Personalised Socially Assistive Robot for Cardiac Rehabilitation

Critical Reflections on Long-Term Interactions in the Real World

Bahar Irfan[1,2,*] · **Nathalia Céspedes**[3,4] ·
Jonathan Casas[3,5] · **Emmanuel Senft**[1,6] · **Luisa F. Gutiérrez**[7] · **Mónica Rincon-Roncancio**[7] ·
Carlos A. Cifuentes[3,8] · **Tony Belpaeme**[9,1] ·
Marcela Múnera[3]

Abstract Lack of motivation and low adherence rates are critical concerns of long-term rehabilitation programmes, such as cardiac rehabilitation. Socially assistive robots are known to be effective in improving motivation in therapy. However, over longer durations, generic and repetitive behaviours by the robot often result in a decrease in motivation and engagement, which can be overcome by personalising the interaction, such as recognising users, addressing them with their name, and providing feedback on their progress and adherence. We carried out a real-world clinical study, lasting 2.5 years with 43 patients to evaluate the effects of using a robot and personalisation in cardiac rehabilitation. Due to dropouts and other factors, 26 patients completed the programme. The results derived from these patients suggest that robots facilitate motivation and adherence, enable prompt detection of critical conditions by clinicians, and improve the cardiovascular functioning of the patients. Personalisation is further beneficial when providing high-intensity training, eliciting and maintaining engagement (as measured through gaze and social interactions) and motivation throughout the programme. However, relying on full autonomy for personalisation in a real-world environment resulted in sensor and user recognition failures, which caused negative user perceptions and lowered the perceived utility of the robot. Nonetheless,

* Corresponding author: Bahar Irfan
E-mail: bahar.irfan@plymouth.ac.uk

[1] Centre for Robotics and Neural Systems, University of Plymouth, Plymouth, UK
[2] *Present address:* Evinoks Service Equipment Industry and Commerce Inc., Bursa, Turkey
[3] Department of Biomedical Engineering, Colombian School of Engineering Julio Garavito, Bogotá, Colombia
[4] *Present address:* Department of Computer Science and Electronic Engineering, Queen Mary University of London, London, UK
[5] *Present address:* Mechanical and Aerospace Engineering Department, Syracuse University, NY, USA
[6] *Present address:* Department of Computer Sciences, University of Wisconsin-Madison, WI, USA
[7] Instituto de Cardiología, Fundación Cardioinfantil, Bogotá, Colombia
[8] *Present address:* School of Engineering, Science and Technology, Universidad del Rosario, Bogotá, Colombia
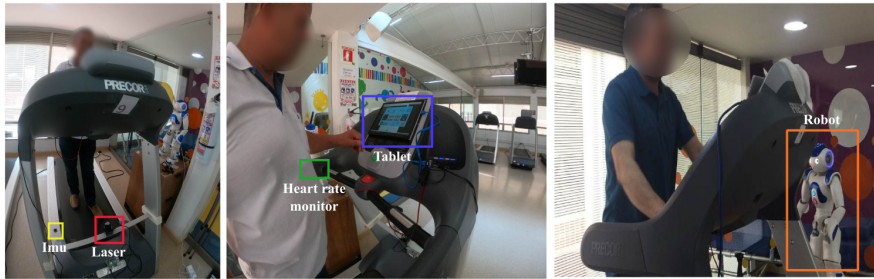[9] IDLab-imec, Ghent University, Ghent, Belgium

Fig. 1 Setup of the sensor system (image on the left), tablet interface (at the middle), and the social robot (on the right), for the cardiac rehabilitation programme at the Fundación Cardioinfantil-Instituto de Cardiología (Bogotá, Colombia).

personalisation was positively perceived, suggesting that potential drawbacks need to be weighed against the various benefits of the personalised interaction.

Keywords Personalisation in social robots · Socially assistive robotics · Real-world study · Cardiac rehabilitation · Long-term interaction · Human-robot interaction

Declaration This paper or a similar version is not currently under review by a journal or conference. This paper is void of plagiarism or self-plagiarism as defined by the Committee on Publication Ethics and Springer Guidelines.

