Designing Long-Term Interaction for Robot-Assisted Recovery after Critical Injury

Carl Bettosi
cb54@hw.ac.uk
Heriot-Watt University
University of Edinburgh
Edinburgh, UK

Lynne Baillie
l.baillie@hw.ac.uk
Heriot-Watt University
Edinburgh, UK

Susan Shenkin
susan.shenkin@ed.ac.uk
University of Edinburgh
Edinburgh, UK

ABSTRACT

Social robots have demonstrated their potential to undertake aspects of a therapist's role, while showing a positive effect on user engagement. To deliver impact in the real world, these robots must have the capability to effect long-term decision making through continued personalization to 1) the patient's progressive exercise performance ability, and 2) their social behavior preferences. We undertook a series of collaborative workshops with domain experts to inform a rich knowledge representation for long-term upper limb recovery. We introduce a planned Hierarchical Reinforcement Learning (HRL) approach to exploit the natural hierarchical structure found in our problem, and promote more sample efficient learning, a common challenge in current RL-based HRI personalization. Our end goal is to deploy this system in-situ, over long-term engagement with recovering patients.

CCS CONCEPTS

• **Computing methodologies** → *Machine learning approaches*; • **Computer systems organization** → **Robotic autonomy**.

KEYWORDS

long-term interaction, social robotics, rehabilitation, reinforcement learning

1 INTRODUCTION

A well-structured progressive exercise program plays a central role in the recovery of patients after critical injury. Equally essential is the maintenance of high engagement levels, as it significantly influences overall program adherence [20]. However, achieving these objectives is often challenging due to resource shortages in struggling healthcare systems and low patient motivation. Recently, Social Robots (SRs) have emerged as a promising area of impact to guide aspects of physical exercise in homes, hospitals, and community centers [15]. Coupled with intelligent learning techniques, their embodiment gives rise to new capabilities that allow them to assume richer societal roles with the ability to personalize to aspects of the user, revealing an increased effect on engagement over the likes of virtual agents [7, 26]. Furthermore, robotic rehabilitation technology is generally used as a means to enhance the repetitive practice aspect of exercise routines, with higher-level decisions regarding the structure and direction of the program fully driven by the human therapist [14]. Although some level of human intervention is beneficial and, in certain instances necessary, greater autonomous decision-making could improve the efficiency of personalized rehabilitation program delivery.

Existing work has provided a glimpse into the potential impact of personalized SRs in the rehabilitation space, yet largely focuses on short-term interactions with limited decision-making capabilities with regard to program structure and direction. However, the relationship between a human patient and a therapist is often long-lasting and involves rich social exchange and learning over time that builds trust from an understanding of the patient's therapeutic needs and social preferences [7]. For robots to adopt such roles effectively, the ability to sustain long-term interaction is essential, and evaluation of these systems in-situ is imperative to advancing the field [2].

We introduce our design of a social robot that aims to personalize aspects of long-term upper limb rehabilitation after critical injury. Specifically, we look to enable the instruction of exercises based on the patient's exercise performance ability and the actioning of preferred social behaviors learned from periodic user feedback. To inform our design, we conducted several workshops with domain experts centered on the structure of the long-term program and the interactive components during a single session. We present the translation of the findings from this design exploration phase to a knowledge representation that can enable long-term decision making for our SR, driven by a Hierarchical Reinforcement Learning (HRL) mechanism. We then outline future steps, including feasibility testing through an initial pilot phase followed by simulation to evaluate the HRL approach. Our final goal is to evaluate our system over a long-term period of engagement with patients at a community rehabilitation center.

2 RELATED WORK

There has been considerable focus on the recent utilization of SRs within the area of physical rehabilitation. Popular applications involve the use of SRs as exercise companions during recovery sessions, providing motivational utterances during upper limb exercise [7, 8, 10, 23].

