Individual Squash Training is More Effective and Social with a Humanoid Robotic Coach*

Ross, Martin K.; Broz, Frank; Baillie, Lynne

DOI
10.1109/RO-MAN57019.2023.10309567

Publication date
2023

Document Version
Final published version

Published in
2023 32nd IEEE International Conference on Robot and Human Interactive Communication, RO-MAN 2023

2023 32nd IEEE International Conference on Robot and
Human Interactive Communication (RO-MAN)
August 28-31, 2023. Paradise Hotel, Busan, Korea

2023 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN) | 979-8-3503-3670-2/23/$31.00 ©2023 IEEE | DOI: 10.1109/RO-MAN57019.2023.10309567

Individual Squash Training is More Effective and Social with a Humanoid Robotic Coach *

Martin K. Ross, Frank Broz, and Lynne Baillie

Abstract—**With the aim of providing extra motivation to adhere to repetitive, individual sports training, this paper presents an autonomous robotic squash coach capable of high-level personalisation. The system was evaluated in person with 16 participants each conducting three 15-minute solo practice sessions. We compared a baseline, non-coaching robotic condition to two conditions in which the robot executed one of 12 different coaching policies, each of which was based on human coaching data. In one of the coaching conditions, the policy was selected based on categories for personalisation and in the other it was selected randomly among policies. The coaching policy conditions were found to be more enjoyable, more socially competent, and perceived as a more effective coach than the baseline.**

I. Introduction

Repetitive, solo practice (i.e. without the supervision of a coach) can improve a player's skill level, and is used regularly by high-performance players in a variety of sports [1]. However, it is used less frequently by players at lower levels [1], indicating a lack of motivation when a coach is not present. One such sport is squash: an intermittent, high-intensity racket sport played on an indoor court (Figure 1).

Previous work has confirmed the potential of the use of a Socially Assistive Robot (SAR) to motivate users to perform physical exercise. Examples include using a Pepper robot as a running coach [2], and a Nao robot as a cycling instructor [3]. However, the current work is the first of its kind to evaluate a SAR in the context of squash coaching, or technical coaching for any skill-based sport (as opposed to fitness training).

An emerging requirement of such a robotic coach is personalisation, both in terms of high-level personalisation to groups of users and low-level adaption to individuals over time [4]. Past studies (e.g. [2], [3]) point to the effectiveness of an embodied device in providing motivation during physical activity. However, the systems do not offer personalised behaviours to users. Throughout this paper the term 'personalised' refers to high-level personalisation based on user information and training context but we acknowledge that continual adaption to individuals over time would also be needed to meet the full personalisation requirement [4].

* This work was funded by the Engineering and Physical Sciences Research Council (EPSRC) Grant ID: EPSRC DTP18.

Martin K. Ross and Lynne Baillie are with Heriot-Watt University, Edinburgh, EH14 4AS, Scotland (corresponding author e-mail: mkr30@hw.ac.uk).

Frank Broz is with Delft University of Technology, Mekelweg 5, 2628 CD Delft, Netherlands.

To the authors' knowledge, this is the first attempt to evaluate a system capable of personalising its coaching approach to users with different information and training contexts. We conducted a within subjects study measuring coaching effectiveness, intrinsic motivation and the robot's social competence during 15-minute training sessions. We aimed to show that using a robotic coach is more effective and enjoyable than a regular solo practice session in squash.

II. Background

A. Robotic Coaching for Sport/Physical Activity

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 [5]. In squash, this independent practice is used by many of the top professionals but the repetitive, individual drills are performed less frequently by non-professional players in a variety of sports [1]. A robotic coaching system that is used by players during solo practice sessions in squash could provide extra motivation for training.

