A Motivational Robotic Coach for Repetitive Individual Training

Martin K Ross
Computer Science Department
Heriot-Watt University
Edinburgh, Scotland
mkr30@hw.ac.uk

Frank Broz
Computer Science Department
Delft University of Technology
Rotterdam, Netherlands
f.broz@tudelft.nl

Lynne Baillie
Computer Science Department
Heriot-Watt University
Edinburgh, Scotland
l.baillie@hw.ac.uk

***Abstract*—Individual, unsupervised practice of repetitive exercises is used frequently in both rehabilitation after stroke and training for squash. In both cases, motivational struggles are common when a physiotherapist or coach is not present and this can cause a lack of adherence over the long-term. In this paper, we present an overview of our work on creating and evaluating a robotic coach capable of motivating users in both of these use cases. We discuss our ongoing work focusing on high level personalisation to groups of users and how this could provide a starting point for a lifelong learning system capable of further adaption to individuals.**

Keywords—coaching, personalisation, squash

I. Introduction

Repetitive, solo practice (i.e. without the supervision of a coach) can improve the skill level of sports players and is used regularly by high-performance players in a variety of sports [1], one of which is squash. However, it is used much less frequently by players at lower levels [1], indicating a lack of motivation when a coach is not present. Similarly, research in rehabilitation techniques after stroke strongly suggest that home-based rehabilitation (i.e. without the supervision of a rehabilitation therapist) is beneficial to the patient [2]. However it is often not adhered to due to (among other reasons) a lack of motivation [3]. We have chosen to consider both of these case studies, from different domains, in the same body of work due to the similarities in the individual, often unsupervised and repetitive nature of practice which helps in making long term functional improvements after stroke [2] and helps high performance sports players improve their skill level [1].

Personalisation (in the form of both high-level personalisation to groups of users based on user type and low-level adaption to individual users over time) has been suggested as a key attribute of systems aiming to motivate users to conduct physical activity [4]. A robotic coaching system that is capable of high-level personalisation, combined with lifelong learning to adapt its behaviour to users during their individual exercise activities could provide extra motivation. A result of this could be an increase in adherence to individual exercise, with the potential of improved rehabilitation results for stroke survivors and squash players improving their skill level. In this paper we present an overview of ongoing studies exploring the use of high-level personalisation in a robotic coach of squash and stroke rehabilitation. We then describe our planned future work to include long-term Reinforcement Learning (RL) in our system.

II. Background

Stroke is one of the leading causes of acquired adult disability with survivors commonly suffering permanent impairments such as fatigue, weakness in the arms and legs, aphasia and forgetfulness [5]. Squash, on the other hand, is an intermittent, high-intensity racket sport (shown in Fig. 11), with matches contested over the best of 5 games in singles competition [6].

In sports coaching, praise for independent practice given by coaches increases the intrinsic motivation of the athlete and their intention to remain physically active in the future [7]. In sports such as squash, this independent practice is used by many of the top professionals and can include repetitive solo drills with little or no input from coaches. This type of repetitive, individual drill can improve a player's skill level in a variety of sports but is used less frequently by non-professional players [1]. In rehabilitation it has also been suggested that social factors such as the behaviour of professionals interacting with stroke survivors, and the relationship held between these two parties, can have both positive and negative impacts on patient motivation [8], [9]. Therefore, a robotic coach which could be used at times when a human practitioner is not available and is capable of providing its user with motivation to conduct these individual exercise sessions in both domains could help non-professional squash players improve their skill level and stroke survivors see improved rehabilitation results.

Previous work has confirmed the potential of the use of an Human Robot Interaction (HRI) system to motivate users to perform physical exercise. Examples include a Pepper robot being used as a running coach [10], a Nao robot as a cycling instructor [11], and a humanoid robot as a motivator for high intensity interval training [12]. These studies point to the effectiveness of an embodied device in providing motivation during physical activity. However, they do not offer technical advice on a specific skill or provide personalisation of behaviours to individual users. This study is the first of its kind to evaluate a Socially Assistive Robot (SAR) in the context of training for squash, or any other skill-based sport (as opposed to fitness training). In the rehabilitation domain, robots have also been considered as rehabilitation coaches, both for children with cerebral palsy [13], and for stroke survivors [14], [15]. Although these studies present promising results, low participant numbers in [15] and short-term interactions in [14] indicate that more research is required, particularly with regards to lifelong learning and personalisation which these studies did not investigate.