1 Introduction

Cardiovascular diseases –disorders of the heart and blood vessels– are the most prominent cause of global deaths, leading to 17.7 million deaths each year (World Health Organization, 2011). Cardiac rehabilitation (CR) is a long-term programme following a cardiovascular episode, which aims to accelerate recovery and reduce the risk of recurrent cardiovascular events. CR is generally composed of three phases (Kraus and Keteyian, 2007): (I) the *Inpatient phase*, involving a medical procedure within the 48 hours after the cardiac event, (II) the *Outpatient phase*, typically 18 weeks (with sessions twice per week) of exercise and education programmes to improve the health of the patient, (III) the *Maintenance phase*, lasting 9 months (with one or two sessions per week) to reinforce the learned behaviour. This programme is demanding both in time and physical effort and consequently results in low adherence rates (with drop-out ranging between 15 to 50%) (Maclean and Pound, 2000; Carlson et al., 2000; Siegert and Taylor, 2004; Bethell et al., 2009; Scane et al., 2012). However, adherence to the programme is vital to achieving a complete recovery and reducing the risk of suffering recurrent events (Jolly et al., 2007; Suaya et al., 2009; Hammill et al., 2010).

Social assistive robotics (SAR) aims to provide monitoring and assistance in physical to cognitive activities, and social interaction during therapy, which improves patient motivation, task performance and clinical progress (Feil-Seifer and Matarić, 2005; Ahmad et al., 2017). However, over a long period of time, repetitive behaviours of the robot may decrease patient interest (Süssenbach et al., 2014; Kidd and Breazeal,

2008), which could reduce the frequency of use and interaction with the robot (Fernaeus et al., 2010). Personalisation, through tailoring the robot's behaviour to each patient, creates an opportunity to break this monotony and helps maintain motivation and facilitates trust over long-term interactions (Castellano et al., 2008; Leite et al., 2013; Irfan et al., 2019).

This work is concerned with the question on *what the impact is of personalisation in socially assistive robotics for long-term cardiac rehabilitation?* We designed and conducted a real-world clinical study at the Fundación Cardioinfantil-Instituto de Cardiología (Bogotá, Colombia) to evaluate the performance and perceptions of the patients throughout the outpatient phase (II) of CR (Lara et al., 2017; Casas et al., 2018). This long-term study explored three conditions: (a) conventional CR in which the patient is monitored using a suite of sensors, (b) SAR using a NAO robot (Soft-Bank Robotics Europe, Fig. 1) to continuously monitor and provide generic feedback to patients during exercise on the treadmill, based on sensory information, or (c) SAR using a NAO robot with personalised features to recognise users (Irfan et al., 2018b, 2021), recall the patients' previous session progress, track their adherence, and give personalised feedback. The results obtained during the first two conditions (Casas et al., 2019; Céspedes et al., 2021) highlighted the benefits of a social robot in comparison to the conventional CR programme in improving adherence, motivation and physical activity performance. However, the studies suggested a need for enhancing the sociability and social presence of the robot to further improve motivation and adherence. We suggest that personalisation of the interaction is key to achieving this. In this paper, we present the complete real-world study that lasted 2.5 years, drawing upon results from 26 patients that completed the programme out of 43 patients recruited, focusing on the benefits and challenges of personalisation in long-term interactions in the real world. This paper evaluates the physiological progress of the patients (i.e., training and recovery heart rate, exertion level, cervical posture), exercise intensity parameters of the sessions (i.e., speed and inclination of the treadmill), interactions with the robot, perceptions of the patients, and adherence to the programme for the personalised robot condition in comparison to the other two conditions. This is the first comprehensive study that explores the effects of personalisation in socially assistive robots for cardiac rehabilitation.

2 Background

Social assistive robotics (SAR) shares with assistive robotics (AR) not only the goal of providing assistance to patients, but also to support the user by offering social interaction, through emotional, cognitive, and social cues to encourage development, learning, or therapy (Feil-Seifer and Matarić, 2005; Okamura et al., 2010; Breazeal, 2011; Matarić and Scassellati, 2016). Because SAR aims to deploy robots in real-world therapy with users who have limited robotics expertise –such as doctors, nurses and patients– the robot needs to perform tasks with a high degree of autonomy. The robots need to provide verbal and non-verbal communication to engage in a natural interaction with the patient (Duffy et al., 1999; Feil-Seifer and Matarić, 2005; Tapus et al., 2007). Other features necessary for real-world deployment of SAR ap-

plications include automated perception of the user's behaviour, quantitative diagnosis and assessment, mobility, sensor-based automated health data acquisition and context-appropriate assistance through user interfaces (Okamura et al., 2010; Prescott and Caleb-Solly, 2017; Johanson et al., 2020). SAR based applications have been developed in a range of clinical areas (e.g., cognitive and developmental disorders, care for elderly, and rehabilitation), all of which share common goals, such as providing physical, cognitive and social support, monitoring and feedback, increasing user motivation, engagement and adherence, and improving task performance and progress (Leite et al., 2013; Ahmad et al., 2017).