Sussenbach et al. showed the potential of using a SAR to engage a user in an individual exercise routine [3]. By first creating a motivational model based on human-human interaction (HHI) observations, a robotic cycling instructor (Nao) was created that elicited better training effects (measured as a decrease in participants' resting heart rate), more intensive workouts, and higher training motivation in participants compared to a textual control system. Ongoing work is also exploring the use of a robot (Pepper) to coach users through the 'couch to 5km' running programme [2]. By having a domain expert manually correct the behaviour of the robotic coach (Pepper) during sessions, the system was able to learn (using a k-nearest neighbours algorithm) and perform very similar behaviours when acting autonomously. The system engaged users for an average of 15.4 hours over 3 months. However, the possibility of using a similar system in a skill-based sport remains unexplored. Furthermore, while both are valid approaches, Winkle et al.'s method required a domain expert to be present during robotic coaching sessions, and the motivational model used by Sussenbach et al. focused more on the structure of a session rather than the coaching behaviours that would be key in a skill-based sport. In our system, we use human data differently to formalise coaching policies [6] that directly control the robot.

B. High-Level Personalisation Methods

Robotic systems employing strategies intended to build a relationship with the user (e.g. by remembering past sessions and using the user's name) have been preferred over systems that are purely functional [7]. Because of this, personalisation has been suggested as a requirement of a robotic coach [4].

One promising method of achieving high-level personalisation to groups of users 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 [8]. In this work, we evaluate the application of a similar strategy to the more interactive, open scenario of coaching.

By combining data collection methods adapted from sports coaching literature with computational techniques and mathematical modelling, we have previously defined a process to formalise human knowledge in the form of 'coaching policies' usable for robotic control [6]. A policy in this context is a mapping from states to actions. In this previous work we considered both squash coaching and rehabilitation after stroke, learning domain independent coaching policies from two separate use cases. We chose to study these two use cases together due to the similarities in the repetitive, individual, and often unsupervised nature of practice over a long period of time which is required to make functional improvements in both. This is a unique feature of our work and is in contrast to previous studies (e.g. [3]) that only learn from the target domain. In the current work, our evaluation focuses only on squash, but the presented system and personalisation method is generalisable across domains due to our novel methodology.

Starting from observations of professional squash coaches and stroke physiotherapists, we used Nikolaidis et al.'s clustering algorithm to generate 12 unique coaching policies [6]. Each of these robot-executable policies is based on real behavioural data. Visualisations of the policies as 'behaviour graphs' (on GitHub[1]) were discussed with coaches and physiotherapists during semi-structured interviews. They confirmed the policies' applicability across domains, and made actionable suggestions as to which policies were likely to be more appropriate for which groups of users. The robotic coach evaluated in the current work uses these suggestions to select a best-fit policy and achieve personalisation based on user information and training context.

This type of personalisation goes beyond that which has previously been explored in the context of robotic coaching. Most past works have focused on customising the utterances of a system with the name and performance history of its user (e.g. [9], [10]). While findings suggest that this can increase adherence to [9], and enjoyment of [10], interactions with a robotic exercise coach, we went a step further. In this work, we attempted to predict the style of interaction (i.e. the behaviours used by the robot during a session) that each user would prefer based on their information and training context.

III. System Description and Implementation

A technical description of the system is beyond the scope of this paper. Its function and personalisation method is given here to aid understanding of the study detailed in Section IV.

[1] https://github.com/M4rtinR/BehaviourGraphVisualisations.

A. System Description

A Pepper robot verbally guides a user through an individual squash session and provides technical coaching on aspects of each shot performed. During the sessions, the robot receives performance data from an IMU sensor attached to the end of the player's racket, which it uses to analyse the quality of the player's swing compared to a professional player. At the start of each session, the robot gives a verbal introduction explaining that it is going to coach the user through a squash session. The sessions comprise sets of a particular shot, so the robot then asks the user to perform e.g. 30 backhand drives. Once the user has completed their 30 shots, the robot asks the user to stop, and gives feedback on the quality of the set just performed before introducing the next set. The whole process is conducted autonomously by Pepper, using actions generated from the coaching policy selected using the process in Section III B.

Robotic behaviours are formulated from the selected actions, incorporating the performance data and stage of the session. Behaviours are primarily animated utterances spoken by the robot, but also include demonstrations via the robot's movements (examples of behaviours are given in TABLE I). For example, the robot might perform a pre-instruction behaviour, praise, or ask a question while demonstrating the correct arm position for a certain technique.