Robotic systems employing strategies intended to build the relationship between the user and the robot (e.g. by using continuity behaviours, and the user's name) have been preferred by users over systems that are purely functional [16]. It has been suggested that both high-level personalisation to groups of users and low-level adaption to individuals is required in a robotic coach [4].

One promising method of achieving high-level personalisation is to learn from human demonstrations. In a collaborative packing task, Nikolaidis et al. showed that by clustering human demonstrations into similar styles and applying inverse reinforcement learning over these clusters, it was possible to learn a reward function that was representative of each user type [17]. In our current work, we focus on high-level personalisation. We have applied a similar strategy to the more interactive, open scenario of coaching. Previous work with domain experts [18] has formalised 12 coaching policies using Nikolaidis et al.'s clustering algorithm. Our robotic coach can select from these 12 policies based on its user's information and training context (see Section III), thus achieving high-level personalisation.

To the author's knowledge, this is the first attempt to evaluate a system capable of personalising its coaching approach to different user types in these contexts. Although we do not specifically focus on lifelong learning in this paper, Silva and Gombolay [19] show that a similar method of encoding domain knowledge can (in their non-interactive scenario) provide a starting point for a Reinforcement Learning (RL) algorithm to further adapt the system's behaviour to individual users. Some possible methods to tackle this problem in the context of our work are discussed further in Section VI.

III. System Desciption

A full description of the system implementation is out with the scope of this paper, but an overview is given here.

The system (implemented on the Pepper robot) provides coaching to users during practice/exercise sessions through the selection of appropriate actions (examples shown in TABLE I) from its chosen coaching policy. These policies are taken from Ross et al.'s work [18]. In that work, the authors combine data collection methods adapted from sports coaching literature with computational techniques and mathematical modelling to define a process used to formalise human knowledge in the form of 'coaching policies'. They also give suggestions on which of the policies were likely to be more appropriate for users with different traits [18]. Using observations of Human-Human Interactions (HHIs), they first obtained action sequences of behaviours exhibited by professional squash coaches and stroke physiotherapists. These were primarily the same behaviours as used by our robotic coach and shown in TABLE I. They then clustered these action sequences into 'behaviour graphs'. Each graph is a visual representation of a coaching policy and can be viewed by following the link in the footnote[1]. Next they obtained coaches' and physiotherapists' reflections on the graphs' applicability to the real world and appropriateness for different training contexts and user information. The policy used by our robotic coach is selected from one of the 12 policies learned in [18] based on the following user information: squash playing ability or effects of stroke (self-provided), number of sessions performed with the robotic coach (i.e. length of the relationship), motivation for training (self-rated), and type of session. The policies used are derived from both coaching and rehabilitation data, so they should be generalisable across these two use cases.

The robotic coaching system was implemented on a Pepper robot using Pepper's Python SDK. In the squash setting, it communicates with a sensor attached to the user's squash racket, via a commercial mobile application, to track the player's swing. To track a user's movements in the rehabilitation setting, a vision system incorporating OpenPose [20] with the video footage captured through Pepper's head-mounted depth camera was implemented. Sessions are comprised of either sets of shots played (squash) or exercises performed (rehabilitation) by the user. By performing a range of coaching behaviours similar to those performed by a human coach/physiotherapist, the robot leads its user through their training session. Behaviours are primarily animated utterances spoken by the robot (i.e. Pepper moves in a human-like way while speaking), but can also include demonstrations via the robot's movements. For example, the robot might perform a pre-instruction behaviour, praise, or a post instruction behaviour while demonstrating the correct arm position for the player's racket preparation or starting point for a specific exercise. After receiving data from the sensor/vision system, the chosen coaching policy generates an appropriate action to be performed by the robot. These actions are grouped into 13 behavioural categories (shown in TABLE I) and are combined to coach a user through their solo practice session.

TABLE I. The 13 behavioural categories the robot can select from.