Beyond short-term interactions, SRs have also been evaluated over multiple sessions and more longitudinal timeframes. Irfan et al. present a robot for use during cardiac rehabilitation program with 26 participants over 2.5 years [11]. Aspects of personalization included progress tracking between sessions, attendance tracking, and automatic recognition of patients. The personalizable SR showed improved cardiovascular function over the long term and robot conditions revealed improved patient adherence. Pulido et al. describe a long-term study over 4 months with 8 pediatric patients, 15 sessions each, using a SAR for upper limb rehabilitation [18]. The system used expert-described poses sent to engineers using

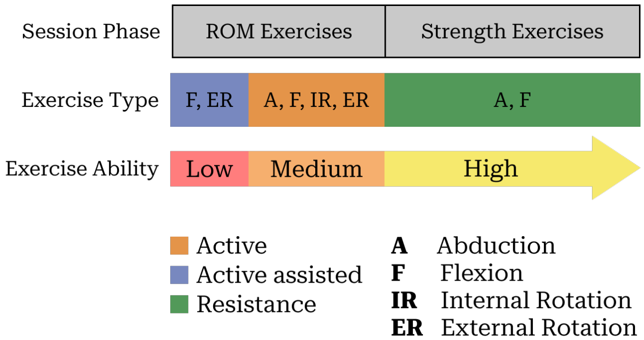

Figure 1: An overview of upper limb exercise progression, as informed through our workshop activities.

Planning Domain Definition Language (PDDL) for automated planning. While reporting increased subjective engagement and a slight rise in Range of Motion (ROM), the need for continuous human monitoring and potential challenges in adapting to unpredictable user behavior pose as potential limitations to achieving full autonomy. Winkle et al. undertook a 27-session study on the application of an Interactive Machine Learning (IML) approach to enhance autonomous SAR behavior during a couch-to-5k exercise session [27]. Challenges were noted in the timing of action selection, likely attributed to the choice of the learning algorithm—a basic adaptation of K-nearest neighbors.

These studies demonstrate the promising potential of SR personalization over long-term/multiple session rehabilitation, however, primarily concentrate on the adaption of lower-level social components rather than directing high-level program structure over the long-term. In wider HRI research, the area of long-term or 'lifelong' learning is an open challenge, yet machine learning techniques such as Reinforcement Learning (RL) can help achieve behavior personalization in complex HRI tasks [1]. For instance, Tapus et al. aims to utilize RL to match robot behavior such as proxemics, speed, and vocal content to user personality types during multiple evaluation studies involving post-stroke rehabilitation [24]. They revealed the RL approach was effective in almost all cases, although limited interaction experience likely hindered convergence to optimal robot behavior. RL, by nature, can prove challenging when training experience relies on real-world data [28]. In some instances, where human behavior can be simulated prior to real-world deployment, starting policies can be learned to boost learning convergence. Tsiakas et al. demonstrate this concept when attempting to adapt task difficulty during a cognitive therapy task based on performance and user engagement [25], where they generate simulated user behavior to evaluate an RL approach for robot personalization, based on an existing dataset.

3 DESIGN INSIGHTS FROM DOMAIN EXPERTS

To initiate the design phase, we first conducted an investigation of human-human physiotherapy practices to inform a rich knowledge representation for our system. To do this, we held a series of design workshops with expert physiotherapists from academia, the health technology industry, and the Scottish National Health Service (NHS). Workshops were divided into two key areas of exploration, *long-term program structure* and *single session interaction.* To ensure robust insight, we specified that all participants: 1) be a fully qualified physiotherapist; 2) have a minimum of 2 years of experience; 3) currently work with upper limb rehabilitation patients or have done so within the past 3 years. The workshops received full ethical approval from the University and all participants gave their consent to participate and be subjected to video recording (only relevant to Section 3.1). We used the Constant Comparative Method (CCM) of thematic analysis [9] to reach our findings discussed in the subsequent sections, allowing us to draw out emergent themes from the data.