TABLE I. THE 13 BEHAVIOURAL CATEGORIES THE ROBOT CAN SELECT FROM.

Action Category	Example
Pre-Instruction	"In the next set, make sure on every shot you play your racket face stays open."
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 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"

B. High-Level Personalisation

In each session, the system executes one of the 12 coaching policies identified in our previous work [6]. The choice of policy to execute is made at the start of each coaching session. It is based on the recommendations of interviewed coaches and physiotherapists as described in [6]. These interviews led to the creation of the following user information categories: the player's ability (self-rated), number of interactions with the robotic coach (i.e. length of the relationship), motivation for training (self-rated), and type of session. In the current work, each of these information categories was split into a 'high' and 'low' value, and if the value matched the recommendations given by [6], the policy received a higher score. The policy with the highest score was used by the robotic system during the coaching session. Thus, the behaviour of the robotic coach is personalised based on real-world human coaching and interview data.

IV. Procedure

A. Participants

Squash players meeting the following inclusion criteria were recruited from local squash clubs:

1. 18 years of age or older.
2. Not a member of the Professional Squash Association (PSA) World Tour (i.e. not a professional player).
3. Has held a membership to a squash club for at least one year (i.e. not a complete beginner).

Each participant was entered into a raffle for a £50 Amazon gift card. In total, 22 participants were recruited, but due to technical faults, as described in Section V, only 16 participants completed all 3 sessions. The demographics of those 16 participants can be found in TABLE II.

TABLE II. Demographic information of the 16 participants.

Gender	4 female, 12 male
Age	Mean = 25.9 ± 6 (range = 19-42)
Ability	**Improver: 3** (have played squash before but not competitively) **Club-lower: 1** (have played competitive squash at a low level) **Club-higher: 7** (frequently play competitive squash in higher club teams) **Regional: 3** (have represented their county/region or play to a similar standard) **Elite : 2** (play to a very high level but have not played professionally)
Hand	1 left-handed, 15 right-handed

B. Conditions

Three conditions were evaluated during this study. In the **Data Selected Policy (DSP)** condition, participants interacted with the robot executing the policy chosen using the method described in Section III. B. In the **Non-Personalised Policy (NPP)** condition, the robot executed a randomly selected policy from the other 11 policies that were not deemed the best-fit for the participant's user information and training context. These two conditions will be referred to from hereon as the "coaching conditions" and comparing them allowed us to discover the effect of personalisation compared to a randomly selected policy. The selection of the random policy and the best-fit policy was performed at the beginning of the first interaction. A **No Coaching Policy (NCP)** baseline condition was also used in which the robot told the user which shot and swing statistic to work on and when to perform each set of shots (see subsection E), but did not give any coaching behaviours. It was therefore the closest condition to a regular solo practice session in squash.

A within-subject design was used. Participants interacted with each condition for around 15 minutes. The interactions were split across two different days (see subsection E) and the condition order was counterbalanced across participants.

C. Hypotheses

Based on our previous observation and interview studies [6], and on the literature surrounding robotic coaching, we made 4 hypotheses regarding the evaluation of our system. We expected that using a player's information and training context for high-level personalisation by selecting the most appropriate policy would outperform a data-based policy that wasn't personalised, and guided practice without any coaching. The following hypotheses relate to the DSP, NPP and NCP conditions described in subsection B.

H1. a) The coaching conditions (DSP and NPP) will be viewed as more effective coaches than the NCP condition.

b) The DSP condition will be viewed as the most effective coach.

H2. a) Participants will be more motivated to conduct solo practice when using the coaching conditions compared to the NCP condition.

b) Participants will be the most motivated to conduct solo practice in the DSP condition.

H3. a) The coaching conditions will be viewed as more socially competent than the NCP condition.

b) The DSP condition will be viewed as the most socially competent.

H4. a) Larger technical improvements will be made during sessions in the coaching conditions than the NCP condition.

b) The largest technical improvements will be made during sessions in the DSP condition.