Category	Example
Pre-Instruction	"In the next set, let's make sure on every shot you play your racket face stays open as you strike the ball."
Concurrent Instruction (Positive)	"Racket up"
Concurrent Instruction (Negative)	"Your racket's not high enough."
Post Instruction (Positive)	"Your racket preparation got better in that practice which was great! You got an average score of 79 and were aiming for 84."
Post Instruction (Negative)	"Today, your follow through didn't manage to improve. You got an average score of 0.17 and were aiming for 0.12."
Questioning	"How did your forehand drive feel there?" (The user would respond using the touch sensors on the robot's head and hands.)
Positive Modelling	Demonstrates correct arm position for racket preparation.
Negative Modelling	Demonstrates swinging the arm but stopping the follow through too quickly.
First Name	"Pepper"
Praise	"Nice!"
Hustle	"Big push!"
Scold	"That was a bad one"
Console	"Hard lines"

[1]Visualisations of the coaching policies used by our robotic system can be found at: https://github.com/M4rtinR/BehaviourGraphVisualisations

IV. HYPOTHESES

Based on previous observation and interview studies by Ross et al. [18], and on the literature surrounding robotic coaching, we have made 4 hypotheses regarding the evaluation of our system and referring to the DSP (Data Selected Policy), NPP (No Personalisation Policy) and NCP (No Coaching Policy) conditions defined in Section V B. We expect that using a participant's information and training context for high-level personalisation by selecting the most appropriate policy will outperform a data-based policy that isn't personalised, and exercise without any coaching.

H1. Participants will make greater improvements in technique/have a higher exercise completion rate in the DSP condition than in the NPP and NCP conditions.

H2. Participants will view the DSP condition as a more effective coach than the NPP and NCP conditions.

H3. Participants will be more motivated to conduct individual exercise when using the DSP condition compared to the NPP and NCP conditions.

H4 (Squash). Participants will view the DSP condition as more socially competent than the NPP and NCP conditions.

H4 (Rehab). Participants will have a lower perceived cognitive and physical workload in the DSP condition than in the NPP and NCP conditions.

V. SHORT-TERM STUDY PROCEDURE

A. Participants

Two separate within-subjects studies will be run to evaluate the hypotheses (**H1** – **H4**) given in Section IV. One has recently concluded (although the results have not yet been analysed) and involved non-professional squash players and the other will begin soon with stroke survivors. We had 16 participants complete all 3 sessions from our squash study and are aiming for a similar number of stroke survivors in our upcoming study.

B. Conditions

In both studies, three conditions will be evaluated. In the **Data Selected Policy (DSP)** condition, participants will interact with the robot executing the policy chosen using the method described in Section III. In the **No Personalisation Policy (NPP)** condition, the robot will execute a randomly selected policy from the other 11 policies that are not the best match for the participant's user information and training context. Comparing these two conditions will allow us to discover the effect of the high-level personalisation derived from choosing an appropriate policy, compared to a randomly selected policy. The selection of the random policy will be performed at the beginning of the interaction. A **No Coaching Policy (NCP)** baseline condition will also be used in which the robot will tell the user which shot/exercise to perform and when to perform each set of shots/exercises. No coaching behaviours will be used by the robot in this condition. It is therefore the closest condition to a regular solo practice session in squash or individual exercise session in rehabilitation.

A within-subject design will be used in both studies, with each participant interacting with all three conditions for 15-30 minutes each. The three interactions will be split across two different days (see subsection D). The order in which participants interact with each of the three conditions will be counterbalanced.

C. Measures

The following measures will be used in both studies to gather appropriate data to evaluate our hypotheses:

a) The "technical skills" subscale of the *Coaching Behaviour Scale for Sport (CBS-S) [21]* will be used to measure participants' subjective opinions on the coaching provided by each robot condition, allowing evaluation of **H2**. The wording will be adapted slightly for use in rehabilitation.

b) The interest/enjoyment, perceived competence, perceived choice and value/usefulness subscales of the *Intrinsic Motivation Inventory (IMI) [22]* will be used to assess the effect of each condition on the participants' intrinsic motivation for conducting individual exercise with the robot, thus allowing the evaluation of **H3**.

c) Exercise/Shot statistics: As described in Section III, a sensor mounted on the end of the participant's racket will provide a score for each shot played in the squash sessions. In the stroke rehabilitation sessions, the number of completed repetitions of a given exercise will be used. Comparing these statistics will allow evaluation of **H1**.