However, most research in SAR has been carried out under laboratory conditions, using short-term interventions that often rely on tele-operation (Leite et al., 2013; Lane et al., 2016; Vandemeulebroucke et al., 2018). This restricts their relevance to long-term therapies in real-world applications and does not address the novelty effect (Gockley et al., 2005) or the challenges faces in the adoption of technology (Riek, 2017; Coninx et al., 2015). For instance, the only prior study with a socially assistive robot in cardiac rehabilitation (for spirometry exercises in inpatient phase I) (Kang et al., 2005) is evaluated under laboratory conditions for one session with a low number (5) of healthy participants, without analysing the physiological progress of the patients. Thus, despite the positive feedback of the participants, these results cannot be generalised to long-term rehabilitation. Nonetheless, real-world studies in other rehabilitation applications showed the benefit of robots over long-term interactions. For instance, a robot was used to support long-term post-stroke rehabilitation, which lasted 5 to 7 weeks (Feingold Polak and Tzedek, 2020). Preliminary results (based on 4 patients) indicated that SAR helps to engage, motivate and support upper limb rehabilitation. Another study in cognitive rehabilitation (Feng et al., 2020) developed a combined platform using augmented reality and an animal-like robot to engage dementia patients within multi-sensory stimulation sessions over four weeks, which was found to elicit positive emotions, increase social bonding and restore communication. Broadbent et al. (2018) evaluated the usefulness of a home-based social robot as a supporting tool in patients with the chronic obstructive pulmonary disease based on a 4-months study, which highlighted the capability of SAR to increase adherence to medication and exercise.

However, long-term interactions can cause a considerable decrease in user interest and motivation compared to the initial interaction (Kidd and Breazeal, 2008; Süssenbach et al., 2014; de Graaf et al., 2016). Thus, a number of studies focused on seeking strategies to promote long-term interactions in real-world environments. For instance, the Autom Robot (Kidd and Breazeal, 2007, 2008) was developed with the aim to address obesity and assist those who are willing to lose or maintain weight. To achieve and maintain these goals in long-term interactions, several key features were used, such as eye contact, hand, head and arms gestures, speech recognition and synthesis, and tracking user progress. The study was conducted with 45 subjects over six weeks at the participants' homes, and the Working Alliance Inventory (Horvath and Greenberg, 1989) (WAI) questionnaire was used to evaluate the interaction. The results showed that the participants assisted by the social robot used the system for longer periods than those who use other methods (i.e. tablet and paper logging of data), had a stronger alliance with the proposed system and a higher interest in know-

ing calorie consumption and exercise performed (Kidd and Breazeal, 2008), which supports the importance of embodiment in SAR. Other studies have shown that embodiment can increase compliance (Bainbridge et al., 2008), likeability (Fasola and Matarić, 2013; Li, 2015), social engagement (Lee et al., 2006; Wainer et al., 2006; Vasco et al., 2019), adherence (Bickmore and Picard, 2005a; Kidd and Breazeal, 2007) and task performance (Vasco et al., 2019; Deng et al., 2019), which are essential in, especially long-term, therapy.

Moreover, a variety of long-term studies on the adaptive and reactive seal-shaped robot PARO showed the positive impact on elder care, such as reducing negative emotions and behavioural symptoms of elderly residents, improving their social bonds and engagement, and promoting positive mood and quality of care experience (Hung et al., 2019). Adaptation and personalisation strategies (e.g., addressing the patient with their name, tracking their progress, and consequently adapting feedback or therapy tasks) were found to play an important role in long-term therapy (Rossi et al., 2017) to elicit and maintain user engagement over extended durations (Tapus et al., 2009; Blanson Henkemans et al., 2013; Scassellati et al., 2018; Winkle et al., 2018; Clabaugh et al., 2019; Richardson et al., 2018; Cao et al., 2019), improve task performance (Tapus et al., 2008; Tapus, 2009; Matarić et al., 2009; Hemminahaus and Kopp, 2017; Andriella et al., 2020), increase perceived familiarity and sociability (Sabelli et al., 2011; Fasola and Matarić, 2013), and perceived competence and trust (Schneider and Kummert, 2021). Moreover, previous studies in other domains showed the benefits of recalling user's personal attributes (e.g., name, gender, age) (Kanda et al., 2004; Gockley et al., 2005; Mutlu et al., 2006; Kanda et al., 2007, 2010; Sabelli et al., 2011; Fasola and Mataric, 2012; Belpaeme et al., 2013; Leite et al., 2014; Kennedy et al., 2015; Churamani et al., 2017; Campos et al., 2018; Zheng et al., 2019; Irfan et al., 2020b), preferences (Ho et al., 2010; Belpaeme et al., 2013; Churamani et al., 2017; Zheng et al., 2019; Irfan et al., 2020b), behaviour patterns (Glas et al., 2017; Zheng et al., 2019), and shared history (Ho et al., 2010; Belpaeme et al., 2013; Matsumoto et al., 2012; Leite et al., 2014, 2017; Campos et al., 2018; Zheng et al., 2019; Ahmad et al., 2019) for improving user experience in long-term interactions.

