30) crushes is taken to be the $cupMax$ for that cup. The crush threshold is set to $crushThreshold = noiseMax + 3/4(cupMax - noiseLevel)$ for each of the three cup stiffnesses. Here we subtract the noiseLevel from the cupMax since fully closed pressure is read from the servo as less than 0.5, so a smaller number represents greater force. The warning threshold was set to $warningthreshold = noiseMax + 2/3(crushThreshold - noiseLevel)$ for each of the three cup stiffnesses. Again, the subtraction is a result of the load reading being directional and the physical orientation of the servo in the hand. So the crush threshold for each cup is set to 75% of the maximum range of the normalized load signal, and the warning is set to 66% of the crush threshold, which is equivalent to 50% of the maximum range of the normalized load signal. Here, this means a break threshold of 0.13 and a warning threshold of 0.25. These values were found to be visually to be functionally appropriate to allow each cup to visibly distort differently from each other, and provide an achievable stop at the warning.

7.5.3 Gaze and Movement Analysis

The Gaze and Movement Assessment (GaMA) protocol uses motion capture and eye tracking to quantify the movement quality and visual attention exhibited by participants as they interact with objects. A 10-camera OptiTrack Flex 13 motion capture system (Natural Point, Corvallis, USA) was used to capture participant movements and task objects at a sampling rate of 120 Hz. Four motion capture markers were placed on the simulated prosthesis hand via a rigid plate (shown in Fig. 7.1), four markers were placed on the cups (also shown in Fig. 7.1), and 6 markers were placed on the table. A Pupil Core head-mounted binocular eye tracker (Pupil Labs GmbH, Berlin, Germany) captured participants' pupil positions at 120 Hz. Participants each performed two gaze calibrations (outlined by Lavoie et al.) to facilitate the generation of gaze vectors. These calibrations were carried out at the beginning and end of each trial (Lavoie et al., 2018).

After motion capture and eye tracking data were collected, gaps in the motion capture data were filled and the data was then filtered. The pupil position data were similarly corrected and filtered. The motion capture and pupil data were then synchronized, and gaze vectors were generated. As per Valevicius et al., 3D objects were created to represent the simulated prosthesis hand, cup, and cart using the motion capture markers placed on each (Valevicius et al., 2018). Then, cup movements were segmented into five phases of reach, grasp, transport, release, and a home phase (which is not relevant to the data) using movements of the hand and cup objects. These phases were used in the calculation of metrics.

A total of 62 metrics were calculated as per previous works (Lavoie et al., 2018; Valevicius et al., 2018). Of note were eye-hand latency measures, which were calculated at instances of phase transitions—at the start of transport (referred to as "pick-up" by Lavoie et al.) and at the end of transport (referred to as "drop-off"). Eye-hand arrival latency and eye-hand leaving latency were calculated as follows:

- **Eye-hand arrival latency at pick-up** is defined as the start of the last continuous look to the cup or its pick-up location. If there are no looks to either the cup or its location during reach or grasp, then this metric is defined as the first look to the hand or cup being moved during transport. The value was positive if the participant looked to the cup or location before pick-up, and negative if they looked at the cup or hand after pick-up.

- **Eye-hand arrival latency at drop-off** is defined as the start of the last continuous look to the cup or its drop-off location before transport end (positive value) or after transport end (negative value).

- **Eye-hand leaving latency at pick-up** is defined as the end of the last continuous look to the cup or its pick-up location. The value was positive if the gaze left the cup or pick-up location after the start of transport, and positive if the gaze left before the start of transport.

- **Eye-hand leaving latency at drop-off** is defined as the end of the last continuous look to the cup or its drop-off location before transport end (positive value) or after transport end (negative value).

These metrics require the phase duration metrics in order to properly assess their meaning. The latency metrics along side the phase duration and fixation metrics were selected from the 62 metrics available as they were considered to be the most suitable to show changes in user behaviour resulting from the different feedback types (Williams, Shehata, Cheng, hebert, et al., 2024).

7.5.4 Machine Learning Methods

All robot signals: load, position, and velocity, were normalized before use with the learning code. Temporal-difference (TD) learning was selected as the machine learning technique to adapt the warning signal in real time. Specifically, TD(0) was selected for its previously demonstrated ability to learn quickly from sensor data in real time in similar domains (Edwards, Dawson, et al., 2016; A. S. R. Parker et al., 2022; A. S. R. Parker et al., 2019a; Pilarski et al., 2022; Pilarski, Dawson, Degris, Carey, et al., 2013). TD(0) is a method from the field of reinforcement learning. It learns a temporally extended expectation, a prediction, of a signal from the real-time information it is provided. Often this signal is a reward, but it does not have to be. In those cases, any

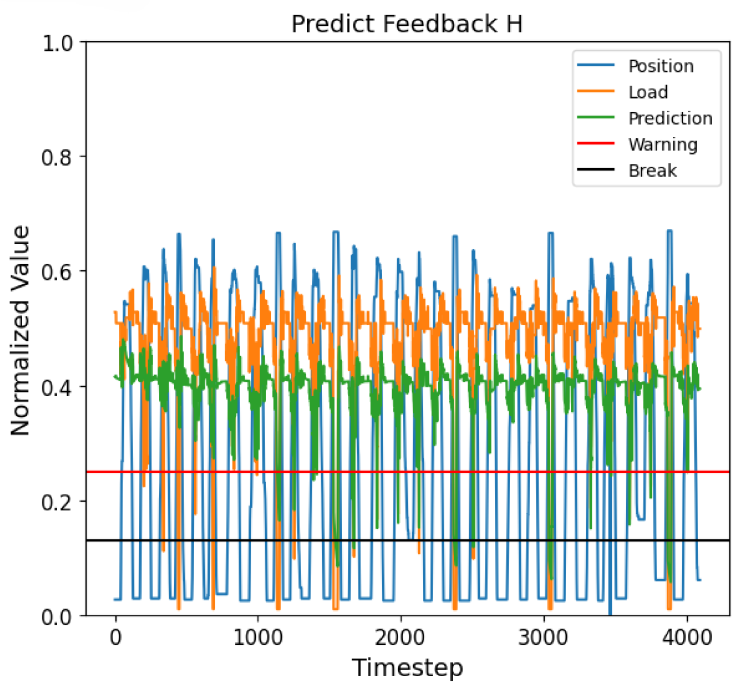


Figure 7.15: Above the machine-learned predictions over a single participant's entire warning trial with the stiffest cup (the first one used) can be seen. Most notably the predictions, in green, can be seen to start low and rise quickly relative to the load, in orange, is being predicted. The load does not change over the course of the trial, only in response to participant motion, which can be interpreted from the position which is shown in blue.

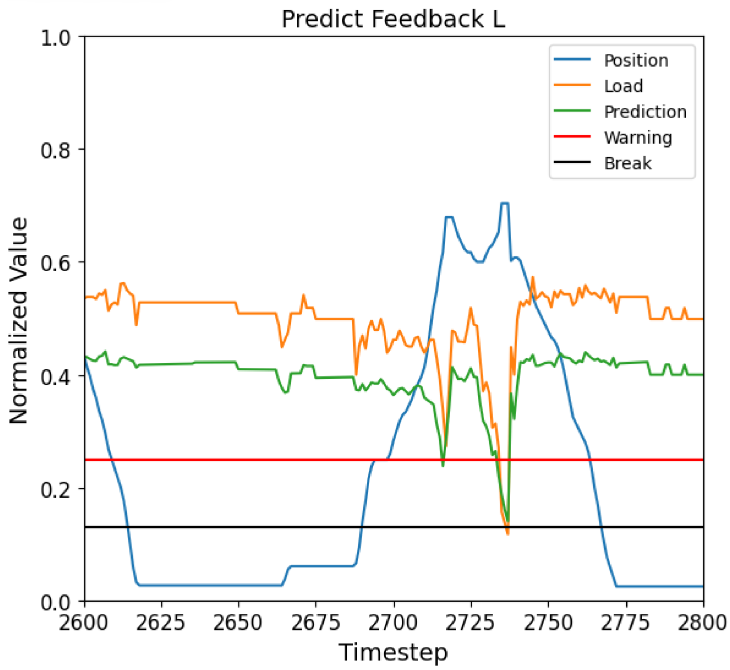


Figure 7.16: The prediction, in green, can be seen to be a similar shape to the load, in orange, in this closer view. Moreover, the participant can be seen to open their hand and close it again by the U-shape in the position, shown in blue. This was presumably in reaction to the warning crossing the threshold, shown in red. Despite this abrupt change in motion, the prediction still moves similarly to the load.

signal in the environment can be the target, called the return or cumulant, and the machine learning agent is learning a general value function (GVF) (Sutton et al., 2011). A GVF encapsulates the expectation the agent has of the future cumulant for the state it is in based on its previous experience. The state, S_t, is a crucial piece of the agent; it is how the agent associates the predictions it is making with what is happening in the world as it experiences it. The learning agent here uses a Selective Kanerva Coded state identified by Travnik and Pilarski (2017) as being suitable to settings such as this (Travnik & Pilarski, 2017). The total usable ranges of the position, velocity, and load of the servo are normalized and become the axis of the state space. Within those axes, 1000 points, called prototypes, are randomly placed. Signals relating to the axis, the position, velocity, and load of the servo, are read and normalized in real-time, and from a point within the axis. The 50 closest prototypes to the current normalized sensor readings are the active features of the state space and will be used for the current time step's predictions, and learning.

$$\delta_t = r_{t+1} + \gamma \mathbf{w}_t^T \mathbf{x}(S_{t+1}) - \mathbf{w}_t^T \mathbf{x}(S_t) \tag{7.1}$$

$$\mathbf{w}_{t+1} = \mathbf{w}_t + \alpha \delta_t \mathbf{x}(S_t) \tag{7.2}$$

The learning is executed using Eqs. 7.1 and 7.2. The first equation, Eq. 7.1, produces the temporal difference error δ, and Eq. 7.2 is the learning update. In the equations, $X(S)$ is the active features of the state representation at time t_{t+1} and t respectively. The values learned are stored in the weight vector, $\mathbf{w}$. Together, $\mathbf{w}_t^T \mathbf{x}(S_t)$ is the value learned prediction for the current time step, t. The learning rate, α, was set to $0.01/number\_of\_active\_features$. As there are 50 active features, the final value of α is $\alpha = 0.0002$. The discount factor, γ, was set to $\gamma = 0.95$

As executed here, the predictions do not add increased temporal information to the signals the participant is receiving. The core concern here was that the machine learning could learn to quickly provide stable signals from real-time observations in the short term. The results of this can be seen in Figs. 7.15 and 7.16. In Fig. 7.15, the green line demonstrates the prediction being learned from its initialized value. The weight vector was initialized to start at approximately the 0 point of the normalized load, $noiseMin/(1-\gamma)/number\_of\_active\_features$ Specifically, the general progression of the prediction in green can be seen to reach lower over repeated interactions. A more detailed view of this can be seen in Fig. 7.16. Here the normalized prediction, again in green, normalized load, in orange. In Fig. 7.16 the participant can be seen to close the hand and receive a warning, and they respond by opening the hand and closing it again. Despite this behaviour, the prediction still predicts the load effectively.

7.5.5 Qualitative Recording

Two (2) semi-structured interviews of approximately ten (10) minutes each were conducted. One after each trial condition. The questions, included in Appendix A.1.1, were designed with a mind to focus the participant on the area of interest of the study without priming them. They start very broad and over the course of the interview the focus on the elements of the study related to the research question increased. These questions are meant as guideliness to conducting a conversation with the participant about the interaction they just completed. Each primary question had several probing questions meant to elicit further information from the participants if they were giving short answers. These questions were not followed precisely in every case; when a participant mentioned something of interest to the study the interviewer attempted to pursue it. The interviews were recorded on a Sony ICD-PX470 stereo digital voice recorder. The audio files were autonomously transcribed using Rev.com. The automated transcripts were then checked and corrected by hand by the research team member who would go on to conduct the coding for analysis.

7.5.6 Experimental Design and Flow

The full checklist followed by the study team can be found in Appendix A.1.2. After the participants arrived, were greeted, and informed consent was acquired, they were helped into the simulated prosthesis brace. Following this, the EMG read from the Myo armband was configured in BrachI/Oplexus for the participant to operate the device with proportional velocity control. First, the participant flexed and extended their wrist several times and the two strongest signals were selected and the thresholds set. The participant then demonstrated opening and closing the hand with those settings. If the participant reported struggling, the thresholds were adjusted. Following this, the participant was asked to open and close the hand several times of the left cup target, then the right cup target, and then over the home position. The EMG thresholds were again adjusted according to observations of use and the participant's reports of operation. Once the EMG was configured appropriately, the eye tracking was configured.