D. Measures

The following measures were used to gather appropriate data to evaluate our hypotheses:

Coaching Behaviour Scale for Sport (CBS-S) [11]*:* The "technical skills" subscale was used to measure participants' opinions on the coaching provided by each robot condition. The 8 questions in this subscale allowed evaluation of **H1**.

Intrinsic Motivation Inventory (IMI) [12]*:* The interest/enjoyment, perceived competence, perceived choice and value/usefulness subscales of the IMI were used to assess the effect of each condition on the participants' intrinsic motivation for conducting a solo practice session with the robot, thus allowing the evaluation of **H2**.

Robotic Social Attributes Scale (RoSAS) [13]*:* The RoSAS was used as a subjective measure of the social competence of the robotic coach to allow evaluation of **H3**.

Statistics from racket sensor: By comparing the average sensor score (Section III. A) of a baseline set to the average score of the final set performed by the participant, the improvement/deterioration of that particular metric during the session could be obtained, allowing evaluation of **H4**.

E. Study Design

The study took place on a hard-back squash court in the university's sports centre. A squash racket with a sensor attached was provided to all participants and was sanitised between sessions. This study took place in August 2021, so additional COVID safety measures were also taken, including the wearing of masks by the researcher and participant (other than when the participant was on court) and 2-metre social distancing at all times. At least a week prior to the study, all participants were sent an information sheet and consent form outlining what would be expected of them and how the data would be collected and used. The study received ethical approval from the university's ethics board.

TABLE III. shows an overview of the study procedure used. Each participant attended the facility on 2 separate days in an attempt to strike the right balance between mitigating against fatigue and mitigating against mid-study COVID restriction changes or self-isolation requirements. To further ensure fatigue did not play a role, on the second day participants were given a minimum of 10 minutes break between sessions while they completed the questionnaires.

TABLE III. SUMMARY OF THE PROCEDURE USED.

Phase	Activity	Duration
Intro	Researcher explained the study and answered any questions. Participant signed the consent form and completed the demographic questionnaire.	7 mins
Warm-up	Participant warmed themselves and the ball up.	3 mins
Squash 1	**Baseline set:** Participant played a set of 30 of the selected shot.	15 mins
	Practice: Participant undertook a solo squash practice session in one of the three conditions (order conterbalanced). Around 120 shots were played.	
	Final set: Participant played a set of 30 of the selected shot.	
Qs 1	Participant completed the post-questionnaires.	10 mins
		Day 1 total: 35 mins
Warm-up	Participant warmed themselves and the ball up.	3 mins
Squash 2	As in **Squash 1** using a different condition. Split into baseline set, practice and final set.	15 mins
Qs 2	As in **Qs 1**. Researcher set up the next condition. Participant completed the questionnaires.	10 mins
Warm-up	Participant warmed themselves and the ball up.	3 mins
Squash 3	As in **Squash 1** using the third condition. Split into baseline set, practice and final set.	15 mins
Qs 3	As in **Qs 1**.	10 mins
Wrap-up	Researcher answered any final questions the participant had.	4 mins
		Day 2 total: 60 mins

During each interaction, participants were given a different shot and swing metric to work on by the robot (either racket preparation during a forehand drive, follow through time during a backhand drive, or racket face angle during a forehand drop). The shots are all commonly played shots in squash which all participants were able to execute.

The setup for the experiment is shown in Figure 1. The researcher observed the sessions from the balcony above the court so as not to interfere. In each of the three conditions, the full session was conducted autonomously by the robot.