In our previous work (Lara et al., 2017; Casas et al., 2018), we described a SAR interface for a long-term cardiac rehabilitation programme. The interface consists of two main modules: (i) the sensor module, in charge of measuring the patients' cardiovascular and spatiotemporal gait parameters through a set of sensors, and (ii) a robot module, which consists of a social robot that monitors and provides immediate feedback and motivation to patients to increase their engagement in CR programmes. The interface was validated with laboratory (Lara et al., 2017) and clinical case studies (Casas et al., 2018; Casas et al., 2018a,b, 2019, 2020). The outcomes of these studies showed that the interface is robust and offers additional monitoring during therapy, and that using a robot improved motivation, trust and adherence in the CR programme. In addition, the clinicians' perception and attitudes towards the robot improved after the demonstration of its potential benefits (Casas et al., 2019). Subsequently, we developed additional features, such as user recognition, adherence tracking and personalised immediate and progress feedback. A case study of a patient assisted by that robot (Irfan et al., 2020a) showed that personalisation helps maintain positive perceptions and social interactions throughout the long-term programme, and

facilitates patient motivation and adherence. Finally, we analysed the overall benefits of the non-personalised socially assistive robot in comparison to conventional cardiac rehabilitation (Céspedes et al., 2021). The results showed that the patients assisted by the robot had a higher adherence, improved faster on their cardiovascular functioning, and showed better physical activity performance. Moreover, the clinicians acknowledged the motivational benefits of the robot for the patients, and the added value of continuous monitoring in cardiac rehabilitation. This work draws on these conclusions, and analyses whether personalisation can improve the benefits of the socially assistive robot, and the perception of the patients.

3 Conventional Cardiac Rehabilitation

Each exercise session in the outpatient phase (II) of CR typically lasts about an hour at the Fundación Cardioinfantil-Instituto de Cardiología clinic, and is generally conducted twice per week for 18 weeks. Each session starts by measuring patients' initial parameters, such as the *initial resting heart rate*, *weight* and *blood pressure*, followed by a *warm-up* consisting of low-paced walking and low-intensity stretching exercises in a group. Afterwards, the patients attend the *training* session based on physical exercises using a treadmill, lasting 15 to 20 minutes. The intensity of the *training* session is determined by the *speed* and *inclination* of the treadmill, which are chosen by the physiatrist at the start of each session. These parameters progressively increase throughout the CR programme depending on the patient's CR performance and progress (Simms et al., 2007). Furthermore, during *training*, the healthcare staff (i.e., physiatrists, occupational therapists, physical therapists, and nurses) manually measures the patients' heart rate (HR), and requests a self-reported exertion level using the Borg scale (Borg, 1998) every 5 to 7 minutes, as shown in Fig. 2, to adjust the intensity if necessary. After *training*, patients step off the treadmill to perform low-intensity exercises for 10 to 15 minutes, referred to as *cooldown*, in order to gradually decrease their HR. During this step, the *resting HR* and *blood pressure* are measured.

The CR sessions are conducted within a large group (e.g., 20 patients), thus, it is challenging for clinicians to provide continuous and individual monitoring of the patients (Turk-Adawi et al., 2019) due to the lack of a telemetry in the CR unit, especially during high-intensity exercises in *training* which can result in critical heart rates. Moreover, clinical studies suggest that providing individual support and supervision during exercise can improve a patient's motivation within the CR programme (Shahsavari et al., 2012). That is why, we designed a SAR system to provide continuous monitoring and feedback during *training* to facilitate prompt intervention from the healthcare staff, in addition to improving patients' motivation and adherence to the programme, as described in the next section.