Next, the participant began their first training session. This had the participant use the robot hand, controlled by proportional EMG, move a stiff cup from the left target to the right target after hearing the queue from GaMA, assume the home position, and move the cup back again upon hearing the GaMA queue, with no audible feedback about the pressure they are applying to the cup. The participant was not only practicing the use of the robot hand but also the motion to be used in the trial and the queues being sent to them to move. This stiff cup could not be appreciably distorted by the torque limits placed on the robot hand. After approximately two minutes of training, a gaze calibration was conducted for GaMA.

At this point, the trial for the first condition began where the cups are now able to be deformed by the robot hand, and a crush noise will play if the normalized load reading from the servo surpasses the thresholds outlined above. The stiffest cup was placed on the left marker, and the participant was asked to move it thirty (30) times. The participant was not told this was the stiffest cup, nor were they permitted to interact with it with their biological hand. For one motion in five (1 in 5), the participant was asked to squeeze the cup as hard as the hand would allow. This was done for the learning experience of both the participant and when applicable the learning agent. A motion in the trial was one-directional, so the first motion was moving the cup from left to right. At the end of the motion, the participant assumed the home position, waiting for the queue from GaMA, and moved the cup back for the second motion. At the end of thirty (30) motions the participant was asked to rest for three (3) minutes. After the rest, the above procedure was done with the medium stiffness cup, and then the light stiffness cup. In all cases, the participant was given no prior knowledge about the stiffness of the cup. Following the completion of all three (3) stiffnesses of cup, a gaze calibration was performed for GaMA. The participant was then asked if they wanted water, and seated while the interview was conducted.

The above block was then repeated for the second trial condition. This condition introduced a warning noise, based on the thresholds outlined above, in addition to the crush noise. Half of the participants had a warning based on the normalized load signal reported by the servo, and half had a machine-learned prediction where the cumulant was the normalized load signal reported by the servo. After the second interview, the participant was helped out of the simulated prosthesis brace and thanked.

7.5.7 Data Analysis

Plots of the GaMA metrics were generated and studied in various groupings of participants and cups in order for the researcher doing the analysis to familiarize themselves with the quantitative data. Automated transcripts were generated from the voice recordings using Rev.com. These transcripts were then corrected by hand, in part as a means of familiarization with the qualitative data. The resulting verbatim transcripts were the qualitatively coded using Quirkos qualitative analysis software. After an initial coding pass from the transcripts, the codes were checked for consistent application and thoroughly examined in separation from the transcripts.

The resultant qualitative codes were studied using the side-by side and overlap views available in Quirkos. Findings from these analyses were noted, and expressed as questions to ask of the full quantitative and qualitative data set. Matches and differences between the qualitative and quantitative data were explored in both data sets and findings were generated that fit the narrative of both data sets. These became the theme elements, which were organized into overarching themes and then compared again to both quantitative and qualitative data to ensure they resonated appropriately.

Chapter 8

Conclusion

Assistive technologies strategies meant to assist users by providing functions such as those that allow them to participate in tasks of daily living. Some assistive technologies are independent agents that are trained to provide assistance, such as seeing-eye dogs, and others are tools that users wield such as wheelchairs and, typically, upper limb prostheses. It is important that assistive technology is developed and applied with a focus on patient-centered care and outcomes in order to best meet that patients specific needs.

Prosthetic limbs are used to provide increased function to people with limb differences. Modern prostheses, particularly for the arm and hand, have many points of articulation that can mimic the motions of a biological limb. Most of this function, however, cannot be accessed by users when the prosthesis is deployed because of challenges mapping the motion available in the device to the control signals available from the user. Adding agency to prosthetic limbs through the application of machine learning gives us a pathway of research to improve users' ability to achieve the full potential of these devices. A prosthetic upper limb with agency can be viewed as the human's collaborative partner, which allows research into ways of improving their collaborative, or joint, action.

8.1 Summary of Contributions

There are several promising steps toward improving assistive technologies that were spearheaded by the work in this dissertation. The first major contribution is championing the reasons and benefits of viewing upper-limb prostheses, and indeed many other assistive technologies, as partners to users rather than resigning them to simply being tools. The capacity of a tool to provide increased function for a user will always be limited by the inability of the tool to adapt to the user, new knowledge, or changes in the task. In this work, we started by outlining why it makes sense to view human-prosthesis interaction as two agents collaborating to achieve an outcome, which was done in Chapter 2. We then went on to discuss the framework of communicative capital, outlined in Chapter 3, which provides the tools to think about how to add collaborative potential and suggests ways to unlock it.

The second major contribution of this work is the demonstration of temporal-difference methods learning and adapting feedback about a task. To our knowledge, this is the first time such feedback has been used to signal a human controlling a robot in the real world. Not only can simple methods learn something that is of value to a user with vastly superior intellect, but they can be learned in real-time, and that learning does not negatively impact the user's experience. Chapter 4 begins these contributions by demonstrating that temporal-difference learning methods can learn something about a task that assists a human that is making decisions about the system. Following this, Chapter 5 provides evidence that such learning can be done in real time. This finding opens the door to research a myriad of ways to use real-time machine learning to signal human users and improve their function and capacity to accomplish tasks. The primary method outlined in this dissertation we called *Pavlovian signalling*, which was applied in Chapter 6 and outlined and formalized in Chapter 5.

As a final major contribution, this dissertation includes the first mixed-methods study to leverage rich data from both humans and machines to study continual learning on the part of both. It would have been reasonable to think that if a signal coming from a machine was changing over time it would cause a user to become disenfranchised with the device; they would lose trust in the device's usefulness and ultimately blame it for difficulties and discard it. In the short term, however, it appears that positive changes to the signals users receive as feedback are given strong weight as improvement to overall function. Chapter 6 begins the this exploration with the addition of qualitative methods to the analysis of human interaction with real-time learning agents. It suggests that the user does not give up on the agent as a result of early failures or learning, and that the signals provided by the agent are used in ways the designer did not expect. Following this, Chapter 7 conducts a first-of-its-kind mixed methods study examining the interaction using rich human and machine data from human use of a simulated prosthesis with real-time adapting signals about the task. This chapter highlights the existence of a disconnect between user experience and performance metrics. Coupled with the finding that participants are drawn to addaptation, this suggests that real-time adaptation may be a path to connecting the actual performance of a system with the users interpretation. The approach used in Chapter 7 provides valuable insights into the interaction between a human using a wearable robot and a machine-learning agent that learns and adapts feedback signals in real time. Such methods gave us far greater insight into the interaction than could have been gained by either method alone, or via the use of less rich data from either part. There were also insights in both mixed-method studies that were surprising to experts in multiple fields, which further highlights the importance of more research in human-machine interaction that utilizes such methods.

There is tremendous potential to improve the functionality available to users of prosthetic limbs through the applications of real-time machine learning such as TD methods from the field of RL. Since these methods can adapt from experience in real-time, they are well aligned with the goal of achieving patient-focused care. The findings of this dissertation, along with the demonstration of the value of mixed-methods studies, chart a path toward fostering strong human-machine collaborations. These collaborations are expected to assist users in utilizing more of the potential function available in modern robotic prostheses. As future prostheses add functions, techniques that facilitate collaboration may scale better than other currently researched approaches. If the device and user collaborate seamlessly, then added functions should not pose a challenge to the pair. This approach has the potential to revolutionize not just prosthetic limbs, but other assistive robotics in rehabilitation, and human-machine interaction in society in general.

8.2 Future Directions

There are three core directions that are natural continuations of this work. First, it is interesting to find that changing the signals that a user receives while they are doing a task in real time does not cause them confusion or lead them to outright ignore the feedback they are receiving as unreliable. The opposite may be what is happening; the adaptation users notice offers benefits to the interaction as it is interpreted by users. This finding should be further explored using studies to find the limits of this both in terms of the length of the benefit and the strength of it. Finding how long this effect can improve user experience of the interactions would give designers a strong sense of how long learning on the part of the system can take before it becomes a burden. It will also be interesting to explore how much "failure" the adaptation can draw focus away from, and in what conditions. It may be that the machine learning agent has to adapt in a way that is beneficial from the perspective of the participant, or it may be that simply changing over time will give users the impression of improvement. These are important to research in people with limb differences, in the case of prosthesis partners, as their views on having to live with the device will differ vastly from participants without limb difference. It will also be interesting to see what other interactions benefit from adapting the feedback in real-time, and how this can be tuned to improve human function and ability in a multitude of areas.

Perhaps the greatest impact would result from the continuation and refinement of studies and methods that mix rich quantitative and qualitative data to explore and develop human-machine interactions. The most interesting findings of this work would not have been possible without rich qualitative data in conjunction with rich quantitative data. Chapter 7 highlights how user experience does not necessarily sync with measurable metrics, and that is a prime example of of the insights that would be unavailable without rich data from both user and device. This disconnect is something that has been anecdotally noticed in other studies such as Williams, Shehata, Cheng, hebert, et al. (2024) where the results of their survey did not favour the measurably better intervention, but has not been expressly addressed. Understanding this the difference between external measures and user experience directly and how to work with it, for example, could be vital to getting users on board with new interventions that can measurably improve their ability to function with their assistive technologies. More studies into this phenomenon, studies which would be far less informative if they were not done with such a rich mixing of data, could yield the keys to achieving strong human-machine collaborations in rehabilitation and beyond.

Interdisciplinary research will be key to advancing these methods in the future. Algorithmic advances in machine learning will need to be tested on robots and with intended users in order to determine the effects that different methods of machine learning have on the perceptions of users and in real world applications. It could be important to know how users notice and internalize changes to the machine agent, such as the way it steps through the state space, the speed at which it learns, or perhaps the model the agent builds. The combination of the expertise of clinicians, engineers, users, and researchers across fields is crucial for developing even greater interventions for patients requiring assistive technologies to ensure the needs of all of the involved agents are met. This collaborative approach, encompassing mixed-methods studies containing rich data from both physical systems and human users, holds great potential for progress. Further, there is the potential for research in one area to assist research in another. As research is pushed further into framing human-machine interaction as joint action, methods will develop that will assist human-human joint action research. In this way joint agency, be it between a human and a human or a human and a machine, will advance and continue to generate knowledge and user-centric solutions that add function and value to humanity.

References

Abramson, J., Ahuja, A., Brussee, A., Carnevale, F., Cassin, M., Clark, S., Dudzik, A., Georgiev, P., Guy, A., Harley, T., Hill, F., Hung, A., Kenton, Z., Landon, J., Lillicrap, T. P., Mathewson, K. W., Muldal, A., Santoro, A., Savinov, N., . . . Zhu, R. (2020). Imitating interactive intelligence. *arXiv:2012.05672 [cs.LG]*. https://doi.org/10.48550/arXiv.2012.05672.

Amershi, S., Cakmak, M., Knox, W. B., & Kulesza, T. (2014). Power to the people: The role of humans in interactive machine learning. *AI Magazine*, *35*, 105–120. https://doi.org/10.1609/aimag.v35i4.2513.

Antfolk, C., D'alonzo, M., Rosén, B., Lundborg, G., Sebelius, F., & Cipriani, C. (2013). Sensory feedback in upper limb prosthetics. *Expert Review of Medical Devices*, *10*(1), 45–54. https://doi.org/10.1586/erd.12.68.

Argall, B. D., Chernova, S., Veloso, M., & Browning, B. (2009). A survey of robot learning from demonstration. *Robotics and Autonomous Systems*, *57*(5), 469–483. https://doi.org/10.1016/j.robot.2008.10.024.

Arthanat, S., Bauer, S. M., Lenker, J. A., Nochajski, S. M., & Wu, Y. W. B. (2007). Conceptualization and measurement of assistive technology usability. *Disability and Rehabilitation: Assistive Technology*, *2*(4), 235–248. https://doi.org/10.1080/17483100701343665.

Ballinger, C., & Payne, S. (2000). Discourse analysis: Principles, applications and critique. *British Journal of Occupational Therapy*, *63*(12), 566–572. https://doi.org/10.1177/030802260006301202.

Belfiore, M. (2010). *The department of mad scientists: How DARPA is remaking our world, from the internet to artificial limbs*. Harper.

Bicho, E., Erlhagen, W., Louro, L., & e Silva, E. C. (2011). Neuro-cognitive mechanisms of decision making in joint action: A human–robot interaction study. *Human Movement Science*, *30*(5), 846–868. https://doi.org/10.1016/j.humov.2010.08.012.

Biddiss, E. A., & Chau, T. T. (2007a). Upper limb prosthesis use and abandonment: A survey of the last 25 years. *Prosthetics and Orthotics International*, *31*(3), 236–257. https://doi.org/10.1080/03093640600994581.

Biddiss, E. A., & Chau, T. T. (2007b). Upper-limb prosthetics: Critical factors in device abandonment. *American Journal of Physical Medicine & Rehabilitation*, *86*, 977–987. https://doi.org/10.1097/PHM.0b013e3181587f6c.