Additionally, in the squash study, the *Robotic Social Attributes Scale (RoSAS)* [23] will be used as a subjective measure of the social competence of the robotic coach to allow evaluation of **H4 (Squash)**. **H4 (Rehab)** will be evaluated using the *NASA Task Load Index (TLX) [24].*

D. Study Design

The squash study will take place on a hard-back squash court in the university's sports centre. A squash racket with a sensor attached will be provided to all participants and sanitised between sessions. The rehabilitation study will take place in a laboratory on the university campus. All necessary equipment to complete the given exercises (e.g. cane, towel, water bottle) will be provided to all participants and sanitised between sessions.

TABLE II shows an overview of the procedure to be used in the studies. Each participant will attend the facility on 2 separate days. This was deemed to strike the right balance between mitigating against fatigue and mitigating against self- isolation requirements of participants or mid-study COVID restriction changes. To further ensure fatigue does not play a role, on the second day participants will be given a minimum of 10 minutes break (squash study) or 45 minutes break (rehabilitation study) between sessions.

The setup to be used during interactions in both sessions is shown in Fig. 1 and Fig 2. In all 3 conditions, the session will be coordinated autonomously by the robot.

TABLE II. SUMMARY OF THE PROCEDURE USED FOR EACH PARTICIPANT.

Day	Phase	Activity	Duration
1	***Intro***	Researcher will explain the study and answer any questions. Participant will sign the consent form and complete the demographic questionnaire.	10 mins
	Session 1	The participant will be guided through a squash/rehabilitation session under one of the three robotic conditions. In the squash study, this will be split into sets of shots. Whereas in the rehabilitation study this will be split into sets of different exercises targetting their stroke-affected limb.	15-30 mins
	Qs 1	Participant will complete the evaluation questionnaires.	10 mins
			Day 1 total: 35 - 50 mins
	Session 2	As in **Session 1** using a different condition.	15-30 mins
	Qs 2	As in **Qs 1**.	10 mins
	Break	To mitigate against fatigue, the participant in the rehabilitation study will be given a 45 minute break before proceeding. In the squash study, the 10 minute break while completing the questionnaires is sufficient rest for this type of training.	0-45 mins
	Session 3	As in **Session 1** using the third condition.	15-30 mins
	Qs 3	As in **Qs 1**.	10 mins
	Wrap-up	Researcher will answer any final questions the participant had.	5 mins
			Day 2 total: 55 mins – 2hrs 10mins

Fig. 1. The squash experimental setup - the robotic coach was on court with the participant, acting autonomously. The researcher observed from the balcony so had a perspective similar to the one shown in this image.

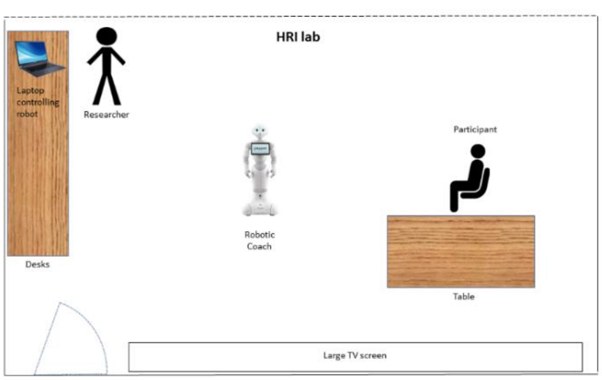

Fig. 2. The rehabilitation experimental setup – the robotic coach will be positioned in front of the participant, who will be seated next to a table (required for some of the exercises).

VI. FUTURE WORK (LIFELONG LEARNING FOR PERSONALISATION)

Much of our work to this point has focussed on high-level personalisation to groups of users during short-term interactions. However, we also plan to include low level adaption to individual users so that our system can continue to personalise its behaviour to a user over the long-term to improve the interaction. Both squash training and recovery after stroke require consistent effort over a long period of time. Therefore personalisation using lifelong learning will be a key part of a successful system in both cases.

A number of different methods have been explored to achieve continued personalisation of a robot's actions. For example, using RL to personalise the teaching behaviours of a robot has been shown to result in higher levels of positive valence towards the robot [25] and more effective teaching [26]. This is a promising method of low-level adaption to individuals within sessions and throughout the life of an interactive system.