3.1 Long-Term Program Structure

We conducted four separate workshops with six physiotherapists (1 practicing NHS physiotherapist/academic, 1 practicing NHS physiotherapist, 3 practicing NHS physiotherapists, and 1 fully trained physiotherapist working in rehabilitative robotics, respectively). Participants had an average of 5.3 years' experience in patient practice. Our objective was to fully understand the progressive stages involved in an upper-limb rehabilitation program, and to structure this information such that it could translate effectively to our system design for long-term interaction.

3.1.1 Procedure. We first held semi-structured interviews with participants. To help stimulate discussions, we prompted participants with three patient personas, each recovering from shoulder fractures at weeks 1-3, 3-6, and 6+ respectively, as aligned with existing NHS materials for self-managed recovery [16]. The talking points for each persona were 1) types of exercise at this stage, 2) details of exercise such as the number of repetitions, 3) measurements used to assess patient progress or ability, and 4) any other information relevant to the particular stage. We then used a Vicon motion capture system to record the participants performing three repetitions of each discussed exercise correctly. This allowed us to formulate a dataset of exercise recordings for use later during development to the robot's perception component (future work).

3.1.2 Findings. The early weeks of recovery (typically weeks 0-3) were deemed unlikely to be suitable for such a robotic intervention as the patients are performing gentle mobilization exercises to relieve stiffness. At weeks 3-5, patients begin to engage in repetitive range of motion (ROM) exercises such as flexion, abduction, internal and external rotations. Active assistive (the patient's affected limb is assisted) movement is started and gradually phased into active (no assistance) movements. Beyond week 6, ROM exercises are continued with the introduction of low-resistance strengthening exercises, gradually increasing to weight-bearing movement. We used this information to build a knowledge representation of the long-term recovery scenario for our system as shown in Figure 1. Participants stressed that the recovery timeline is highly flexible from patient-to-patient, but the general progression remains the same. Pain management is also of importance, using assessments such as the Visual Analogue Scale [21], and frequent assessment of ROM and strength to understand progress. Generalization of our system to other causes of upper limb injury were also of interest, such to maximize potential impact. Participants communicated that the proposed robot system will likely be able to use this program structure in other variations of shoulder fracture, stroke, frozen shoulder, and postoperative recovery.

3.2 Single Session Interaction

A workshop was carried out involving a group of seven physiotherapists who work in the NHS. Among them, two were also academics. The average experience of the participants was 14 years. Six participants took part in the long-term program structure workshop (Section 3.1). The aim of this session was to understand physiotherapist behaviors inside a single rehabilitation session, with a focus on any social actions or short-term considerations and how these may manifest during a one-to-one session.

3.2.1 Procedure. Ross et al. previously conducted systematic observations of upper limb exercise sessions and generated a dataset of physiotherapist behavior models, representing various styles of delivery [19]. The behavior actions of these models (i.e. *instruction*, *hustle*, *praise*, *console*, etc.) were influenced by the Arizona State University Observation Instrument [13], a popular tool to systematically record expert actions during sports coaching sessions. We utilized a subset of these models as design probes during our workshop, asking participants to provide granular feedback on the therapist behaviors as annotations on the models. Prior to this annotation session, we provided a short demonstration of the ARI humanoid robot by PAL robotics [6], the proposed social robot embodiment for this project. An example of an annotated behavior model can be seen in Figure 2. We additionally recorded any verbal discussions that arose during the workshop for later analysis.