Boser, Q. A. (2019). *Characterizing the visuomotor behaviour of upper limb body-powered prosthesis users* (Master's thesis). University of Alberta Libraries. https://doi.org/10.7939/r3-jf4r-9056.

Brady, J. V., & Jonsen, A. R. (2014). The Belmont Report. Ethical principles and guidelines for the protection of human subjects of research. *The Journal of the American College of Dentists*, *81*, 4–13.

Braun, V., & Clarke, V. (2021a). Can I use TA? should I use TA? should I not use TA? comparing reflexive thematic analysis and other pattern-based qualitative analytic approaches. *Counselling and Psychotherapy Research*, *21*(1), 37–47. https://doi.org/10.1002/capr.12360.

Braun, V., & Clarke, V. (2021b). One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative Research in Psychology*, *18*(3), 328–352. https://doi.org/10.1080/14780887.2020.1769238.

Brennan, A. A., & Enns, J. T. (2015). What's in a friendship? Partner visibility supports cognitive collaboration between friends. *PloS one*, *10*(11), 1–15. https://doi.org/10.1371/journal.pone.0143469.

Brenneis, D. J. A., Parker, A. S. R., Johanson, M. B., Butcher, A., Davoodi, E., Acker, L., Botvinick, M. M., Modayil, J., White, A., & Pilarski, P. M. (2022). Assessing human interaction in virtual reality with continually learning prediction agents based on reinforcement learning algorithms: A pilot study. *Adaptive and Learning Agents (ALA) Workshop at AAMAS 2022*, 1–8. https://doi.org/10.48550/arXiv.2112.07774.

Brooks, R. (2002). *Flesh and machines: How robots will change us*. Pantheon Books.

Buonomano, D. (2017). *Your brain is a time machine: The neuroscience and physics of time*. WW Norton & Company.

Butcher, A., Johanson, M. B., Davoodi, E., Brenneis, D. J. A., Acker, L., Parker, A. S. R., White, A., Modayil, J., & Pilarski, P. M. (2022). Pavlovian signalling with general value functions in agent-agent temporal decision making. *Adaptive and Learning Agents (ALA) Workshop at AAMAS 2022*. https://doi.org/10.48550/arXiv.2201.03709.

Candidi, M., Curioni, A., Donnarumma, F., Sacheli, L. M., & Pezzulo, G. (2015). Interactional leader–follower sensorimotor communication strategies during repetitive joint actions. *Journal of The Royal Society Interface*, *12*(110), 20150644. https://doi.org/10.1098/rsif.2015.0644.

Cao, K., Lazaridou, A., Lanctot, M., Leibo, J. Z., Tuyls, K., & Clark, S. (2018). Emergent communication through negotiation. *arXiv:1804.03980 [cs.AI]*, 1–15. https://doi.org/10.48550/arXiv.1804.03980.

Carmena, J. M. (2012). Becoming bionic. *IEEE Spectrum*, *49*(3), 24–29. https://doi.org/10.1109/MSPEC.2012.6156860.

Castellini, C., Artemiadis, P., Wininger, M., Ajoudani, A., Alimusaj, M., Bicchi, A., Caputo, B., Craelius, W., Dosen, S., Englehart, K., et al. (2014). Proceedings of the first workshop on peripheral machine interfaces: Going beyond traditional surface electromyography. *Frontiers in Neurorobotics*, *8*, 22. https://doi.org/10.3389/fnbot.2014.00022.

Castellini, C., & Van Der Smagt, P. (2009). Surface EMG in advanced hand prosthetics. *Biological Cybernetics*, *100*(1), 35–47. https://doi.org/10.1007/s00422-008-0278-1.

Chao, C., Cakmak, M., & Thomaz, A. L. (2010). Transparent active learning for robots. *Proceedings of the 5th ACM/IEEE International Conference on Human-Robot Interaction*, 317–324. https://doi.org/10.1109/HRI.2010.5453178.

Cheng, K. Y., Chapman, C. S., & Hebert, J. S. (2022). Spatiotemporal coupling of hand and eye movements when using a myoelectric prosthetic hand. *2022 International Conference on Rehabilitation Robotics (ICORR)*, 1–6. https://doi.org/10.1109/ICORR55369.2022.9896491.

Chernova, S., & Thomaz, A. L. (2014). *Robot learning from human teachers* (Vol. 8). Morgan & Claypool Publishers LLC. https://doi.org/10.1007/978-3-031-01570-0.

Childress, D. S. (1985). Historical aspects of powered limb prostheses. *Clinical Prosthetics Orthotics*, *9*(1), 2–13. https://www.oandplibrary.org/cpo/1985_01_002.asp.

Clark, A. (2008). *Supersizing the mind: Embodiment, action, and cognitive extension*. Oxford University Press.

Clemente, F., D'Alonzo, M., Controzzi, M., Edin, B. B., & Cipriani, C. (2015). Non-invasive, temporally discrete feedback of object contact and release improves grasp control of closed-loop myoelectric transradial prostheses. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *24*(12), 1314–1322. https://doi.org/10.1109/TNSRE.2015.2500586.

Cordella, F., Ciancio, A. L., Sacchetti, R., Davalli, A., Cutti, A. G., Guglielmelli, E., & Zollo, L. (2016). Literature review on needs of upper limb prosthesis users. *Frontiers in Neuroscience*, *10*, 209. https://doi.org/10.3389/fnins.2016.00209.

Couraud, M., Cattaert, D., Paclet, F., Oudeyer, P.-Y., & Rugy, A. D. (2018). Model and experiments to optimize co-adaptation in a simplified myoelectric control system. *Journal of Neural Engineering*, *15*(2), 026006. https://doi.org/10.1088/1741-2552/aa87cf.

Crandall, J. W., Oudah, M., Tennom, F. I.-O., Abdallah, S., Bonnefon, J.-F., Cebrian, M., Shariff, A., Goodrich, M. A., & Rahwan, I. (2018). Cooperating with machines. *Nature Communications*, *9*(233), 12. https://doi.org/10.1038/s41467-017-02597-8.

Creswell, J. W. (2021). *A concise introduction to mixed methods research*. SAGE Publications.

Cultural General Intelligence Team, Bhoopchand, A., Brownfield, B., Collister, A., Lago, A. D., Edwards, A., Everett, R., Frechette, A., Oliveira, Y. G., Hughes, E., Mathewson, K. W., Mendolicchio, P., Pawar, J., Pislar, M., Platonov, A., Senter, E., Singh, S., Zacherl, A., & Zhang, L. M. (2022). Learning robust real-time cultural transmission without human data. https://doi.org/10.48550/arXiv.2203.00715.

Dalrymple, A. N., Roszko, D. A., Sutton, R. S., & Mushahwar, V. K. (2020). Pavlovian control of intraspinal microstimulation to produce over-ground walking. *Journal of Neural Engineering*, *17*(3), 036002. https://doi.org/DOI10.1088/1741-2552/ab8e8e.

Dawson, M. R., Parker, A. S. R., Williams, H. E., Shehata, A. W., Hebert, J. S., Chapman, C. S., & Pilarski, P. M. (2024). Joint action is a framework for understanding partnerships between humans and upper limb prostheses. *International Conference on Biomedical Robotics and Biomechatronics (BioRob); also arXiv:2212.14124 [cs.HC]*. https://doi.org/10.48550/arXiv.2212.14124.

Dawson, M. R., Sherstan, C., Carey, J. P., Hebert, J. S., & Pilarski, P. M. (2014). Development of the bento arm: An improved robotic arm for myoelectric training and research. *Proc. of MEC14 Symposium*, *14*, 57–61. https://conferences.lib.unb.ca/index.php/mec/issue/view/2/1.

Dawson, M. R., Williams, H. E., Murgatroyd, G., Hebert, J. S., & Pilarski, P. M. (2020). BrachIOplexus: Myoelectric training software for clinical and research applications. *Proc. of MEC20 Symposium*, *20*, 87–90. https://conferences.lib.unb.ca/index.php/mec/issue/view/9/9.

Dewdney, C. (1998). *Last flesh: Life in the transhuman era*. HarperCollins.

Dewey, J. (1908). What does pragmatism mean by practical? *The Journal of Philosophy, Psychology and Scientific Methods*, *5*(4), 85–99. https://doi.org/10.2307/2011894.

Doidge, N. (2007). *The brain that changes itself: Stories of personal triumph from the frontiers of brain science*. Viking.

Edwards, A. L. (2016). *Adaptive and autonomous switching: Shared control of powered prosthetic arms using reinforcement learning* (Master's thesis). University of Alberta Libraries. https://doi.org/10.7939/R35Q4RX49.

Edwards, A. L., Dawson, M. R., Hebert, J. S., Sherstan, C., Sutton, R. S., Chan, K. M., & Patrick, P. M. (2016). Application of real-time machine learning to myoelectric prosthesis control: A case series in adaptive switching. *Prosthetics and Orthotics International*, *40*(5), 573–581. https://doi.org/10.1177/0309364615605373.

Edwards, A. L., Hebert, J. S., & Pilarski, P. M. (2016). Machine learning and unlearning to autonomously switch between the functions of a myoelectric arm. *2016 6th IEEE International Conference on Biomedical Robotics and Biomechatronics (BioRob)*, 514–521. https://doi.org/10.1109/BIOROB.2016.7523678.

Faridi, P., Mehr, J. K., Wilson, D., Sharifi, M., Tavakoli, M., Pilarski, P. M., & Mushahwar, V. K. (2022). Machine-learned adaptive switching in voluntary lower-limb exoskeleton control: Preliminary results. *2022 International Conference on Rehabilitation Robotics (ICORR)*, 1–6. https://doi.org/10.1109/ICORR55369.2022.9896611.

Feil-Seifer, D., & Matarić, M. J. (2011). Socially assistive robotics. *IEEE Robotics & Automation Magazine*, *18*(1), 24–31. https://doi.org/10.1109/MRA.2010.940150.

Fishman, G. A. (2003). When your eyes have a wet nose: The evolution of the use of guide dogs and establishing the seeing eye. *Survey of Ophthalmology*, *48*(4), 452–458. https://doi.org/10.1016/S0039-6257(03)00052-3.

Flanagan, J. R., Vetter, P., Johansson, R. S., & Wolpert, D. M. (2003). Prediction precedes control in motor learning. *Current Biology*, *13*(2), 146–150. https://doi.org/10.1016/S0960-9822(03)00007-1.

Fougner, A., Stavdahl, Ø., Kyberd, P. J., Losier, Y. G. L., & Parker, P. A. (2012). Control of upper limb prostheses: Terminology and proportional myoelectric control—A review. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *20*(5), 663–677. https://doi.org/10.1109/TNSRE.2012.2196711.

Gallivan, J. P., & Chapman, C. S. (2014). Three-dimensional reach trajectories as a probe of real-time decision-making between multiple competing targets. *Frontiers in Neuroscience*, *8*, 215. https://doi.org/10.3389/fnins.2014.00215.

Geary, J. (2002). *The body electric: An anatomy of the new bionic senses*. Rutgers University Press.

Goldkuhl, G. (2012). Pragmatism vs interpretivism in qualitative information systems research. *European Journal of Information Systems*, *21*(2), 135–146. https://doi.org/10.1057/ejis.2011.54.

González, D. S., & Castellini, C. (2013). A realistic implementation of ultrasound imaging as a human-machine interface for upper-limb amputees. *Frontiers in Neurorobotics*, *7*, 17. https://doi.org/10.3389/fnbot.2013.00017.

Griffith, S., Subramanian, K., Scholz, J., Isbell, C. L., & Thomaz, A. L. (2013). Policy shaping: Integrating human feedback with reinforcement learning. In C. Burges, L. Bottou, M. Welling, Z. Ghahramani, & K. Weinberger (Eds.), *Advances in neural information processing systems*. Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2013/file/e034fb6b66aacc1d48f445ddfb08da98-Paper.pdf.

Hallworth, B. W., Shehata, A. W., Dawson, M. R., Sperle, F., Connan, M., Friedl, W., Vodermayer, B., Castellini, C., Hebert, J., & Pilarski, P. M. (2022). A transradial modular adaptable platform for evaluating prosthetic feedback and control strategies. *Proc. of MEC22 Symposium*, *22*, 136–139. https://conferences.lib.unb.ca/index.php/mec/article/view/1948.

Hancock, P., Billings, D., Schaefer, K., Chen, J., de Visser, E., & Parasuraman, R. (2011). A meta-analysis of factors affecting trust in human-robot interaction. *Human Factors*, *53*, 517–27. https://doi.org/10.1177/0018720811417254.

Hargrove, L. J., Miller, L. A., Turner, K., & Kuiken, T. A. (2017). Myoelectric pattern recognition outperforms direct control for transhumeral amputees with targeted muscle reinnervation: A randomized clinical trial. *Scientific Reports*, *7*(1), 13840. https://doi.org/10.1038/s41598-017-14386-w.