A key challenge in our work (as with many RL applications) will be choosing an appropriate reward function. Previous work [25] has used a facial expression analysis system to measure children's valence and engagement in a second-language learning task and based the reward function on these factors to achieve personalisation. Another work [26] instead utilised a user performance parameter as the basis of the reward function. The details of the implementation for our system are still being considered but previous works such as these indicate that our system could continue to improve and adapt over an extended period of time.

VII. CONCLUSION

Personalisation of a robotic coaching system could provide extra motivation for non-professional squash players and people recovering from a stroke during individual exercise. We have presented our method for high-level personalisation of a robotic coach usable in both domains, as well as discussing how an RL algorithm could be used for lifelong learning to further adapt the behaviour of such a system to individuals over the long term.

ACKNOWLEDGMENT

Thank you to Racketware (https://www.racketware.co.uk/) for their help with this project and for providing their sensor for use during the study. This work is funded by the Engineering and Physical Sciences Research Council (EPSRC), Grant.ID: EPSRC DTP18.

REFERENCES

[1] J. Baker, J. Côté, and B. Abernethy, "Learning from the Experts: Practice Activities of Expert Decision Makers in Sport," *Research Quarterly for Exercise and Sport*, vol. 74, no. 3, pp. 342–347, Sep. 2003, doi: 10.1080/02701367.2003.10609101.

[2] S. Hillier and G. Inglis-Jassiem, "Rehabilitation for community-dwelling people with stroke: home or centre based? A systematic review," *International Journal of Stroke*, vol. 5, pp. 178–186, 2010, doi: 10.1111/j.1747-4949.2010.00427.x.

[3] K. K. Miller, R. E. Porter, E. DeBaun-Sprague, M. V. Puymbroeck, and A. A. Schmid, "Exercise after Stroke: Patient Adherence and Beliefs after Discharge from Rehabilitation," *Topics in Stroke Rehabilitation*, vol. 24, no. 2, pp. 142–148, Feb. 2017, doi: 10.1080/10749357.2016.1200292.

[4] K. Winkle, P. Caleb-Solly, A. Turton, and P. Bremner, “Social Robots for Engagement in Rehabilitative Therapies: Design Implications from a Study with Therapists,” in *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction - HRI ’18*, 2018, pp. 289–297. doi: 10.1145/3171221.3171273.

[5] World Health Organisation, “WHO EMRO | Stroke, Cerebrovascular accident | Health topics,” 2018. http://www.emro.who.int/health-topics/stroke-cerebrovascular-accident/index.html

[6] N. Gibson, P. Bell, A. Clyne, G. Lobban, L. Aitken, and K. Gibbon, “Physical Preparation for Elite-Level Squash Players : Monitoring , Assessment , and Training Practices for the Strength and Conditioning Coach,” *Strength and Conditioning Journal*, 2019.

[7] B. J. Almagro, P. Sáenz-López, and J. A. Moreno, “Prediction of sport adherence through the influence of autonomy-supportive coaching among spanish adolescent athletes,” *Journal of Sports Science and Medicine*, vol. 9, no. 1, pp. 8–14, 2010.

[8] N. Maclean, P. Pound, C. Wolfe, and A. Rudd, “Qualitative analysis of stroke patients’ motivation for rehabilitation,” *BMJ*, vol. 321, no. 7268, pp. 1051–1054, Oct. 2000, doi: 10.1136/bmj.321.7268.1051.

[9] N. Maclean, P. Pound, C. Wolfe, and A. Rudd, “The Concept of Patient Motivation: A Qualitative Analysis of Stroke Professionals’ Attitudes,” *Stroke*, vol. 33, no. 2, pp. 444–448, Feb. 2002, doi: 10.1161/hs0202.102367.

[10] K. Winkle, S. Lemaignan, P. Caleb-Solly, U. Leonards, A. Turton, and P. Bremner, “Couch to 5km Robot Coach: An Autonomous, Human-Trained Socially Assistive Robot,” in *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, Cambridge, United Kingdom, Mar. 2020, pp. 520–522. doi: 10.1145/3371382.3378337.

[11] L. Sussenbach *et al.*, “A robot as fitness companion: Towards an interactive action-based motivation model,” *IEEE RO-MAN 2014 - 23rd IEEE International Symposium on Robot and Human Interactive Communication: Human-Robot Co-Existence: Adaptive Interfaces and Systems for Daily Life, Therapy, Assistance and Socially Engaging Interactions*, pp. 286–293, 2014, doi: 10.1109/ROMAN.2014.6926267.