3.2.2 Findings. Participants discussed the importance of balancing certain social actions such as praise, console, and hustle to the correct personality type, "Approach can differ per person. Some need more support or encouragement than others. It is very specific to the patient". Avoiding the repetition of phrases was also a concern for participants when considering the proposed robot system, "Being repetitive is fine for exercises, but not for communication". Close monitoring of patient emotional and physical state is also of high importance for both the effective direction of rehabilitation and patient safety, with specific focus on pain, fatigue, engagement, and frustration, "pain monitoring is important, [therapists] use a visual scale". Participants additionally stressed the importance of progress feedback to long-term engagement, "remind them of long-term improvements", "remind them that [they have] managed 100 reps. If you think about a week ago, [they] probably wouldn't have been able to do 10". With regard to the occurrences and flow of such social actions during the sessions, our model annotations helped clarify when and how frequently these should be applied. The key findings were the assessment of the physical and emotional state at the beginning of a session and the heavy use of verbal support during and between exercise repetitions to maintain participation and gather feedback on the patient's feelings during the session.

4 LONG-TERM PERSONALIZATION APPROACH

We introduce a Reinforcement Learning (RL) approach to address our goal of personalization. RL describes a problem where an agent

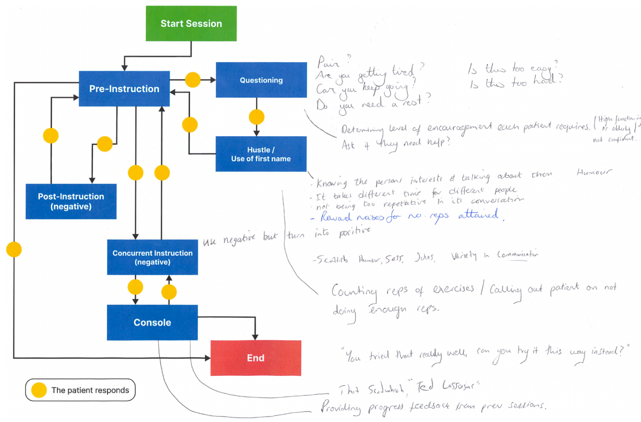

Figure 2: An example of an annotated behavior model from our single session interaction design workshop

interacting with an environment seeks to learn a policy of state-to-action mappings in an effort to maximize a cumulative reward over time [22]. This learning process manifests as interaction experience between the agent and the environment, progressively improving its behavior policy. Such an approach fits well to our use case, were we wish for our robot to learn the most optimal actions to take over time to maximize patient engagement and therapeutic outcome.

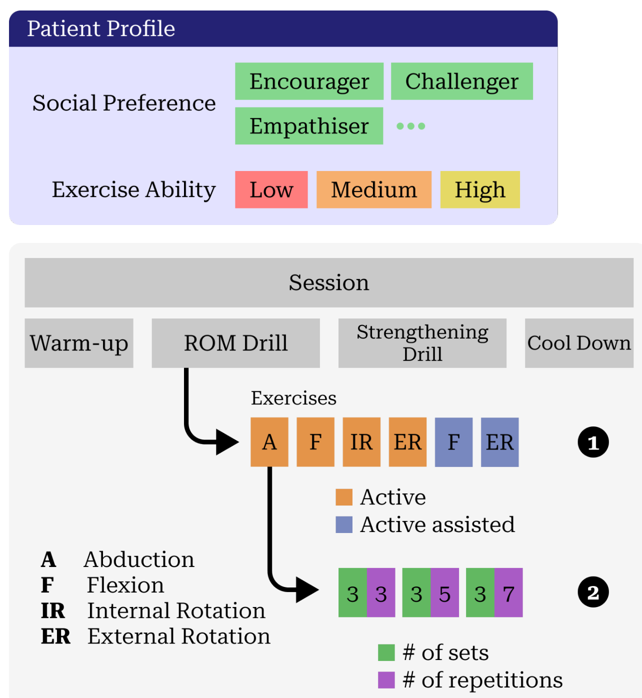

Figure 3: Our hierarchical learning structure. ROM Drill has been expanded to show an example of task decomposition. (1) and (2) show levels where learning will take place.