Hebert, J. S., Boser, Q. A., Valevicius, A. M., Tanikawa, H., Lavoie, E. B., Vette, A. H., Pilarski, P. M., & Chapman, C. S. (2019). Quantitative eye gaze and movement differences in visuomotor adaptations to varying task demands among upper-extremity prosthesis users. *JAMA Network Open*, *2*(9), 13. https://doi.org/10.1001/jamanetworkopen.2019.11197.

Hebert, J. S., Elzinga, K., Chan, K. M., Olson, J., & Morhart, M. (2014). Updates in targeted sensory reinnervation for upper limb amputation. *Current Surgery Reports*, *2*(3), 45. https://doi.org/10.1007/s40137-013-0045-7.

Hebert, J. S., Olson, J. L., Morhart, M. J., Dawson, M. R., Marasco, P. D., Kuiken, T. A., & Chan, K. M. (2013). Novel targeted sensory reinnervation technique to restore functional hand sensation after transhumeral amputation. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *22*(4), 765–773. https://doi.org/10.1109/TNSRE.2013.2294907.

Hebert, J. S., & Shehata, A. W. (2022). The effect of sensory feedback on the temporal allocation of gaze using a sensorized myoelectric prosthesis. *Proc. of MEC22 Symposium*, *22*, 37–40. https://www.unb.ca/ibme/_assets/documents/mec22conferenceprogram.pdf.

Hebert, J. S., Wolfe, D. L., Miller, W. C., Deathe, A. B., Devlin, M., & Pallaveshi, L. (2009). Outcome measures in amputation rehabilitation: Icf body functions. *Disability and Rehabilitation*, *31*(19), 1541–1554. https://doi.org/10.1080/09638280802639467.

Hedlund, E., Johnson, M., & Gombolay, M. (2021). The effects of a robot's performance on human teachers for learning from demonstration tasks. *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*, 207–215. https://doi.org/10.1145/3434073.3444664.

Herberts, P., Almström, C., Kadefors, R., & Lawrence, P. D. (1973). Hand prosthesis control via myoelectric patterns. *Acta Orthopaedica Scandinavica*, *44*(4), 389–409. https://doi.org/10.3109/17453677308989075.

Herr, H. (2009). Exoskeletons and orthoses: Classification, design challenges and future directions. *Journal of NeuroEngineering and Rehabilitation*, *6*(1), 21. https://doi.org/10.1186/1743-0003-6-21.

Heyes, C. M. (1998). Theory of mind in nonhuman primates. *Behavioral and Brain Sciences*, *21*(1), 101–114. https://doi.org/10.1017/S0140525X98000703.

Hochberg, L. R., Serruya, M. D., Friehs, G. M., Mukand, J. A., Saleh, M., Caplan, A. H., Branner, A., Chen, D., Penn, R. D., & Donoghue, J. P. (2006). Neuronal ensemble control of prosthetic devices by a human with tetraplegia. *Nature*, *442*(7099), 164–171. https://doi.org/10.1038/nature04970.

Holzinger, A. (2018). From machine learning to explainable AI. *2018 World Symposium on Digital Intelligence for Systems and Machines (DISA)*, 55–66. https://doi.org/10.1109/DISA.2018.8490530.

Jackson, M. (2002). Familiar and foreign bodies: A phenomenological exploration of the human-technology interface. *Journal of the Royal Anthro-*

pological Institute, *8*(2), 333–346. https://doi.org/10.1111/1467-9655.00006.
Jaques, N., Lazaridou, A., Hughes, E., Gulcehre, C., Ortega, P., Strouse, D., Leibo, J. Z., & De Freitas, N. (2019). Social influence as intrinsic motivation for multi-agent deep reinforcement learning. In K. Chaudhuri & R. Salakhutdinov (Eds.), *Proceedings of the 36th international conference on machine learning* (pp. 3040–3049). Proceedings of Machine Learning Research (PMLR). https://proceedings.mlr.press/v97/jaques19a.html.
Jette, A. M. (2006). Toward a common language for function, disability, and health. *Physical Therapy*, *86*(5), 726–734. https://doi.org/10.1093/ptj/86.5.726.
Judah, K., Roy, S., Fern, A., & Dietterich, T. (2010). Reinforcement learning via practice and critique advice. *24*(1), 481–486. https://doi.org/10.1609/aaai.v24i1.7690.
Kalinowska, A., Pilarski, P. M., & Murphey, T. D. (2023). Embodied communication: How robots and people communicate through physical interaction. *Annual Review of Control, Robotics, and Autonomous Systems*, *6*, 205–232. https://doi.org/10.1146/annurev-control-070122-102501.
Kaplan, F., Oudeyer, P.-Y., Kubinyi, E., & Miklósi, Á. (2002). Robotic clicker training. *Robotics and Autonomous Systems*, *38*(3), 197–206. https://doi.org/10.1016/S0921-8890(02)00168-9.
Kaushik, V., & Walsh, C. A. (2019). Pragmatism as a research paradigm and its implications for social work research. *Social Sciences*, *8*(9), 255. https://doi.org/10.3390/socsci8090255.
Kehoe, E. J., & Macrae, M. (2002). Fundamental behavioral methods and findings in classical conditioning. In J. W. Moore (Ed.), *A neuroscientist's guide to classical conditioning* (pp. 171–231). Springer New York. https://doi.org/10.1007/978-1-4419-8558-3_6.
Knoblich, G., Butterfil, S., & Sebanz, N. (2011). Psychological research on joint action: Theory and data. In B. Ross (Ed.), *Advances in research and theory* (pp. 59–101). Academic Press. https://www.sciencedirect.com/science/article/pii/B9780123855275000036.
Knox, W. B., & Stone, P. (2009). Interactively shaping agents via human reinforcement: The tamer framework. *Proceedings of the Fifth International Conference on Knowledge Capture*, 9–16. https://doi.org/10.1145/1597735.1597738.
Knox, W. B., & Stone, P. (2012). Reinforcement learning from simultaneous human and mdp reward. *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems - Volume 1*, 475–482. 10.5555/2343576.2343644.
Knox, W. B., & Stone, P. (2013). Learning non-myopically from human-generated reward. *Proceedings of the 2013 International Conference on Intelligent User Interfaces*, 191–202. https://doi.org/10.1145/2449396.2449422.

Knox, W. B., Stone, P., & Breazeal, C. (2013). Training a robot via human feedback: A case study. *International Conference on Social Robotics*, 460–470.

Kollar, T., Tellex, S., Roy, D., & Roy, N. (2010). Toward understanding natural language directions. *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 259–266. https://doi.org/10.1109/HRI.2010.5453186.

Langley, P. (1997). Machine learning for adaptive user interfaces. In G. Brewka, C. Habel, & B. Nebel (Eds.), *Advances in artificial intelligence* (pp. 53–62). Springer Berlin Heidelberg.

Lavoie, E. B., Valevicius, A. M., Boser, Q. A., Kovic, O., Vette, A. H., Pilarski, P. M., Hebert, J. S., & Chapman, C. S. (2018). Using synchronized eye and motion tracking to determine high-precision eye-movement patterns during object-interaction tasks. *Journal of Vision*, *18*(6), 18. https://doi.org/10.1167/18.6.18.

Lazaridou, A., & Baroni, M. (2020). Emergent multi-agent communication in the deep learning era. *arXiv:2006.02419 [cs.CL]*, 1–4. https://doi.org/10.48550/arXiv.2006.02419.

Lazaridou, A., Peysakhovich, A., & Baroni, M. (2016). Multi-agent cooperation and the emergence of (natural) language. *arXiv:1612.07182 [cs.CL]*, 1–11. https://doi.org/10.48550/arXiv.1612.07182.

Li, G., Gomez, R., Nakamura, K., & He, B. (2019). Human-centered reinforcement learning: A survey. *IEEE Transactions on Human-Machine Systems*, *49*(4), 337–349. https://doi.org/10.1109/THMS.2019.2912447.

Licklider, J. C. R. (1960). Man-computer symbiosis. *IRE Transactions on Human Factors in Electronics*, *HFE-1*(1), 4–11. https://doi.org/10.1109/THFE2.1960.4503259.

Light, C. M., Chappell, P. H., & Kyberd, P. J. (2002). Establishing a standardized clinical assessment tool of pathologic and prosthetic hand function: Normative data, reliability, and validity. *Archives of Physical Medicine and Rehabilitation*, *83*(6), 776–783. https://doi.org/10.1053/apmr.2002.32737.

Lin, L. J. (1991). Programming robots using reinforcement learning and teaching. *AAAI*, 781–786.

Lin, L. J. (1993). Hierarchical learning of robot skills by reinforcement. *IEEE International Conference on Neural Networks*, *1*, 181–186. https://doi.org/10.1109/ICNN.1993.298553.

Liu, C., & Hedrick, J. K. (2016). Cooperative search using human-UAV teams. In *AIAA infotech at aerospace* (p. 1653). American Institute of Aeronautics; Astronautics. https://doi.org/10.2514/6.2016-1653.

Loftin, R., MacGlashan, J., Peng, B., Taylor, M. E., Littman, M. L., Huang, J., & Roberts, D. L. (2014). A strategy-aware technique for learning behaviors from discrete human feedback. *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence*, *28*(1), 937–943. https://doi.org/10.1609/aaai.v28i1.8839.

Longo, L., Brcic, M., Cabitza, F., Choi, J., Confalonieri, R., Ser, J. D., Guidotti, R., Hayashi, Y., Herrera, F., Holzinger, A., Jiang, R., Khosravi, H., Lecue, F., Malgieri, G., Páez, A., Samek, W., Schneider, J., Speith, T., & Stumpf, S. (2024). Explainable artificial intelligence (XAI) 2.0: A manifesto of open challenges and interdisciplinary research directions. *Information Fusion*, 102301. https://doi.org/10.1016/j.inffus.2024.102301.

Maes, P. (1995). Agents that reduce work and information overload. In R. M. Baecker, J. Grudin, W. A. S. Buxton, & S. Greenberg (Eds.), *Readings in human–computer interaction* (pp. 811–821). Morgan Kaufmann. https://doi.org/https://doi.org/10.1016/B978-0-08-051574-8.50084-4.

Mann, R. W., & Reimers, S. D. (1970). Kinesthetic sensing for the EMG controlled "Boston Arm". *IEEE Transactions on Man-Machine Systems*, *11*(1), 110–115. https://doi.org/10.1109/TMMS.1970.299971.

Marasco, P. D., Hebert, J. S., Sensinger, J. W., Beckler, D. T., Thumser, Z. C., Shehata, A. W., Williams, H. E., & Wilson, K. R. (2021). Neurorobotic fusion of prosthetic touch, kinesthesia, and movement in bionic upper limbs promotes intrinsic brain behaviors. *Science Robotics*, *6*(58), eabf3368. https://doi.org/10.1126/scirobotics.abf3368.

Markoff, J. (2015). *Machines of loving grace: The quest for common ground between humans and robots*. ECCO, an imprint of HarperCollinsPublishers.

Marković, M., Dosen, S., Cipriani, C., Popović, D. B., & Farina, D. (2014). Stereovision and augmented reality for closed-loop control of grasping in hand prostheses. *Journal of Neural Engineering*, *11*(4), 046001. https://doi.org/10.1088/1741-2560/11/4/046001.

Mathewson, K. W., Parker, A. S. R., Sherstan, C., Edwards, A. L., Sutton, R. S., & Pilarski, P. M. (2023). Communicative capital: A key resource for human–machine shared agency and collaborative capacity. *Neural Computing and Applications*, *35*(23), 16805–16819. https://doi.org/10.1007/s00521-022-07948-1.

Mathewson, K. W., & Pilarski, P. M. (2016). Simultaneous control and human feedback in the training of a robotic agent with actor-critic reinforcement learning. *arXiv:1606.06979 [cs.HC]*, *abs/1606.06979*. https://doi.org/10.48550/arXiv.1606.06979.

Mathewson, K. W., & Pilarski, P. M. (2022). A brief guide to designing and evaluating human-centered interactive machine learning. *Machine Learning Evaluation Standards Workshop at ICLR 2022*, 1–7. https://doi.org/10.48550/arXiv.2204.09622.

Mathiowetz, V. G., Volland, G., Kashman, N., & Weber, K. (1985). Adult norms for the box and block test of manual dexterity. *The American Journal of Occupational Therapy: Official Publication of the American Occupational Therapy Association*, *39*(6), 386–391. https://doi.org/10.5014/ajot.39.6.386.

Micera, S., Carpaneto, J., & Raspopovic, S. (2010). Control of hand prostheses using peripheral information. *IEEE Reviews in Biomedical Engineering*, *3*, 48–68. https://doi.org/10.1109/RBME.2010.2085429.