[12] D. J. Rea, S. Schneider, and T. Kanda, “‘Is this all you can do? Harder!’: The Effects of (Im)Polite Robot Encouragement on Exercise Effort,” in *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*, Boulder CO USA, Mar. 2021, pp. 225–233. doi: 10.1145/3434073.3444660.

[13] N. A. Malik, F. A. Hanapiah, R. A. A. Rahman, and H. Yussof, “Emergence of Socially Assistive Robotics in Rehabilitation for Children with Cerebral Palsy: A Review,” *International Journal of Advanced Robotic Systems*, vol. 13, no. 3, 2016, doi: 10.5772/64163.

[14] E. Wade, A. R. Parnandi, and M. J. Matarić, “Using Socially Assistive Robotics to Augment Motor Task Performance in Individuals Post – Stroke,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2011, pp. 2403–2408.

[15] R. Feingold Polak and S. L. Tzedek, “Social Robot for Rehabilitation: Expert Clinicians and Post-Stroke Patients’ Evaluation Following a Long-Term Intervention,” in *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, Cambridge, United Kingdom, Mar. 2020, pp. 151–160. doi: 10.1145/3319502.3374797.

[16] B. Irfan, M. Hellou, A. Mazel, and T. Belpaeme, “Challenges of a Real-World HRI Study with Non-Native English Speakers: Can Personalisation Save the Day?,” in *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, Cambridge, United Kingdom, Mar. 2020, pp. 272–274. doi: 10.1145/3371382.3378278.

[17] S. Nikolaidis, R. Ramakrishnan, K. Gu, and J. Shah, “Efficient Model Learning from Joint-Action Demonstrations for Human-Robot Collaborative Tasks,” in *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction - HRI ’15*, Mar. 2015, pp. 189–196. Accessed: Mar. 02, 2020. [Online]. Available: https://dl.acm.org/doi/10.1145/2696454.2696455

[18] M. K. Ross, F. Broz, and L. Baillie, “Observing and Clustering Coaching Behaviours to Inform the Design of a Personalised Robotic Coach,” presented at the 23rd International Conference on Human-Computer Interaction with Mobile Devices and Services, Virtual (originally Toulouse, France), Sep. 2021. doi: 10.1145/3447526.3472043.

[19] A. Silva and M. Gombolay, “Encoding Human Domain Knowledge to Warm Start Reinforcement Learning,” *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, no. 6, Art. no. 6, May 2021.

[20] Z. Cao, G. Hidalgo, T. Simon, S.-E. Wei, and Y. Sheikh, “OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields,” *ArXiv: 1611.08050*, May 2019, Accessed: Feb. 01, 2022. [Online]. Available: https://arxiv.org/abs/1812.08008v2

[21] J. Côté, J. K. Yardley, J. Hay, W. A. Sedgwick, and J. Baker, “An Exploratory Examination of the Coaching Behavior Scale for Sport,” *AVANTE*, vol. 5, no. 3, pp. 82–92, 1999.

[22] R. M. Ryan, “Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory. - PsycNET,” *Journal of Personality and Social Psychology*, vol. 43, no. 3, pp. 450–361, 1982.

[23] C. M. Carpinella, A. B. Wyman, M. A. Perez, and S. J. Stroessner, “The Robotic Social Attributes Scale (RoSAS): Development and Validation,” in *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, Vienna Austria, Mar. 2017, pp. 254–262. doi: 10.1145/2909824.3020208.

[24] S. G. Hart and L. E. Staveland, “Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research,” in *Advances in Psychology*, vol. 52, P. A. Hancock and N. Meshkati, Eds. North-Holland, 1988, pp. 139–183. doi: 10.1016/S0166-4115(08)62386-9.

[25] G. Gordon *et al.*, “Affective personalization of a social robot tutor for children’s second language skills,” in *Proceedings of the 30th Conference on Artificial Intelligence (AAAI 2016)*, 2016, pp. 3951–3957. doi: 10.1016/j.bbmt.2012.12.001.

[26] S. Roy, E. Kieson, C. Abramson, and C. Crick, “A reinforcement learning model for robots as teachers,” in *27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 2018, pp. 294–299.