4.1 A Hierarchical RL Framework

Hierarchical Reinforcement Learning (HRL) has demonstrated potential within HRI personalization [3–5], being particularly well suited for tasks characterized by a natural hierarchy, wherein complex objectives can be systematically decomposed into smaller, more manageable subtasks [17]. We explore the potential of utilizing HRL as the structure of a recovery session, shown in Figure 3, naturally breaks down into sub-tasks which can be independently solved and, importantly, can be reused at different stages of the session. Furthermore, in Section 4.2 we discuss how HRL provides various advantages of traditional RL approaches that will help combat some of the common issues which arise in RL-based HRI personalization.

During interaction, a perception component will deliver input to the learning agent concerning real-time exercise performance levels and periodic user feedback. There are a wide array of potential techniques towards acquiring motion-tracking data such as computer vision or wearable sensor technologies. User feedback can be gathered between rests in exercise sets and could act as a strong indicator towards user perception of the current session. One of our next steps in this overarching work, discussed in Section 5, is to design and evaluate these perception technologies.

4.1.1 High-level controller. HRL allows for various behavior policies to be learned at each level of the task hierarchy, such that tasks will control the selection and execution of the sub-tasks below. In our framework, learning will take place at levels 1 and 2 as depicted in Figure 3. Furthermore, profile states that persist over multiple interaction episodes are necessary to enable ongoing learning, specifically, current exercise ability and the patient's social preference. As users naturally progress through varying physical abilities, they will enter new profile states, thus allowing our system to provide ongoing personalization.

The set of states pertaining to the high-level control tasks are as follows. ROM Drill, Strengthening Drill:
s{current exercise, exercise ability, previous completion rate}. Exercise ability is used to restrict lower abilities from accessing exercises which are out of scope and *previous completion rate* provides information on how well the patient performed previously inside each phase of the session, thus helping the policy to select exercises for the current session. The reward function at this level will be based on relative completion rate compared to the patient's previous session.

4.1.2 Low-level controller. The low-level controller will decide upon more fine-grained instructions such as the number of sets and repetitions (represented as pair *{set, repetition}*), as well as social actions. The set of states pertaining to the low-level control tasks is as follows. *s{previous performace, social preference}. Previous performance* will allow for a more fine-grained understanding of the patients performance in individual exercises, while *social preference* will guide the agent's social behaviors closer to that style of interaction. The reward function at this level will be based on relative performance compared to the patient's previous session, and high engagement sensed through periodic patient feedback.

4.2 Challenges

Attempting to optimize robot behavior from human interactions within real-world constraints introduces a variety of practical and theoretical challenges.

4.2.1 Sample efficiency. Learning optimal policies in such complex scenarios traditionally requires large amounts of training data. In the case of real-world HRI, data from human interaction experience is difficult to obtain [28]. There are various techniques that can be employed to mitigate the need for extensive learning: abstraction through simplified state space design and task decomposition enabled through the HRL framework; transfer learning were parts of the HRL allow for policies to be reused across various sub-tasks; learning through simulated users can generate starting policies to boost learning at the beginning of deployment.

4.2.2 User safety. Controlling autonomous decision-making in high-risk scenarios such as rehabilitation is paramount. Our robot should not be permitted access to actions that pose potential harm to users, for instance, the instruction of strengthening exercises for patients in the early stages of recovery. Several safeguarding measures will be taken to mitigate risk, specifically: frequent assessment of physical ability and user profiling; rule-based constraints to block certain profiles from entering high ability phases of the session; expert oversight to supervise system actions.

4.2.3 Moving goals. Since patients' physical abilities and therapeutic goals evolve over time, we would expect that the underlying relationship between actions and resulting rewards of our system would be non-stationary, thus threatening efficient learning. To mitigate this, our system will focus on relative progress from session to session, with a reward function which incentivizes performance and engagement progress.