Millán, J. d. R., Rupp, R., Mueller-Putz, G., Murray-Smith, R., Giugliemma, C., Tangermann, M., Vidaurre, C., Cincotti, F., Kubler, A., Leeb, R., Neuper, C., Mueller, K., & Mattia, D. (2010). Combining brain–computer interfaces and assistive technologies: State-of-the-art and challenges. *Frontiers in Neuroscience*, *4*, 1–15. https://doi.org/10.3389/fnins.2010.00161.

Misselhorn, C. (2015). Collective agency and cooperation in natural and artificial systems. In C. Misselhorn (Ed.), *Collective agency and cooperation in natural and artificial systems: Explanation, implementation and simulation* (pp. 3–24). Springer International Publishing. https://doi.org/10.1007/978-3-319-15515-9_1.

Modayil, J., & Sutton, R. S. (2014). Prediction driven behavior: Learning predictions that drive fixed responses. *The AAAI-14 Workshop on Artificial Intelligence and Robotics*, 36–42.

Modayil, J., White, A., & Sutton, R. S. (2014). Multi-timescale nexting in a reinforcement learning robot. *Adaptive Behavior*, *22*(2), 146–160. https://doi.org/10.1177/1059712313511648.

Mook, D. G. (1983). In defense of external invalidity. *American Psychologist*, *38*(4), 379–387. https://doi.org/10.1037/0003-066X.38.4.379.

Moss, F. (2011). *The sorcerers and their apprentices: How the digital magicians of the mit media lab are creating the innovative technologies that will transform our lives*. Crown Business.

Müller, J. S., Vidaurre, C., Schreuder, M., Meinecke, F. C., Von Bünau, P., & Müller, K.-R. (2017). A mathematical model for the two-learners problem. *Journal of Neural Engineering*, *14*(3), 036005.

Murray, C. D. (2004). An interpretative phenomenological analysis of the embodiment of artificial limbs. *Disability and Rehabilitation*, *26*(16), 963–973. https://doi.org/10.1080/09638280410001696764.

Ortiz-Catalan, M., Mastinu, E., Sassu, P., Aszmann, O., & Brånemark, R. (2020). Self-contained neuromusculoskeletal arm prostheses. *New England Journal of Medicine*, *382*(18), 1732–1738. https://doi.org/10.1056/NEJMoa1917537.

Ortiz-Catalán, M., Håkansson, B., & Brånemark, R. (2014). An osseointegrated human-machine gateway for long-term sensory feedback and motor control of artificial limbs. *Science Translational Medicine*, *6*(257), 257re6–257re6. https://doi.org/10.1126/scitranslmed.3008933.

Osiurak, F., & Badets, A. (2016). Tool use and affordance: Manipulation-based versus reasoning-based approaches. *Psychological Review*, *123*(5), 534–568. https://doi.org/10.1037/rev0000027.

Papadimitriou, C. (2008). Becoming en-wheeled: The situated accomplishment of re-embodiment as a wheelchair user after spinal cord injury. *Disability & Society*, *23*(7), 691–704. https://doi.org/10.1080/09687590802469420.

Parker, A. S. R., Dawson, M. R., & Pilarski, P. M. (2022). Continually learned Pavlovian signalling without forgetting for human-in-the-loop robotic control. *NeurIPS Workshop on Human in the Loop Learning (HiLL). also arXiv:2305.14365 [cs.LG]*, 1–12. https://doi.org/10.48550/arXiv.2305.14365.

Parker, A. S. R., Edwards, A. L., & Pilarski, P. M. (2014). Using learned predictions as feedback to improve control and communication with an artificial limb: Preliminary findings. *arXiv:1408.1913 [cs.AI]*. https://doi.org/10.48550/arXiv.1408.1913.

Parker, A. S. R., Edwards, A. L., & Pilarski, P. M. (2018). Exploring communication as actions in human-machine partnerships [Abstract and Poster]. *12th Annual Canadian Neuroscience meeting Satellite 1: CAPnet/CPS*.

Parker, A. S. R., Edwards, A. L., & Pilarski, P. M. (2019a). Exploring the impact of machine-learned predictions on feedback from an artificial limb. *2019 16th IEEE/RAS-EMBS International Conference on Rehabilitation Robotics (ICORR)*, 1239–1246. https://doi.org/10.1109/ICORR.2019.8779424.

Parker, A. S. R., Edwards, A. L., & Pilarski, P. M. (2019b). Machine-learned predictions assisting human control of an artificial limb [Abstract and Poster]. *4th Multidisciplinary Conference on Reinforcement Learning and Decision Making (RLDM)*.

Parker, A. S. R., & Pilarski, P. M. (2021). Position statement: Assistive technology as partners through machine-learned communication. *Workshop on Reinforcement Learning for Humans, Computer, and Interaction (RL4HCI), ACM CHI 2021*, 1–3. https://sites.google.com/view/rl4hci/position-papers.

Parker, A. S. R., Williams, H. E., Phelan, S. K., Hebert, J. S., Shehata, A. W., & Pilarski, P. M. (2024). *Understanding human interaction with real-time adaptive feedback during simulated prosthesis use* [In preparation].

Parker, P. A., Englehart, K. B., & Hudgins, B. (2006). Myoelectric signal processing for control of powered limb prostheses [Special Section (pp. 541–610): 2006 ISEK Congress]. *Journal of Electromyography and Kinesiology*, *16*(6), 541–548. https://doi.org/10.1016/j.jelekin.2006.08.006.

Parr, J. V. V., Vine, S. J., Harrison, N. R., & Wood, G. (2018). Examining the spatiotemporal disruption to gaze when using a myoelectric prosthetic hand. *Journal of Motor Behavior*, *50*(4), 416–425. https://doi.org/10.1080/00222895.2017.1363703.

Paton, J. J., & Buonomano, D. V. (2018). The neural basis of timing: Distributed mechanisms for diverse functions. *Neuron*, *98*(ue 4), 687–705. https://doi.org/10.1016/j.neuron.2018.03.045..

Peerdeman, B., Boere, D., Witteveen, H., Huis in‘t Veld, R., Hermens, H., Stramigioli, S., Rietman, H., Veltink, P., & Misra, S. (2011). Myoelectric forearm prostheses: State of the art from a user-centered perspective. *The Journal of Rehabilitation Research and Development*, *48*(6), 719–738. https://doi.org/10.1682/JRRD.2010.08.0161.

Pesquita, A., Whitwell, R. L., & Enns, J. T. (2018). Predictive joint-action model: A hierarchical predictive approach to human cooperation. *Psychonomic Bulletin & Review*, *25*(5), 1751–1769. https://doi.org/10.3758/s13423-017-1393-6.

Pezzulo, G., & Dindo, H. (2011). What should I do next? Using shared representations to solve interaction problems. *Experimental Brain Research*, *211*(3), 613–630. https://doi.org/10.1007/s00221-011-2712-1.

Pezzulo, G., Donnarumma, F., & Dindo, H. (2013). Human sensorimotor communication: A theory of signaling in online social interactions. *PLoS ONE*, *8*(11), 1–11. https://doi.org/10.1371/journal.pone.0079876.

Pfaffenberger, C. J. (1976). *Guide dogs for the blind: Their selection, development, and training*. Elsevier Scientific Publishing Company.

Pilarski, P. M., & Sutton, R. S. (2012). Between instruction and reward: Human-prompted switching. *AAAI 2012 Fall Symposium on Robots Learning Interactively from Human Teachers (RLIHT)*, 46–52. https://cdn.aaai.org/ocs/5496/5496-23886-1-PB.pdf.

Pilarski, P. M., Butcher, A., Davoodi, E., Johanson, M. B., Brenneis, D. J. A., Parker, A. S. R., Acker, L., Botvinick, M. M., Modayil, J., & White, A. (2022). The Frost Hollow experiments: Pavlovian signalling as a path to coordination and communication between agents. *arXiv:2203.09498 [cs.AI]*. https://doi.org/10.48550/arXiv.2203.09498.

Pilarski, P. M., Butcher, A., Johanson, M., Botvinick, M. M., Bolt, A., & Parker, A. S. R. (2019). Learned human-agent decision-making, communication and joint action in a virtual reality environment. *The 4th Multidisciplinary Conference on Reinforcement Learning and Decision Making; also arXiv:1905.02691 [cs.AI]*. https://doi.org/10.48550/arXiv.1905.02691.

Pilarski, P. M., Dawson, M. R., Degris, T., Carey, J. P., Chan, K. M., Hebert, J. S., & Sutton, R. S. (2013). Adaptive artificial limbs: A real-time approach to prediction and anticipation. *IEEE Robotics & Automation Magazine*, *20*(1), 53–64. https://doi.org/10.1109/MRA.2012.2229948.

Pilarski, P. M., Dawson, M. R., Degris, T., Carey, J. P., & Sutton, R. S. (2012). Dynamic switching and real-time machine learning for improved human control of assistive biomedical robots. *2012 4th IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob)*, 296–302. https://doi.org/10.1109/BioRob.2012.6290309.

Pilarski, P. M., Dawson, M. R., Degris, T., Fahimi, F., Carey, J. P., & Sutton, R. S. (2011). Online human training of a myoelectric prosthesis controller via actor-critic reinforcement learning. *2011 IEEE International Conference on Rehabilitation Robotics (ICORR)*, 1–7. https://doi.org/10.1109/ICORR.2011.5975338.

Pilarski, P. M., Dick, T. B., & Sutton, R. S. (2013). Real-time prediction learning for the simultaneous actuation of multiple prosthetic joints. *2013 IEEE International Conference on Rehabilitation Robotics (ICORR)*, 1–8. https://doi.org/10.1109/ICORR.2013.6650435.

Pilarski, P. M., & Hebert, J. S. (2017). Upper and lower limb robotic prostheses. In P. Encarnacao & A. M. Cook (Eds.), *Robotic assistive technologies: Principles and practice* (pp. 99–144). CRC Press.

Pilarski, P. M., & Sherstan, C. (2016). Steps toward knowledgeable neuroprostheses. *2016 6th IEEE International Conference on Biomedical Robotics and Biomechatronics (BioRob)*, 220–220. https://doi.org/10.1109/BIOROB.2016.7523626.

Pilarski, P. M., Sutton, R. S., & Mathewson, K. W. (2015). Prosthetic devices as goal-seeking agents. *Second Workshop on Present and Future of Non-Invasive Peripheral-Nervous-System Machine Interfaces: Progress in Restoring the Human Functions (PNS-MI)*, 4.

Ponterotto, J. G. (2005). Qualitative research in counseling psychology: A primer on research paradigms and philosophy of science. *Journal of Counseling Psychology*, *52*(2), 126–136. https://doi.org/10.1037/0022-0167.52.2.126.

Pulliam, C. L., Lambrecht, J. M., & Kirsch, R. F. (2011). Emg-based neural network control of transhumeral prostheses. *Journal of Rehabilitation Research and Development*, *48*(6), 739. https://doi.org/10.1682/jrrd.2010.12.0237.

Rabinowitz, N., Perbet, F., Song, F., Zhang, C., Eslami, S. M. A., & Botvinick, M. (2018). Machine theory of mind. *Proceedings of the 35th International Conference on Machine Learning*, *80*, 4218–4227. https://proceedings.mlr.press/v80/rabinowitz18a.html.

Radmand, A., Scheme, E., & Englehart, K. (2016). High-density force myography: A possible alternative for upper-limb prosthetic control. *Journal of Rehabilitation Research & Development (JRRD)*, *53*(4), 443–456. https://doi.org/10.1682/jrrd.2015.03.0041.

Rafiee, B., Abbas, Z., Ghiassian, S., Kumaraswamy, R., Sutton, R., Ludvig, E., & White, A. (2023). From eye-blinks to state construction: Diagnostic benchmarks for online representation learning. *Adaptive Behavior*, *31*(1), 3–19. https://doi.org/10.1177/10597123221085039.

Rashidi, P., & Mihailidis, A. (2013). A survey on ambient-assisted living tools for older adults. *IEEE Journal of Biomedical and Health Informatics*, *17*(3), 579–590. https://doi.org/10.1109/JBHI.2012.2234129.

Resnik, L. (2011). Development and testing of new upper-limb prosthetic devices: Research designs for usability testing. *The Journal of Rehabilitation Research and Development*, *48*(6), 697–706. https://doi.org/10.1682/JRRD.2010.03.0050.

Resnik, L., Meucci, M. R., Lieberman-Klinger, S., Fantini, C., Kelty, D. L., Disla, R., & Sasson, N. (2012). Advanced upper limb prosthetic devices: Implications for upper limb prosthetic rehabilitation. *Archives of Physical Medicine and Rehabilitation*, *93*(4), 710–717. https://doi.org/10.1016/j.apmr.2011.11.010.

Retzlaff, C. O., Das, S., Wayllace, C., Mousavi, P., Afshari, M., Yang, T., Saranti, A., Angerschmid, A., Taylor, M. E., & Holzinger, A. (2024).