4 Personalised Patient-Robot Interface

As previously described, cardiac rehabilitation is a long-term programme, where adherence is critical to ensure a complete recovery. However, repetitive and rigorous

exercises may cause a decrease in patient motivation, and consequently, dropouts from the programme. Thus, based on prior research in other areas of SAR and rehabilitation, we developed a personalised patient-robot interface to provide continuous monitoring and individualised support to patients during exercise for improving their motivation and engagement in the programme. All behaviours and responses of the robot were created in collaboration with medical specialists to avoid incorrect medical assessments and prevent negative perceptions of the programme, and with the aim to resemble the behaviours of the clinicians towards the patients. The interface, as presented in Fig. 3, consists of a sensor interface to measure the patient's performance, the socially assistive robot to provide immediate feedback and motivation based on sensor values, and personalisation features to track the patient's progress throughout the CR programme and provide a tailored experience to the patients.
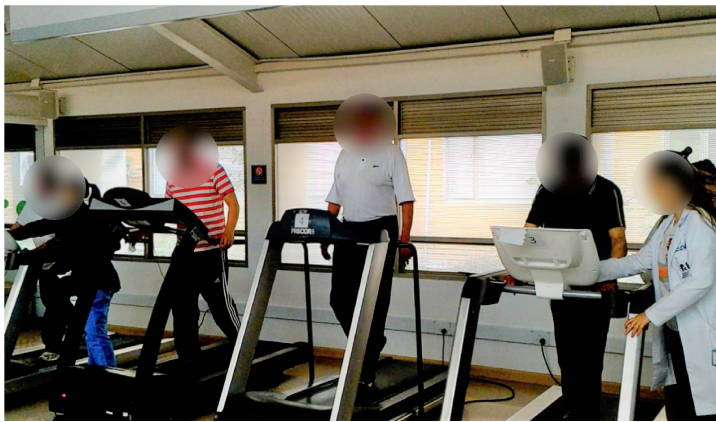

Fig. 2 Conventional cardiac rehabilitation session during *training* exercises at the Fundación Cardioinfantil-Instituto de Cardiología.
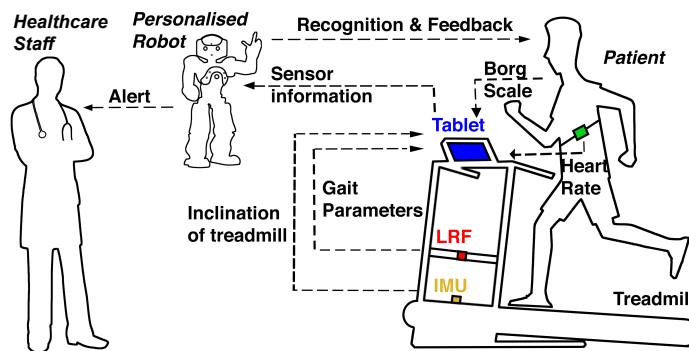

Fig. 3 The architecture of the personalised patient-robot interface for cardiac rehabilitation programme.

4.1 Sensor Interface

The patient-robot interface integrates a set of sensors to measure the patients' physiological progress regarding their cardiovascular functioning and exercise performance. *Cardiovascular parameters* (i.e., training and resting HR) of the patient were measured by a heart rate monitor (Zephyr HxM, Medtronic, New Zealand). *Gait spatiotemporal parameters* (i.e., cadence, step length, speed) were measured with a laser range finder (Hokuyo-URG 04LX-UG01, Hokuyo, Japan), shortly LRF. The *treadmill inclination* was measured using an Inertial Measurement Unit (MPU9150, Invensense, USA), shortly IMU. Finally, to visualise the data registered, receive self-reported data from the patient and control the sessions' flow, a graphical user interface (GUI) was developed. A tablet (SurfacePro, Microsoft, USA) was used to display the GUI and the tablet camera was used to measure the *cervical posture* of the patient using a head gaze detection algorithm (Lemaignan et al., 2016) and take images for user recognition during the sessions.

4.2 Socially Assistive Robot

We use a NAO (SoftBank Robotics Europe, France) robot, as shown in Fig. 1, which is the most commonly used platform for human-robot interaction (HRI) research (Lambert et al., 2020). The robot was located at the same place (in front of the patient) during the study, for the conditions that include the robot. The robot's goal is to support and monitor the patient throughout the training (Fig. 4). At the start of the interaction, the robot will greet the patient and describe the intensity of the session (*treadmill inclination* and *speed*). During the session, the robot will periodically provide a randomly selected verbal encouragement to facilitate motivation from three pre-scripted responses for the first five minutes (e.g., *"Let's start well today!"*), 13 responses until mid-session (e.g., *"Let's go! You can do it!"*), and eight responses until the cooldown period (e.g., *"Only a few minutes left!"*). Furthermore, similarly to the medical staff in conventional therapy, the robot requests the patient to report their exertion level on a Borg scale every seven minutes, using a randomly selected phrase from 10 re-
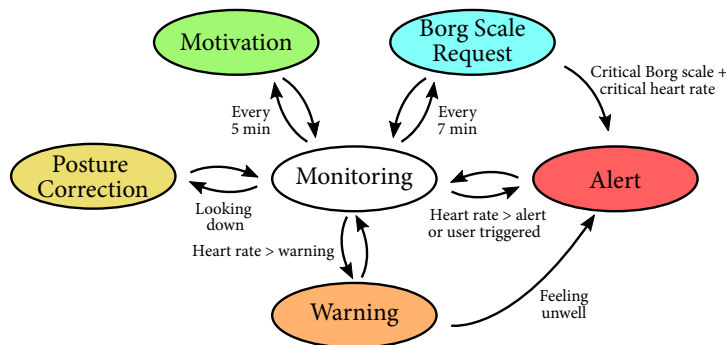