5 FUTURE WORK AND CONCLUSION

Future steps include the development of an initial prototype of the system in which a second round of domain expert collaboration will be conducted to gather feedback on the interaction design, and validate the technical feasibility of a perception component for real-time data gathering on exercise performance quality and user feedback. Previous work in HRI has demonstrated how domain experts roleplaying as end users can create useful datasets for early feasibility testing [12]. We aim to exploit this technique to develop a set of simulated users at various stages of recovery, enabling the safe evaluation of our HRL framework in a simulated environment prior to real-world deployment. Our final goal is to evaluate our system over long-term engagement in a community rehabilitation setting.

In this work, we describe our learnings from workshop activities with domain experts in the upper limb rehabilitation domain. From this, we show how aspects of session interaction and, more interestingly, long-term program progression have translated into a comprehensive knowledge representation for the basis of a long-term robot personalization framework that will form the basis of our planned hierarchical reinforcement learning approach. We aim to optimize our system to user profiles based on ongoing assessments of physical performance ability during exercise, as well as personal behavior preference.

REFERENCES

[1] Neziha Akalin and Amy Loutfi. 2021. Reinforcement learning approaches in social robotics. *Sensors* 21, 4 (2021), 1292.

[2] Antonio Andriella, Carme Torras, and Guillem Alenya. 2020. Short-term human–robot interaction adaptability in real-world environments. *International Journal of Social Robotics* 12 (2020), 639–657.

[3] Jeanie Chan and Goldie Nejat. 2011. A learning-based control architecture for an assistive robot providing social engagement during cognitively stimulating activities. In *2011 IEEE International Conference on Robotics and Automation*. IEEE, 3928–3933.

[4] Jeanie Chan and Goldie Nejat. 2011. Minimizing task-induced stress in cognitively stimulating activities using an intelligent socially assistive robot. In *2011 RO-MAN*. IEEE, 296–301.

[5] Jeanie Chan and Goldie Nejat. 2012. Social intelligence for a robot engaging people in cognitive training activities. *International Journal of Advanced Robotic Systems* 9, 4 (2012), 113.

[6] Sara Cooper, Alessandro Di Fava, Carlos Vivas, Luca Marchionni, and Francesco Ferro. 2020. ARI: The social assistive robot and companion. In *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 745–751.

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

[8] Ronit Feingold-Polak, Oren Barzel, and Shelly Levy-Tzedek. 2021. A robot goes to rehab: a novel gamified system for long-term stroke rehabilitation using a socially assistive robot—methodology and usability testing. *Journal of neuroengineering and rehabilitation* 18, 1 (2021), 1–18.

[9] Barney G Glaser, Anselm L Strauss, and Elizabeth Strutzel. 1968. The discovery of grounded theory; strategies for qualitative research. *Nursing research* 17, 4 (1968), 364.

[10] Min Hun Lee, Daniel P Siewiorek, Asim Smailagic, Alexandre Bernardino, and Sergi Bermudez i Badia. 2023. Design, development, and evaluation of an interactive personalized social robot to monitor and coach post-stroke rehabilitation exercises. *User Modeling and User-Adapted Interaction* 33, 2 (2023), 545–569.

[11] Bahar Irfan, Nathalia Céspedes, Jonathan Casas, Emmanuel Senft, Luisa F Gutiérrez, Mónica Rincon-Roncancio, Carlos A Cifuentes, Tony Belpaeme, and Marcela Múnera. 2023. Personalised socially assistive robot for cardiac rehabilitation: Critical reflections on long-term interactions in the real world. *User Modeling and User-Adapted Interaction* 33, 2 (2023), 497–544.

[12] Joe Jeffcock, Mark Hansen, and Virginia Ruiz Garate. 2023. Transformers and human-robot interaction for delirium detection. In *Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*. 466–474.

[13] Alan C Lacy and Paul W Darst. 1984. Evolution of a systematic observation system: The ASU coaching observation instrument. *Journal of teaching in physical education* 3, 3 (1984), 59–66.