Human-in-the-loop reinforcement learning: A survey and position on requirements, challenges, and opportunities. *Journal of Artificial Intelligence Research*, *79*, 359–415. https://doi.org/10.1613/jair.1.15348.

Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences*, *20*, 676–688. https://doi.org/10.1016/j.tics.2016.07.002.

Santoro, A., Lampinen, A. K., Mathewson, K. W., Lillicrap, T. P., & Raposo, D. (2021). Symbolic behaviour in artificial intelligence. *arXiv:2102.03406 [cs.AI]*, *abs/2102.03406*. https://doi.org/10.48550/arXiv.2102.03406.

Saridis, G. N., & Stephanou, H. E. (1977). A hierarchical approach to the control of a prosthetic arm. *IEEE Transactions on Systems, Man, and Cybernetics*, *7*(6), 407–420. https://doi.org/10.1109/TSMC.1977.4309737.

Sartori, M., Durandau, G., Došen, S., & Farina, D. (2018). Robust simultaneous myoelectric control of multiple degrees of freedom in wrist-hand prostheses by real-time neuromusculoskeletal modeling. *Journal of Neural Engineering*, *15*(6), 066026. https://doi.org/10.1088/1741-2552/aae26b.

Saunders, I., & Vijayakumar, S. (2011). The role of feed-forward and feedback processes for closed-loop prosthesis control. *Journal of Neuroengineering and Rehabilitation*, *8*(60), 1–12. https://doi.org/10.1186/1743-0003-8-60.

Scheme, E., & Englehart, K. (2011). Electromyogram pattern recognition for control of powered upper-limb prostheses: State of the art and challenges for clinical use. *Journal of Rehabilitation Research & Development*, *48*(6), 643–660. https://doi.org/10.1682/JRRD.2010.09.0177.

Scheme, E., Fougner, A., Stavdahl, Ø., Chan, A. D. C., & Englehart, K. (2010). Examining the adverse effects of limb position on pattern recognition based myoelectric control. *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, 6337–6340. https://doi.org/10.1109/IEMBS.2010.5627638.

Schofield, J. S., Battraw, M. A., Parker, A. S. R., Pilarski, P. M., Sensinger, J. W., & Marasco, P. D. (2021). Embodied cooperation to promote forgiving interactions with autonomous machines. *Frontiers in Neurorobotics*, *15*(661603). https://doi.org/10.3389/fnbot.2021.661603.

Schofield, J. S., Evans, K. R., Carey, J. P., & Hebert, J. S. (2014). Applications of sensory feedback in motorized upper extremity prosthesis: A review. *Expert Review of Medical Devices*, *11*(5), 499–511. https://doi.org/10.1586/17434440.2014.929496.

Schwandt, T. A. (1994). Constructivist, interpretivist approaches to human inquiry. In *Handbook of qualitative research* (pp. 118–137). Sage Publications, Inc.

Scott-Phillips, T. C. (2014). *Speaking our minds: Why human communication is different, and how language evolved to make it special*. Palgrave Macmillan.

Scott-Phillips, T. C., Kirby, S., & Ritchie, G. R. (2009). Signalling signalhood and the emergence of communication. *Cognition*, *113*(2), 226–233. https://doi.org/10.1016/j.cognition.2009.08.009.

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: Bodies and minds moving together. *Trends in Cognitive Sciences*, *10*(2), 70–76. https://doi.org/10.1016/j.tics.2005.12.009.

Sebanz, N., & Knoblich, G. (2009). Prediction in joint action: What, when, and where. *Topics in Cognitive Science*, *1*(2), 353–367. https://doi.org/10.1111/j.1756-8765.2009.01024.x.

Sensinger, J. W., & Dosen, S. (2020). A review of sensory feedback in upper-limb prostheses from the perspective of human motor control. *Frontiers in Neuroscience*, *14*, 345. https://doi.org/10.3389/fnins.2020.00345.

Serpell, J., & Hsu, Y. (2001). Development and validation of a novel method for evaluating behavior and temperament in guide dogs. *Applied Animal Behaviour Science*, *72*(4), 347–364. https://doi.org/https://doi.org/10.1016/S0168-1591(00)00210-0.

Shehata, A. W., Rehani, M., Jassat, Z. E., & Hebert, J. S. (2020). Mechanotactile sensory feedback improves embodiment of a prosthetic hand during active use. *Frontiers in Neuroscience*, *14*, 513550. https://doi.org/10.3389/fnins.2020.00263.

Shehata, A. W., Scheme, E. J., & Sensinger, J. W. (2018a). Audible feedback improves internal model strength and performance of myoelectric prosthesis control. *Scientific Reports*, *8*(8541), 1–10. https://doi.org/10.1038/s41598-018-26810-w.

Shehata, A. W., Scheme, E. J., & Sensinger, J. W. (2018b). Evaluating internal model strength and performance of myoelectric prosthesis control strategies. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *26*(5), 1046–1055. https://doi.org/10.1109/TNSRE.2018.2826981.

Sherstan, C. (2015). *Prosthetic arms as wearable intelligent robots* (Master's thesis). University of Alberta.

Sherstan, C., Modayil, J., & Pilarski, P. M. (2015). A collaborative approach to the simultaneous multi-joint control of a prosthetic arm. *2015 IEEE/RAS-EMBS International Conference on Rehabilitation Robotics (ICORR)*, 13–18. https://doi.org/10.1109/ICORR.2015.7281168.

Smail, L. C., Neal, C., Wilkins, C., & Packham, T. L. (2021). Comfort and function remain key factors in upper limb prosthetic abandonment: Findings of a scoping review. *Disability and Rehabilitation: Assistive Technology*, *16*(8), 821–830. https://doi.org/10.1080/17483107.2020.1738567.

Smith, L. H., Kuiken, T. A., & Hargrove, L. J. (2016). Myoelectric control system and task-specific characteristics affect voluntary use of simultaneous control. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *24*(1), 109–116. https://doi.org/10.1109/TNSRE.2015.2410755.

Sutton, R. S. (1988). Learning to predict by the methods of temporal differences. *Machine Learning*, *3*(1), 9–44. https://doi.org/10.1007/BF00115009.

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction, second edition*. MIT Press.

Sutton, R. S., Modayil, J., Delp, M., Degris, T., Pilarski, P. M., White, A., & Precup, D. (2011). Horde: A scalable real-time architecture for learning knowledge from unsupervised sensorimotor interaction. *Proceedings of 10th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 761–768.

Svensson, P., Wijk, U., Björkman, A., & Antfolk, C. (2017). A review of invasive and non-invasive sensory feedback in upper limb prostheses. *Expert Review of Medical Devices*, *14*(6), 439–447. https://doi.org/10.1080/17434440.2017.1332989.

Taylor, M. E., Nissen, N., Wang, Y., & Navidi, N. (2023). Improving reinforcement learning with human assistance: An argument for human subject studies with hippo gym. *Neural Computing and Applications*, *35*(32), 23429–23439. https://doi.org/10.1007/s00521-021-06375-y.

Thomas, N., Fazlollahi, F., Brown, J. D., & Kuchenbecker, K. J. (2021). Sensorimotor-inspired tactile feedback and control improve consistency of prosthesis manipulation in the absence of direct vision. *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 6174–6181. https://doi.org/10.1109/IROS51168.2021.9635885.

Thomaz, A. L., & Breazeal, C. (2008). Teachable robots: Understanding human teaching behavior to build more effective robot learners. *Artificial Intelligence*, *172*(6), 716–737. https://doi.org/10.1016/j.artint.2007.09.009.

Tosic, P. T., & Agha, G. A. (2004). Towards a hierarchical taxonomy of autonomous agents. *2004 IEEE International Conference on Systems, Man and Cybernetics (IEEE Cat. No.04CH37583)*, *4*, 3421–3426 vol.4. https://doi.org/10.1109/ICSMC.2004.1400871.

Travnik, J. B., & Pilarski, P. M. (2017). Representing high-dimensional data to intelligent prostheses and other wearable assistive robots: A first comparison of tile coding and selective Kanerva coding. *2017 IEEE-RAS-EMBS International Conference on Rehabilitation Robotics (ICORR)*, 1443–1450. https://doi.org/10.1109/ICORR.2017.8009451.

Valevicius, A. M., Boser, Q. A., Lavoie, E. B., Murgatroyd, G. S., Pilarski, P. M., Chapman, C. S., Vette, A. H., & Hebert, J. S. (2018). Characterization of normative hand movements during two functional upper limb tasks. *PLOS ONE*, *13*(6), 1–21. https://doi.org/10.1371/journal.pone.0199549.

Veeriah, V., Pilarski, P. M., & Sutton, R. S. (2016). Face valuing: Training user interfaces with facial expressions and reinforcement learning. *2016 IJCAI Workshop on Interactive Machine Learning; also arXiv:1606.02807 [cs.HC]*, *abs/1606.02807*. https://doi.org/10.48550/arXiv.1606.02807.

Vesper, C., Butterfill, S., Knoblich, G., & Sebanz, N. (2010). A minimal architecture for joint action. *Neural Networks*, *23*(8), 998–1003. https://doi.org/https://doi.org/10.1016/j.neunet.2010.06.002.

Viswanathan, P., Bell, J., Wang, R. H., Adhikari, B., Mackworth, A. K., Mihailidis, A., Miller, W. C., & Mitchell, I. M. (2014). A wizard-of-oz intelligent wheelchair study with cognitively-impaired older adults: Attitudes toward user control. *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems, Workshop on Assistive Robotics for Individuals with Disabilities: HRI Issues and Beyond, September 14, 2014, Chicago, Illinois, USA*.

Vu, P. P., Vaskov, A. K., Irwin, Z. T., Henning, P. T., Lueders, D. R., Laidlaw, A. T., Davis, A. J., Nu, C. S., Gates, D. H., Gillespie, R. B., Kemp, S. W. P., Kung, T. A., Chestek, C. A., & Cederna, P. S. (2020). A regenerative peripheral nerve interface allows real-time control of an artificial hand in upper limb amputees. *Science Translational Medicine*, *12*(533), eaay2857. https://doi.org/10.1126/scitranslmed.aay2857.

Wells, E. D., Carpenter, S., Dawson, M. R., Shehata, A. W., Carey, J. P., & Hebert, J. S. (2020). Development of a modular simulated prosthesis and evaluation of a compliant grip force sensor. *Proc. of MEC20 Symposium*, *20*, 179–182. https://conferences.lib.unb.ca/index.php/mec/issue/view/9/9.

Wells, E. D., Shehata, A. W., Dawson, M. R., Carey, J. P., & Hebert, J. S. (2022). Preliminary evaluation of the effect of mechanotactile feedback location on myoelectric prosthesis performance using a sensorized prosthetic hand. *Sensors*, *22*(10), 19. https://doi.org/10.3390/s22103892.

White, A. (2015). *Developing a predictive approach to knowledge* (Doctoral dissertation). University of Alberta Libraries. https://doi.org/10.7939/R3FF3M75H.

Wilding, C., & Whiteford, G. (2005). Phenomenological research: An exploration of conceptual, theoretical, and practical issues. *OTJR: Occupation, Participation and Health*, *25*(3), 98–104. https://doi.org/10.1177/153944920502500303.

Williams, H. E., Chapman, C. S., Pilarski, P. M., Vette, A. H., & Hebert, J. S. (2019). Gaze and movement assessment (GaMA): Inter-site validation of a visuomotor upper limb functional protocol. *PLOS ONE*, *14*(12), e0219333. https://doi.org/10.1371/journal.pone.0219333.

Williams, H. E., Chapman, C. S., Pilarski, P. M., Vette, A. H., & Hebert, J. S. (2021). Myoelectric prosthesis users and non-disabled individuals wearing a simulated prosthesis exhibit similar compensatory movement strategies. *Journal of NeuroEngineering and Rehabilitation*, *18*(1), 72. https://doi.org/10.1186/s12984-021-00855-x.

Williams, H. E., Günther, J., hebert, J. S., Pilarski, P. M., & Shehata, A. W. (2022). Composite recurrent convolutional neural networks offer a position-aware prosthesis control alternative while balancing predictive accuracy with training burden. *2022 International Conference on Rehabilitation*

Robotics (ICORR), 1–6. https://doi.org/10.1109/ICORR55369.2022.9896593.

Williams, H. E., Shehata, A. W., Cheng, K. Y., Hebert, J. S., & Pilarski, P. M. (2024). A multifaceted suite of metrics for comparative myoelectric prosthesis controller research. *PLoS ONE*, *19*(5), 1–32. https://doi.org/10.1371/journal.pone.0291279.

Williams, H. E., Shehata, A. W., Cheng, K. Y., hebert, J. S., & Pilarski, P. M. (2024). Myoelectric prosthesis control using recurrent convolutional neural network regression mitigates the limb position effect. *bioRxiv*, 2024–02. https://www.biorxiv.org/content/early/2024/02/08/2024.02.05.578477.