Fig. 4 Finite State Machine presenting the different behaviours of the robot during the monitoring phase.

sponses. Additionally, the robot monitors the patient to ensure that their physiological parameters remain in a healthy range. For example, if the patient's *cervical posture* is incorrect (i.e., when the patient is looking down instead of straight ahead), the robot reminds the patient to look straight ahead through a randomly selected feedback from six responses. If the patient's HR is above a warning threshold set by the therapists at the start of the session, the robot asks the patient whether they feel fine or need assistance from the medical staff. If the HR is above a critical threshold, as calculated by the medical staff using the Karvonen formula (She et al., 2014), or the patient requests it, the robot directly alerts the medical staff verbally (e.g., *"Your heart rate is too high, I am calling for help. Doctor, could you please come here?"*), waving its hand. The warning behaviours have a cooldown period of three minutes to prevent overloading the patient if a warning state is maintained. More details about the robot behaviours can be found in Casas et al. (2018); Casas et al. (2020).

4.3 Personalisation

As discussed in Section 2, several studies showed the importance of personalisation for socially assistive robotics in facilitating perceived familiarity and sociability, in addition to improving user performance and engagement within long-term interactions. Personalisation in long-term interactions should not only focus on the inherent differences, such as a person's name, it should also focus on the long-term changes, such as their therapy progress, and incrementally learn and adapt to the user for maintaining user engagement over the duration of the interaction (Tapus et al., 2007). Moreover, as discussed in the introduction, low adherence is a big concern of the CR programmes. To address these needs, we developed personalisation features focusing on users' personal attributes (i.e., appearance and name) and behaviour patterns (i.e., adherence and rehabilitation progress) by (i) recognising users autonomously, (ii) personalising the verbal (motivational and sensory) feedback of the social robot by occasionally (every 2 to 4 minutes) using their name, (iii) tracking their progress between sessions, and (iv) keeping referring to their attendance to improve their motivation in the programme and facilitate adherence. These personalisation features were developed based on the suggestions of the medical specialists, and they correlate with therapists' approaches in improving the motivation, engagement and compliance of the patient, through feedback, positive reinforcement, reminders and prompts (Winkle et al., 2018). The personalised features of the robot were validated in a pilot study under laboratory conditions before deployment in the clinic.

4.3.1 Multi-modal User Recognition

In a real-world environment such as a hospital, it is necessary to have a robust system that can autonomously recognise and learn new users that require minimum efforts from the patients and the doctors (Reig et al., 2021). Moreover, user recognition allows more natural interaction with the robot. Thus, we applied Multi-modal Incremental Bayesian Network (MMIBN) with online learning (Irfan et al., 2018b, 2021), which is the first method for sequential and incremental learning of users that does
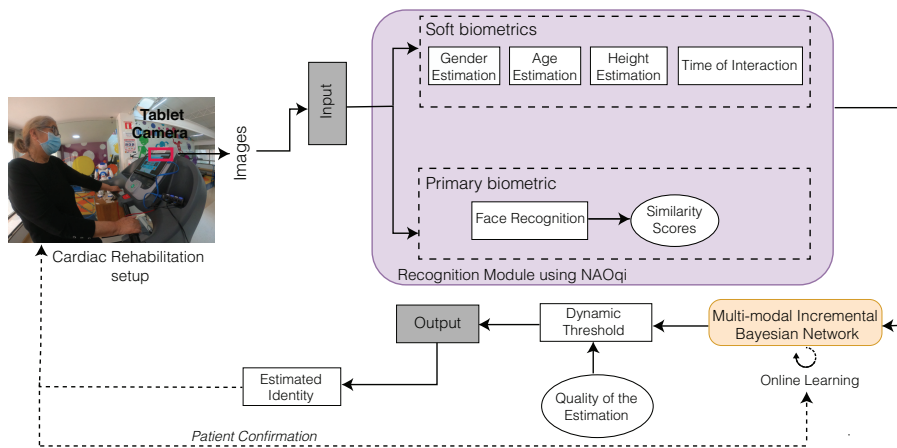
Fig. 5 Architecture for multi-modal user recognition in CR programme using the Multi-modal Incremental Bayesian Network (MMIBN) with online learning.