[14] Rui CV Loureiro, William S Harwin, Kiyoshi Nagai, and Michelle Johnson. 2011. Advances in upper limb stroke rehabilitation: a technology push. *Medical & biological engineering & computing* 49 (2011), 1103–1118.

[15] Abolfazl Mohebbi. 2020. Human-robot interaction in rehabilitation and assistance: a review. *Current Robotics Reports* 1 (2020), 131–144.

[16] NHS-Inform. 2021. Exercises for shoulder muscle and joint problems. https://www.nhsinform.scot/illnesses-and-conditions/muscle-bone-and-joints/exercises/exercises-for-shoulder-problems#seated-forearm-turn. Accessed: 2021-12-12.

[17] Shubham Pateria, Budhitama Subagdja, Ah-hwee Tan, and Chai Quek. 2021. Hierarchical reinforcement learning: A comprehensive survey. *ACM Computing Surveys (CSUR)* 54, 5 (2021), 1–35.

[18] Jose Carlos Pulido, Cristina Suarez-Mejias, Jose Carlos Gonzalez, Alvaro Duenas Ruiz, Patricia Ferrand Ferri, Maria Encarnacion Martinez Sahuquillo, Carmen Echevarria Ruiz De Vargas, Pedro Infante-Cossio, Carlos Luis Parra Calderon, and Fernando Fernandez. 2019. A socially assistive robotic platform for upper-limb rehabilitation: a longitudinal study with pediatric patients. *IEEE Robotics & Automation Magazine* 26, 2 (2019), 24–39.

[19] Martin Ross, Frank Broz, and Lynne Baillie. 2021. Observing and clustering coaching behaviours to inform the design of a personalised robotic coach. In *Proceedings of the 23rd international conference on mobile human-computer interaction*. 1–17.

[20] Martin Saebu, Marit Sørensen, and Hallgeir Halvari. 2013. Motivation for physical activity in young adults with physical disabilities during a rehabilitation stay: a longitudinal test of self-determination theory. *Journal of Applied Social Psychology* 43, 3 (2013), 612–625.

[21] Jane Scott and EC Huskisson. 1976. Graphic representation of pain. *pain* 2, 2 (1976), 175–184.

[22] Richard S Sutton and Andrew G Barto. 2018. *Reinforcement learning: An introduction*. MIT press.

[23] Katelyn Swift-Spong, Elaine Short, Eric Wade, and Maja J Matarić. 2015. Effects of comparative feedback from a socially assistive robot on self-efficacy in post-stroke rehabilitation. In *2015 IEEE International Conference on Rehabilitation Robotics (ICORR)*. IEEE, 764–769.

[24] Adriana Tapus, Cristian Ţăpuş, and Maja J Matarić. 2008. User—robot personality matching and assistive robot behavior adaptation for post-stroke rehabilitation therapy. *Intelligent Service Robotics* 1 (2008), 169–183.

[25] Konstantinos Tsiakas, Maher Abujelala, and Fillia Makedon. 2018. Task engagement as personalization feedback for socially-assistive robots and cognitive training. *Technologies* 6, 2 (2018), 49.

[26] Valentina Vasco, Cesco Willemse, Pauline Chevalier, Davide De Tommaso, Valerio Gower, Furio Gramatica, Vadim Tikhanoff, Ugo Pattacini, Giorgio Metta, and Agnieszka Wykowska. 2019. Train with me: a study comparing a socially assistive robot and a virtual agent for a rehabilitation task. In *International Conference on Social Robotics*. Springer, 453–463.

[27] Katie Winkle, Séverin Lemaignan, Praminda Caleb-Solly, Paul Bremner, Ailie J Turton, and Ute Leonards. 2020. In-Situ Learning from a Domain Expert for Real World Socially Assistive Robot Deployment.. In *Robotics: Science and Systems*.

[28] Yang Yu. 2018. Towards Sample Efficient Reinforcement Learning.. In *IJCAI*. 5739–5743.