Williams III, T. W. (2011). Progress on stabilizing and controlling powered upper-limb prostheses. *The Journal of Rehabilitation Research and Development*, *48*(6), ix. https://doi.org/10.1680/JRRD.2010.05.0080.

Winance, M. (2006). Trying out the wheelchair: The mutual shaping of people and devices through adjustment. *Science, Technology, & Human Values*, *31*(1), 52–72. https://doi.org/10.1177/0162243905280023.

Wolpert, D. M., Ghahramani, Z., & Flanagan, J. R. (2001). Perspectives and problems in motor learning. *Trends in Cognitive Sciences*, *5*(11), 487–494. https://doi.org/10.1016/S1364-6613(00)01773-3.

World Health Organization. (2007). *International classification of functioning, disability and health (ICF)*.

Xu, W., Huang, J., Wang, Y., & Cai, H. (2013). Study of reinforcement learning based shared control of walking-aid robot. *Proceedings of the 2013 IEEE/SICE International Symposium on System Integration*, 282–287. https://doi.org/10.1109/SII.2013.6776656.

Yamamoto, M., Chung, K. C., Sterbenz, J., Shauver, M. J., Tanaka, H., Nakamura, T., Oba, J., Chin, T., & Hirata, H. (2019). Cross-sectional international multicenter study on quality of life and reasons for abandonment of upper limb prostheses. *Plastic and Reconstructive Surgery - Global Open*, *7*(5), e2205. https://doi.org/10.1097/GOX.0000000000002205.

Zacks, J. M., Kurby, C. A., Eisenberg, M. L., & Haroutunian, N. (2011). Prediction error associated with the perceptual segmentation of naturalistic events. *Journal of Cognitive Neuroscience*, *23*(12), 4057–4066. https://doi.org/10.1162/jocn_a_00078.

Zhu, H., Neubig, G., & Bisk, Y. (2021). Few-shot language coordination by modeling theory of mind. *Proceedings of the 38th International Conference on Machine Learning*, *139*, 12901–12911. https://proceedings.mlr.press/v139/zhu21d.html.

Ziegler-Graham, K., MacKenzie, E. J., Ephraim, P. L., Travison, T. G., & Brookmeyer, R. (2008). Estimating the prevalence of limb loss in the united states: 2005 to 2050. *Archives of Physical Medicine and Rehabilitation*, *89*(3), 422–429. https://doi.org/https://doi.org/10.1016/j.apmr.2007.11.005.

Zuo, K. J., & Olson, J. L. (2014). The evolution of functional hand replacement: From iron prostheses to hand transplantation. *Plastic Surgery*, *22*(1), 44–51. https://doi.org/10.1177/229255031402200111.

Appendix A

Background Material

A.1 Additional Material for Chapter 7

A.1.1 Interview Questions and Script

Participant Data

`Handedness`

`Vision`

`   corrected`

`   uncorrected`

`   normally corrected`

`EMG experience`

Rapport Builders

`What made you interested in participating?`

`What's your favourite robot?`

`What's your favourite (media)?`

`What tasks energize you?`

`What tasks drain you?`

`(Share about self based on participant responses)`

Training Script

So, first is "hand on home, eyes on neutral". From here, you'll hear a sound from Heather at the console. That is your queue to reach for the cup, and move it from one side to the other, then go back to home and neutral, and wait for the next beep.

There are 2 other sounds. They won't occur here, but this is the easiest point to demonstrate them. They are the break sound, and the warning sound. While your goal is to try not to hear the break sound, there are no failure conditions. If you hear a break sound, just continue the trial. If you drop the cup, I will place it on the side you were moving to and then you go back to home/neutral as if you completed the motion, okay?

Trial 1 Script

In this first trial you'll be doing the same thing you just did in training, except with this cup with the motion capture markers. Also, that break sound I mentioned earlier will play if you squeeze the cup too hard. Only the break sound, no warnings in this trial. Also, every now and then Heather will call out to "Crush the Cup". When this happens, do just that; crush it.

Part 1

Intro: Thank you for agreeing to chat with me today: it is nice to meet you. We are here to talk to you about the trial you just did, with a focus on your interaction with the robot while trying to complete the task rather than on the task completion itself. This information you share with me today will help with a research project relating to human-AI interactions. Your individual perspective is invaluable to the goal of improving interactions between humans and devices. You can take a break whenever you want.

If you want to stop the interview at any point, you can tell me you'd like to stop. If you want to skip a question, you can.

Do you have any questions for me before we get started?

What was it like to participate in this research today?
- What were you thinking about?
- What were you excited about?
- What were you nervous about?

What did you think about your interaction with the power hand?
- Why do you think you noticed <participant named theme>?
- Why was <participant named theme> good/bad/noteworthy/matter?
- Tell me more about...

Did anything else stand out to you about your experience; what did you notice about the trial, sounds, or power hand?
- Why do you think you noticed that?
- How much did you notice it?
- How did it interfere?
- How was there something different at the beginning/end?

What were you thinking about or focused on as you grabbed the cup?
- Why do you think you noticed <participant named theme>?
- What else did you notice?

What were you thinking about or focused on when you were asked to crush
the cup?
- How come?

What about when you accidentally crushed the cup?

- Why do you think you noticed <participant named theme>?

How did your experience of the interaction change after you crushed the cup?
- How did that change over time?
- Why do you think you noticed <participant named theme>?

How did you respond to the crush sound?
- What did you do?

What was your goal or objective, in your own words?

What was your strategy for grabbing the cup?
- How did you arrive at that strategy?

Trial 2 Script

Alright. This time there will be a warning sound as well as the break sound. Do you remember what the warning sound is?

That sound doesn't mean a break has happened, but it does mean that a break sound is impending. This warning sound could be triggered directly by the sensor reading, or it could be a machine-learned prediction of the sensor reading. I'll tell you which one after, not right now. Does that all make sense?

Part 2

Do you have any questions for me before we get started?
What thoughts or feelings do you have about this trial vs the last one, or the last one vs this one?
- How did they compare?

What did you think about your interaction with the power hand?

- Why do you think you noticed <participant named theme>?
- Why was <participant named theme> good/bad/noteworthy/matter?
- Tell me more about...

Did anything else stand out to you about your experience; what did you notice about the trial, sounds, or power hand?
- Why do you think you noticed that?
- How much did you notice it?
- How did it interfere?
- Was there something different at the beginning/end?

What were you thinking about or focused on as you grabbed the cup?
- Why do you think you noticed <participant named theme>?
- What else did you notice?

What were you thinking about or focused on when you were asked to crush the cup?
- How come?
- What were you observing?

What about when you accidentally crushed the cup?
- Why do you think you noticed <participant named theme>?

How did your experience of the interaction change after you crushed the cup?
- How did that change over time?
- Why do you think you noticed <participant named theme>?

What was your goal or objective, in your own words?

What was your strategy for grabbing the cup?
- How did you arrive at that strategy?

How did you respond to the warning sound?

- What about the crush beep?

- How did the warning sound change your strategy?

Did the sounds make sense to you? How/How not? Over time?

- Was one more consistent than the other?

What did the warning mean to you?

What type of feedback do you think you just used [outline]?

- Why do you think that?

Does that change anything we've just talked about?

If you had to use a device like this regularly, what would you want it to do?

Is there anything else you would like to talk about that we didn't get a chance to discuss?

After Final Interview

Thank you so much for your time, we really appreciate your participation in this study.

A.1.2 Study Checklist

Upon Receiving Participant Interest

- ☐ Reply to the participant to schedule their timeslot
  - ☐ Ask that if the participant could wear contacts if they

typically wear glasses

☐ If participant can't wear contacts, ask if they can see 1 m in front of them w/o glasses

☐ Ask that the participant wear a short-sleeved shirt

☐ Ask participant about metal allergies or other skin sensitivities

☐ Inform the participant the the lab requests the use of masks while in the lab

Pre-Participant-Arrival

☐ Prepare consent form and pen

☐ Designate participant code and trial type

☐ Prepare error tracking sheets: fill info at top of the page and the trigger trials

☐ Set up computers

☐ Turn on motion capture computer

☐ Launch Lab Remote

☐ Connect blue ethernet cable to 1st port on the router and connect white ethernet cable to motion capture computer

☐ Connect blue ethernet cable to robot laptop

☐ Launch BrachIOplexus on robot computer

☐ Run Adam's Modified Adaptive Switching

☐ Refresh streams on Lab Remote to see if PowerHand shows up

☐ Pre-adjust belt

☐ Set up Lab Remote:

☐ Set file path

☐ Set first gaze cal file name and ensure that the trail number is set to 1: IDA_GazeCal_Stat (see Participant Code section at the end of the document)

☐ Motive Setup

☐ Launch Motive

☐ Enable NatNet in the Streaming Settings

☐ Set the file path in the Capture layout

☐ Remove everything from the motion capture space and close the curtain

☐ Calibrate the motion capture system in the Calibration layout

☐ Move cart and chair back into the motion capture space

☐ Remove marker 2 (affixed with velcro) from cart and set it aside

☐ Delete all rigid bodies in the Capture layout

☐ Load in rigid bodies from desktop (RHND, Cart, Cup, and Wand)

☐ Unselect the Cup rigid body

☐ Pupil Labs Setup

☐ Launch Pupil Labs

☐ Restart with default settings in the General Settings

☐ Detect eye 0 and eye 1 in the General Settings

☐ Turn on Pupil LSL Relay in the Plugin Manager

☐ Set the Calibration Mode to Single Marker Calibration

☐ In the eye 0 window, click Flip image display in General Settings

☐ Synchronization Setup on Motion Capture Computer

☐ Launch OptiTrack Steamer

☐ Refresh streams in Lab Remote

☐ Check that the Pupil Primitive Data and OptiTrackFrameID streams are selected to record

☐ Select OptiTrackFrameID to control

☐ Turn on Auto Play Beep, with a delay of 0

☐ Mute computer speaker

☐ Myo Armband Setup

☐ Briefly plug in the Myo armband to turn it on

☐ Launch Myo Connect on the laptop

☐ Connect Myo armband

☐ Turn on "Show Myo Gestures"

☐ Marker and Task Setup

☐ Ensure that all motion capture markers are placed on the eye tracker and the RHND rigid marker plate is placed on the power hand

☐ Ensure that task cups are prepared and accessible

☐ Put up the quiet sign on lab door

☐ Clean the eye tracker, Myo armband, measuring tape, and simulated prosthesis hand brace with an alcohol swab

☐ Ensure that the button on the simulated prosthesis belt is off

☐ Put out the the simulated prosthesis for the participant don and place a piece of alpha liner in the connecting ring

☐ Ensure that the following are available:

- ☐ Connecting ring bolts
- ☐ Hex key
- ☐ Socket cushions
- ☐ Velcro strap
- ☐ Sleeve material
- ☐ Scissors
- ☐ Transpore tape|3 pieces
- ☐ Alcohol swabs
- ☐ Recorder for interviews
- ☐ Water, juice, and snacks

Participant Arrival

☐ Welcome participant

☐ Ask participant if they'd like to use the washroom first

- ☐ Ask participant about metal allergies or other skin sensitivities

☐ Do letter of consent

☐ Collect participant data (EMG experience, Vision, Handedness)

☐ Explain EMG and muscle contractions

☐ Donning

☐ Measure forearm length - medial epicondyle to ulnar styloid process

☐ Clean the participant's arm with an alcohol swab

☐ Place the Myo armband on the participant

☐ Sync the Myo armband

☐ Turn off Show Myo Gestures

☐ Measure placement of Myo armband - medial epicondyle to top of armband

☐ Have the participant don the belt

☐ Prepare a fabric sleeve and have the participant don it

☐ Have the participant place their hand in the simulated prosthesis brace

☐ Tape the alphaliner

☐ Close the connecting ring with bolts, and tape the heads of the bolts

☐ Place cushions throughout the simulated prosthesis and tighten the straps

☐ Connect the simulated prosthesis to the belt via a cable

☐ Use one small piece velcro to secure the cable

☐ Secure velcro strap over fingers

☐ Confirm that the simulated prosthesis is secure by having the participant hold their arm at their side and shake their arm.

☐ Explain the button to the participant

☐ Turn on the simulated prosthesis via the button.