V. RESULTS

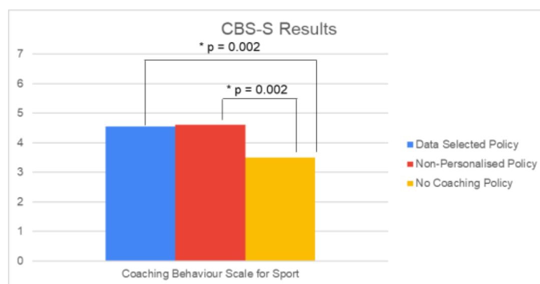
In total, 60 of the 66 squash sessions ran without problems. However, during 2 sessions, connection was lost with the sensor, and during another 4 an error caused the system to terminate prematurely. We have removed the data from the 6 affected participants from our analysis. A repeated measures ANOVA was used to test for significant differences (relating to the hypotheses given in Section IV C.) between conditions for all measures. If a significant result was found, a post-hoc pairwise comparison with a Bonferroni adjustment was then conducted to compare individual conditions.

Figure 1. The experimental setup: Pepper was on court with the player, acting autonomously, while the researcher observed from the balcony.

A. Coaching Behaviour Scale for Sport

The mean scores for the CBS-S are shown in Figure 2. A repeated measures ANOVA with a Greenhouse-Geisser correction returned a significant result ($F(1.43, 21.45) = 15.25, p < 0.001$). Post-hoc pairwise comparisons revealed that both coaching conditions (DSP: $M = 4.5$, $SD = 0.9$, NPP: $M = 4.6$, $SD = 0.9$) scored significantly higher ($p = 0.002$ for both) than the NCP condition ($M = 3.5$, $SD = 1.5$), supporting **H1 a)**. There was no significant difference between the DSP and NPP conditions.

Figure 2. Column chart showing the mean scores from the CBS-S. Significant results are indicated by an asterisk.



B. Intrinsic Motivation Inventory

Each of the 4 IMI subscales (interest/enjoyment, perceived competence, perceived choice and value/ usefulness) were analysed separately. This is standard practice (e.g. [14]). The mean scores for each can be seen in TABLE IV. A significant difference in the interest/enjoyment subscale was found ($F(2, 30) = 3.48, p = 0.044$) with post-hoc testing showing a trend ($p = 0.076$) towards the NPP condition ($M = 5.9$, $SD = 1$) being perceived as more interesting/enjoyable than the NCP condition ($M = 5.3$, $SD = 1.1$). This partially supports **H2 a)** as a coaching condition performed better than the NCP condition. No significant differences were found between the DSP condition and the other conditions, or in the scores of the other subscales.

C. Robotic Social Attributes Scale

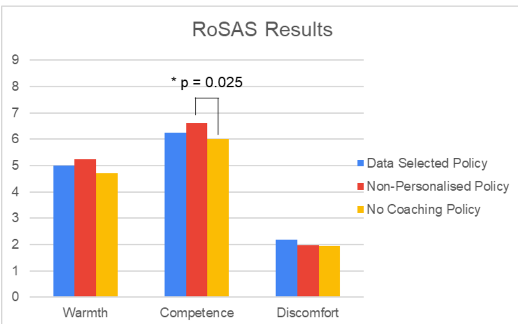
Each subscale of the RoSAS was analysed separately, with the mean values displayed in Figure 3. A significant

TABLE IV. THE MEAN (WITH STANDARD DEVIATION) SCORES GIVEN BY PARTICIPANTS TO EACH SUBSCALE OF THE IMI, WITH SIGNIFICANT RESULTS INDICATED BY AN ASTERISK *.

	Interest	Competence	Choice	Value
DSP	5.7 ± 0.6	4.6 ± 0.6	6.6 ± 0.4	5.7 ± 0.2
RSP	5.9 ± 0.6	4.8 ± 0.6	6.6 ± 0.3	5.8 ± 0.2
NCP	5.3 ± 0.5	4.7 ± 0.5	6.4 ± 0.3	5.8 ± 0.3
p	**0.044***	0.621	0.343	0.790

difference was identified in the competence subscale (F(2, 30) = 3.36, p = 0.048), with the NPP condition (M = 6.6, SD = 1.3) being viewed as significantly more competent (p = 0.025) than the NCP condition (M = 6, SD = 1.6). This result partially supports **H3 a)** and is another example of a coaching condition performing better than the NCP condition. No significant differences were found between the other conditions or in the other subscales.