not require any preliminary training for user recognition. It combines face recognition with *soft biometrics*, which are ancillary physical or behavioural characteristics, such as gender, age, height and time of interaction, that can be used to improve the recognition performance (Jain et al., 2004; Dantcheva et al., 2016). The structure of user recognition can be seen in Fig. 5.

An image is taken from the tablet camera and transferred to the robot to obtain the face recognition similarity scores, and gender, age and height estimations through the NAOqi[1] proprietary software of the robot. Time of interaction is a suitable ancillary behavioural characteristic, because patients in the CR programme have set appointments twice a week on certain days and times. The biometric data are combined using MMIBN. If the quality of the estimation (i.e., the difference between the highest and the second-highest probability scores, multiplied by the number of known users) is above a certain threshold, then the identity that corresponds to the highest probability score is returned; otherwise, the user is believed to be a new patient. Explicit confirmation of the identity is obtained after each recognition (e.g., *"Hello PATIENT_NAME, it is nice to see you again! Could you confirm that it is you please?"*) through the tablet interface, in order to avoid any errors in personalisation of the session, and to improve the user recognition through online learning. Online learning of biometric data helps adapt to the changes of user appearances (e.g., different hair styles, and glasses) and the interaction patterns (e.g., time of interaction). If the user enters an identity that is not in the system for the confirmation (i.e., the user is new), ground truth values (i.e., name, age, and height) are requested from the tablet interface to apply incremental learning on MMIBN and face recognition database. Overall, the system is able to recognise and learn new users without the need for intrusive methods or external devices, such as QR codes or access cards, to offer a more natural interaction, and it is suitable for non-expert users. Moreover, MMIBN was found to significantly outperform NAOqi face recognition and a state-of-the-art

---

[1] http://doc.aldebaran.com/2-4

open world recognition method (Rudd et al., 2018) on a long-term (4 weeks) HRI study in the real world with 14 participants (93.2% identification rate) and on a large artificial multi-modal dataset with 200 users (65.7% identification rate) (Irfan et al., 2018b, 2021).

4.3.2 Progress Tracking

CR aims to improve cardiovascular functioning and recovery of the patients throughout the long-term programme. The session's intensity (as determined by the treadmill speed and inclination) progressively increases, but can be scaled back based on the patient's progress determined by a variety of physiological factors, such as the recovery and training heart rate and exertion level, and whether these parameters stay within healthy levels. Thus, it is challenging to determine the progress of the patient from session to session. Correspondingly, the medical specialists suggested session-based feedback for progress, i.e., comparing the current session to the previous session of the patient using the alerts for critical heart rate and exertion level, and the cervical posture corrections, such that the patient can track whether they are responding well to the rehabilitation on a session per session basis, which is expected to improve their motivation.

In order to prepare and motivate the patient for the session, the *relative* session intensity is indicated by the personalised robot at the beginning of each session after the announcement of session intensity parameters, such as *"Today, we are starting with a speed of 2.1 miles per hour with an inclination of 0.8 degrees, which will be more intense than the last time."*. The relative intensity is defined by the clinicians at Fundación Cardioinfantil-Instituto de Cardiología, to be higher (i.e., more intense) if the treadmill speed or inclination is higher than that of the previous session, and lower (i.e., less intense) if both of these parameters are lower. Subsequently, the previous session progress is mentioned followed by a motivational phrase, such as *"In the previous session, you experienced difficulty with your heart rate. I am sure it will be all fine this time!"* or *"I am sure that it will be as good as last time!"*. At the end of the session, the performance of the patient is compared to the previous session. To avoid demotivating the patient, the relative session intensity is also noted only if it is higher. For instance, *"We had a lower number of difficulties in this session than the previous one, even though the session intensity was higher. Let's keep up the good work, PATIENT_NAME!"*, or *"We had a higher number of difficulties this session than the previous one. Next time will be better, PATIENT_NAME!"*.