☐ Connect to the Bento Arm in brachI/Oplexus

☐ Connect to the Myo armband in brachI/Oplexus

☐ Configure EMG control for participant in brachI/Oplexus

☐ Adjust cart height

☐ Check/calibrate EMG signals in the 3 major task positions

☐ Eye Tracking Setup

☐ Affix marker 2 to the cart

☐ Have the participant stand at the task cart

☐ Place the eye tracker on the participant and plug the eye tracker in to the computer

☐ Adjust each eye camera (check pupil visibility in the 4 task areas of interest)

☐ Adjust the world view camera so the entire cart can be seen

☐ Tell the participant not to touch the eye tracker or their face

☐ Set world camera settings to (640,480) and 120 Hz

☐ Set eye camera settings to (192,192) and 120 Hz

☐ Adjust pupil min, max, and intensity range for each eye camera (check pupil detection in the 4 task areas of interest)

☐ Set the eye camera mode to Camera Image

☐ Close curtain

☐ Collect the Pupil Labs calibration|press c on the keyboard to start and end recording in Pupil Labs

☐ Create the HD rigid body|1: front, 2: participant's left, 3: participant's right, 4: lower marker

☐ Training

  ☐ Adjust computer speaker volume

  ☐ Demonstrate Lab Remote Sound

  ☐ Set feedback to direct heavy

☐ Demonstrate break sound

  ☐ Demonstrate warning sound

  ☐ Enable powerhand in brachI/Oplexus

☐ Set feedback to none

(Training procedure)

☐ Gaze Calibrations

☐ Suspend powerhand in brachI/Oplexus

☐ Refresh streams in Lab Remote

☐ Ensure PowerHand is NOT in LSL

☐ Close curtain

☐ Record stationary gaze calibration

☐ Record paint gaze calibration

☐ Record check gaze validation

☐ Block 1

☐ In Motive, ensure that 4 rigid bodies are selected (HD, Cart, Cup, and RHND)

☐ Set the Lab Remote file name to IDCF_Crush (where ID is the participant ID as xxx-y; C denotes the cup stiffness of L-light, M-medium, or S-strong;
F denotes the feedback type of N-none, D-direct, or P-prediction)

☐ Enable powerhand in brachI/Oplexus

☐ Launch Adam's modified adaptive switching code

☐ Refresh streams in Lab Remote

☐ Confirm PowerHand is in LSL

☐ Place first cup

☐ Ensure that participant understands the following:

☐ How to place simulated prosthesis hand at home

☐ Stay looking at neutral in between trials

☐ They can ask for breaks at any point

☐ Don't push the cup into the target - just place it and do your best to get it in the target

☐ To continue trials if they crush the cup (regardless of if they're told to crush it)

☐ Close the curtain

☐ Set correct threshold (BrachIO ‘Motion Sequencer’ tab)

(Trial procedure)

☐ 3 minute break

☐ Place medium cup

☐ Set correct threshold (BrachIO ‘Motion Sequencer’ tab)

(Trial procedure)

☐ 3 minute break

☐ Place final cup

(Trial procedure)

☐ Gaze Calibrations

☐ Suspend powerhand in brachI/Oplexus

☐ Turn off Adam’s modified adaptive switching code

☐ In Motive, ensure that 3 rigid bodies are selected (HD, Cart, and Wand)

☐ Set the Lab Remote file name to IDA_GazeCal_Stat or _Paint and set trial number

☐ Refresh streams in Lab Remote

☐ Ensure PowerHand is NOT in LSL

☐ Close curtain

☐ Record stationary gaze calibration or paint gaze calibration

☐ Record check gaze validation

☐ Interview

☐ Offer water/juice/snack

☐ Label any error/trigger trials from Block 1

☐ Training

☐ Demonstrate Lab Remote sound

☐ Demonstrate break sound

☐ Demonstrate warning sound

☐ Enable powerhand in brachI/Oplexus

(Training procedure)

☐ Gaze Calibrations

☐ Suspend powerhand in brachI/Oplexus

☐ In Motive, ensure that 3 rigid bodies are selected (HD, Cart, and Wand)

☐ Set the Lab Remote file name to IDA_GazeCal_Stat and set trial number

☐ Refresh streams in Lab Remote

☐ Ensure PowerHand is NOT in LSL

☐ Close curtain

☐ Record stationary gaze calibration

☐ Record paint gaze calibration

☐ Record check gaze validation

☐ Block 2 | With direct or prediction feedback

☐ In Motive, ensure that 4 rigid bodies are selected (HD, Cart, Cup, and RHND)

☐ Set the Lab Remote file name to IDCF_Crush

☐ Enable powerhand in brachI/Oplexus

☐ Launch Adam's modified adaptive switching code

☐ Select appropriate feedback type

- ☐ Pause learning if learning trial

☐ Refresh streams in Lab Remote

☐ Ensure PowerHand is in LSL

☐ Place first cup

☐ Ensure that participant understands the following:

- ☐ How to place simulated prosthesis hand at home
- ☐ Stay looking at neutral in between trials
- ☐ They can ask for breaks at any point

☐ Close the curtain

☐ Set correct threshold (BrachIO 'Motion Sequencer' tab)

☐ Unpause learning if learning trial

(Trial procedure)

☐ Pause learning if learning trial

☐ 3 minute break (pause learning in 'Motion Sequencer' tab of BrachIO)

☐ Place medium cup

☐ Set correct threshold (BrachIO 'Motion Sequencer' tab)

☐ Unpause learning if learning trial

(Trial procedure)

☐ Pause learning if learning trial

☐ 3 minute break (pause learning in 'Motion Sequencer' tab of BrachIO)

☐ Place final cup

☐ Unpause learning if learning trial

(Trial procedure)

☐ Gaze Calibrations

☐ Suspend powerhand in brachI/Oplexus

☐ Turn off Adam's modified adaptive switching code

☐ In Motive, ensure that 3 rigid bodies are selected (HD, Cart, and Wand)

☐ Set the Lab Remote file name to IDA_GazeCal_Stat or _Paint and set trial number

☐ Refresh streams in Lab Remote

☐ Ensure PowerHand is NOT in LSL

☐ Close curtain

☐ Record stationary gaze calibration or paint gaze calibration

☐ Record check gaze validation

☐ Interview

☐ Offer water

☐ Doffing:

  ☐ Unplug eye tracker, ensuring that you ground yourself first

  ☐ Turn off the button

  ☐ Unplug all cables

  ☐ Remove the bolts from the connecting ring

  ☐ Undo all velcro straps

  ☐ Open the simulated prosthesis and have the participant remove their arm

  ☐ Remove the sleeve and the Myo armband

  ☐ Have the participant remove the belt

  ☐ Have the participant remove the eye tracker

☐ Thank participant

Cleanup

☐ Clean eye tracker, myo armband, and simulated prosthesis brace with an alcohol swab

☐ Using the Bulk Rename Utility, label erroneous trials (excluding crushing the cup errors) with the suffix _error

☐ Using the Bulk Rename Utility, label trials trials where the participants were triggered to crush the cup with the suffix _trigger

☐ Save motion capture trials to a new folder called Raw

☐ Duplicate motion capture trials to a new folder called Cleaning

☐ Remove the motion capture system calibration files and the _error trials from the Cleaning folder

☐ Save xdf trials to a new folder called Exported

☐ Clean and export motion capture trials to the Exported folder, following these general steps:

  ☐ Open the files in the Cleaning folder in the Motive Edit view

  ☐ Check each trial to ensure that there are not large gaps in any labelled markers (quickly visible via the Tracks view)

☐ Correct any mislabelling and fill large gaps in any necessary trials

☐ If any trial has large gaps with all markers missing, try right
clicking on the trial and selecting \Reconstruct and Auto-Label".
But ensure that the auto-labelling is correct.

☐ Select all trials, right click, and click Solve All Assets

☐ Select all trials, right click, and click Export Tracking Data. Ensure that units are set millimeters, and save to the Exported folder.
This may take around 5 minutes, with the GazeCal files taking longer to
export.

☐ Select all trials, right click, and click Save. This may take around
7 minutes, and Motive may stop responding as this continues.
Monitor the Cleaning folder to check that the files are being saved recently, and you may need to close Motive if it freezes on any trial.

☐ Create two folders in the Exported folder - L2R and R2L. Move the odd Crush trials to L2R and move the even Crush trials to R2L

☐ Using the Bulk Raname Utility, label Crush trials with the suffix _L2R or _R2L

☐ Organize files in the Exported folder and subfolders into GaMA Project Creation folders: Gaze Calibration, Gaze Validation, Joint Calibration Landmarks, Joint Calibration Pose, Trials

☐ Launch GaMA, and create a New Project for L2R then R2L. This may take around 8 minutes for each project.

☐ Set task to Custom

☐ Set PavSig_PowerHand channels 8-10 interpolation to Previous

☐ Set output directory to the Projects folder

☐ Move the participant's folders from Project Creation to the respective project folders in ALL Project Creation

☐ Copy each folder in Projects and pasta in the Combined folder

☐ Shut down the mocap computer

☐ Fill Participant Metrics spreadsheet (handedness, vision, EMG experience, forearm length, and Myo armband location)

Procedures

Training

Feedback setting: none

Participants will move an object from left to right, then place the robot hand on home

After Lab Remote beep, participants will move the object from right to left and place the robot hand on home

Repeat for approx. 2 minutes

Trial

With correct feedback setting

1-3 for break-only feedback

4-6 for direct feedback

7-9 for predictive feedback

Participants will move an object from left to right, then place the robot hand on home

After Lab Remote beep, participants will move the object from right to left and place the robot hand on home

Repeat for 30 motions

For 1 motion in every 5 participant will be verbally asked to squeeze the cup harder

3 minute break for participant

Change Durometer

Participant Code

x1x - Direct feedback

    heavy > medium > light

x2x - Prediction feedback

    heavy > medium > light

A.1.3 Additional Figures

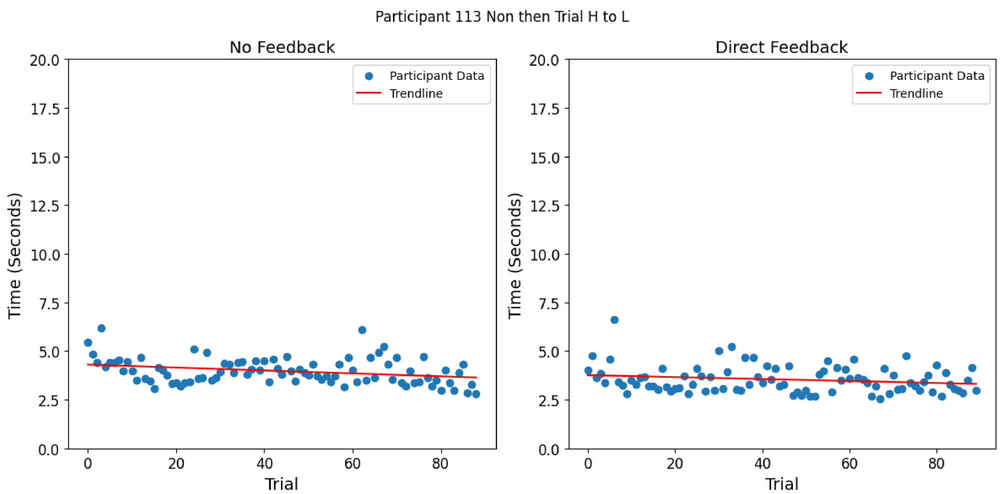


Figure A.1: The total trial time for every trial participant 113 did. This participant used direct feedback. The minor slope in the trend line in blue suggests that while the participant would have benefited from more training to further stabilize learning, there was some stability. Comments made by participants about control, therefore, may still occur in participants with more practice.

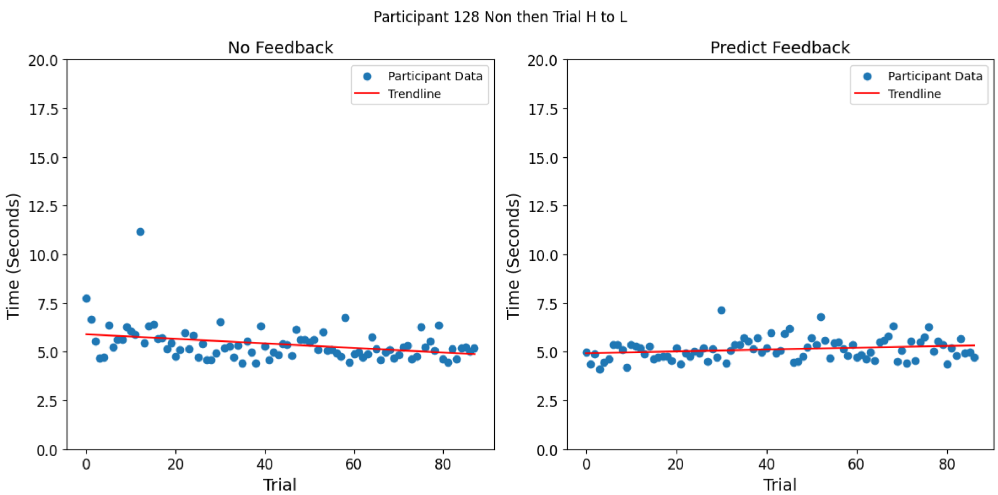


Figure A.2: The total trial time for every trial participant 128 did. This participant used predictive feedback. The minor slope in the trend line in blue suggests that while the participant would have benefited from more training to further stabilize learning, there was some stability. Comments made by participants about control, therefore, may still occur in participants with more practice.