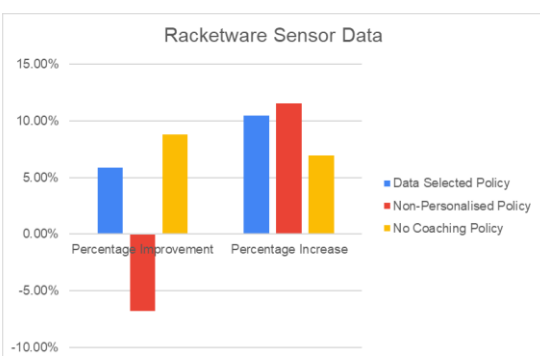
Figure 3. Column chart showing the mean scores given by participants on the 3 RoSAS subscales. Significant results are indicated by an asterisk.



D. Technical Swing Improvements

The score from the racket sensor was analysed as a percentage of the target score given by the creators of the sensor[2] for that specific swing statistic. The score the participant achieved in the baseline set was subtracted from the score achieved in the final set, giving a percentage improvement for the session. The mean percentage improvement for each condition is shown in Figure 4. No significant differences between the three conditions were found ($F(2, 30) = 3.064, p = 0.062$). The main result is close to significance but does not support **H4**.

Figure 4. Column chart showing the percentage improvement (how much closer to the target the final score was) and the percentage increase (how much higher the final score was) for each condition.



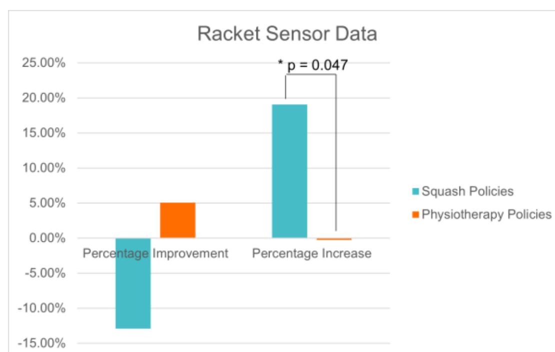
The percentage improvement represents how much closer a player got to the target score given by the system. However, players who performed well in the baseline set often got further away from the target because the robot's coaching influenced them to over exaggerate a particular aspect of their swing. This resulted in a negative percentage improvement (i.e. their score got further away from the target) despite them following the robot's instructions. The performance increase of a player indicates how much their score changed between the baseline and final set regardless of whether their new score was closer to the target. The mean percentage increase for each of the three conditions is also shown in Figure 4. A non-normal distribution meant that a log transformation was performed on the data before analysis. Again, no significant differences were found between the 3 conditions ($F(2, 14) = 0.094, p = 0.911$) so **H4** was not supported.

[2] Racketware: https://www.racketware.co.uk/

E. Domain Comparison

To learn more about the coaching policies used, for 10 participants a comparison was made between policies learned from squash coaching observation data and policies learned from stroke physiotherapy observation data (as per [6]). The other 6 participants interacted with 2 policies derived from the same domain. A Wilcoxon signed-rank test revealed no significant difference in technical improvement made by participants in either condition ($Z = -1.785, p = 0.074$) but there was a significantly bigger increase in score ($Z = -1.989$, $p = 0.047$) when participants interacted with the robot executing a squash coaching policy ($M = 19$, $SD = 19.5$) compared to a stroke physiotherapy policy ($M = -0.3$, $SD = 24.5$). This information is visualised in Figure 5 and suggests that the coaching policies may not be domain independent even though they were perceived this way by domain professionals [6].

Figure 5. Column chart showing the percentage improvement and increase of participants interacting with policies derived from squash coaching observation data compared to stroke physiotherapy observation data.



VI. DISCUSSION

Our findings support **H1 a)** and partially support **H2 a)** and **H3 a)**, providing strong evidence that the coaching conditions were more effective than the NCP condition. They performed better in perceived coaching effectiveness, interest/enjoyableness, and social competency. It is important to note that while the DSP condition did not perform the best, the NPP condition also used a policy based on human coaching data. The rigid experimental setup and inclusion criteria resulted in few differences in the user information and training context factors identified in our previous work [6]. For example, there was only one type of session possible (a technical session) and for all participants this was their first time interacting with the robot. The little high-level personalisation that was possible did not have the expected effect during these short sessions, evidenced by the lack of significant results supporting **H1 b)**, **H2 b)**, **H3 b)** and **H4 b)**.

Another factor which could have impacted the effect of our personalisation method is highlighted in [15], which used Nikolaidis' clustering method to assign a group to each of their simulated users. The simulated users were engaged for significantly longer when a robot selected actions given their group, compared to taking random actions, demonstrating the validity of this approach. However, when the robot had a flawed model of the user (i.e. it was using the wrong cluster), the users were significantly less engaged over the course of the interaction compared to randomly selecting actions. It is possible that the recommendations given by coaches and physiotherapists in our previous work [6] did not result in the robot choosing the correct coaching policy.

Despite these limitations, this is the first study to analyse technical improvements in sport using an autonomous robotic coach. The increases in racket sensor scores made by players in the DSP and NPP conditions in just 10-25 minutes, and the increase in enjoyment and perceived social competence in the NPP condition, demonstrates the potential of this system to motivate users and help them improve their playing ability.

Future work will further adapt the selected policy to individuals over time using reinforcement learning (RL). This combination of high-level personalisation with low-level adaption has been suggested as the best personalisation method for a robotic coach [4]. The negative results for high-level personalisation presented in this work adds to the evidence that personalisation is a complex and difficult problem, requiring more research. Therefore, further evaluations of the system (including low-level adaption using RL) over a longer duration and multiple sessions, as was seen to have great effect in [3], will also be conducted.

VII. Conclusion

This paper has presented the evaluation of a novel autonomous robotic coach for solo practice in a skill-based sport (squash). Interestingly, high-level personalisation of the coaching policy used by the system did not offer significant performance improvements. However, the robotic coach was perceived as a more effective coach, more enjoyable to use, and more socially competent than the baseline condition in which the robot did not offer any coaching behaviours. This indicates the potential to use a system such as this to motivate players and help them make technical improvements in sport.

Acknowledgment

This work's funder is 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] 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*, in HRI '20. Cambridge, United Kingdom: Association for Computing Machinery, Mar. 2020, pp. 520–522. doi: 10.1145/3371382.3378337.

[3] 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.

[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] 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.

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

[7] 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*, in HRI '20. Cambridge, United Kingdom: Association for Computing Machinery, Mar. 2020, pp. 272–274. doi: 10.1145/3371382.3378278.

[8] 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*, ACM, Mar. 2015, pp. 189–196. Accessed: Mar. 02, 2020. [Online]. Available: https://dl.acm.org/doi/10.1145/2696454.2696455

[9] N. Céspedes *et al.*, "A Socially Assistive Robot for Long-Term Cardiac Rehabilitation in the Real World," *Frontiers in Neurorobotics*, vol. 15, 2021, Accessed: Jul. 28, 2022. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fnbot.2021.633248

[10] J. Fasola and M. J. Matarić, "Using Socially Assistive Human-Robot Interaction to Motivate Physical Exercise for Older Adults," in *Household Service Robotics*, 2014, pp. 487–516. doi: 10.1016/B978-0-12-800881-2.00022-0.

[11] 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.

[12] 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.

[13] 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: ACM, Mar. 2017, pp. 254–262. doi: 10.1145/2909824.3020208.

[14] M. Ayoade and L. Baillie, "A Novel Knee Rehabilitation System for the Home," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*, Toronto, 2014, pp. 2521–2530. doi: 10.1145/2556288.2557353.

[15] N. Dennler, C. Yunis, J. Realmuto, T. Sanger, S. Nikolaidis, and M. Matarić, "Personalizing User Engagement Dynamics in a Non-Verbal Communication Game for Cerebral Palsy," in *2021 30th IEEE International Conference on Robot & Human Interactive Communication (RO-MAN)*, Aug. 2021, pp. 873–879. doi: 10.1109/RO-MAN50785.2021.9515466.