4.3.3 Adherence Tracking

Patients are prescribed two sessions per week for the outpatient phase of CR for a total of 18 weeks (4.5 months). However, the medical records in the Fundación Cardioinfantil-Instituto de Cardiología clinic[2] show that patients take 5.7 months on average to finish the outpatient phase. The long duration of the programme also decreases the willingness to continue, resulting in dropouts, as mentioned in the introduction. Hence, in order to encourage patients to come to their appointed sessions

---

[2] Obtained from 14 CR patients that are not part of the study.

and improve adherence, we tracked the patient's attendance per week. Since the lack of attendance can be either due to justifiable reasons, such as sickness or leaving town, as well as negligence, the robot comments on the missed sessions (excluding holidays) in a positive manner, such as *"You didn't come to the last (X) session(s). I hope everything is all right!"*. Moreover, to increase the sociability of the robot and familiarity, we tracked weekends and the national holidays, with comments such as *"I hope you had a nice weekend/holiday!"*.

5 Experimental Procedure

A longitudinal study was carried out at the Fundación Cardioinfantil-Instituto de Cardiología (Bogotá, Colombia) for 2.5 years to evaluate the impact of socially assistive robots and personalisation in the outpatient phase (II) of cardiac rehabilitation, which is designed to last 18 weeks (36 sessions) per patient.

5.1 Hypotheses and Predictions

In general, our study was designed to evaluate the benefits of using a socially assistive robot for long-term cardiac rehabilitation with the aim to improve user motivation and adherence to the programme. However, as previous research outlined in Section 2 shows, user motivation and interest towards a generic robot can wane over long-term interactions, which could be overcome by the personalisation of the interaction. Moreover, previous studies in SAR showed that personalisation leads to an improvement in task performance and user perceptions (e.g., competence, trust, sociability, familiarity). In addition, our previous work showed that using a social robot improves adherence, physical activity performance and cardiovascular functioning in comparison to conventional CR programme (Céspedes et al., 2021) and leads to a significant increase in the patients' perceptions of the robot (i.e., perceived trust, utility, usefulness and ease of use) (Casas et al., 2019). Correspondingly, this work aims to address the research question, *what the impact is of personalisation in socially assistive robotics for long-term cardiac rehabilitation?* and the consequent hypotheses and predictions:

H1 Personalisation will improve patient motivation and adherence to the CR programme.
  - *Prediction P_{1a}*: A higher ratio of patients will complete the CR programme with the personalised robot than the control or social robot.
  - *Prediction P_{1b}*: Patients will report higher intrinsic motivation to improve in their sessions with the personalised robot than the other conditions.

H2 Personalisation will improve the cardiovascular performance of the patients.
  - *Prediction P_{2a}*: Patients will have a higher gain of normalised recovery heart rate in the personalised condition than the other conditions.
  - *Prediction P_{2b}*: The personalised robot will lead to a lower number of alerts to medical staff than the social robot.

H3 Interaction with the personalised robot will be maintained throughout the long-term programme.
  - *Prediction P_{3a}*: Patients will comply with the robot's posture correction requests throughout the CR programme.
  - *Prediction P_{3b}*: There will be no significant decrease in the gazing behaviour to the robot throughout the programme.
  - *Prediction P_{3c}*: Social interaction between the human and the robot will be maintained throughout the programme.
  - *Prediction P_{3d}*: Patients will maintain their bond with the robot (as measured by the Working Alliance Inventory) throughout the programme.

H4 Personalisation will improve patients' perceptions of the robot.
  - *Prediction P_{4a}*: The personalised robot will be rated as more useful than the other conditions.
  - *Prediction P_{4b}*: Patients will enjoy the personalised features of the robot.

5.2 Conditions

Based on our research question, we designed three conditions for the study:

– **Control:** The patients perform conventional CR sessions (i.e., without a robot), where they are supervised by the healthcare staff. In order to compare the physiological progress between the groups, the sensor interface (as described in Section 4.1) is used to measure the patients' physiological parameters. Patients only interact with the tablet to enter their Borg scale, when it is requested through the tablet with an audible signal and a change of colour. While the health parameters (i.e., heart rate, gait speed, cadence, step length, treadmill inclination and the self-reported Borg scale value) are visible on the tablet interface for informing the medical staff, patients do not receive any motivational or physiological verbal feedback to emulate the conventional CR sessions.

– **Social Robot:** The patients perform the CR sessions assisted by the robot, as described in Section 4.2, and the sensor interface.

– **Personalised Robot:** The patients perform the CR sessions assisted by the personalised robot, as described in Section 4.3, and the sensor interface.

The *control* and *social robot* conditions started in August 2017, and the *personalised robot* condition started in October 2019. The non-personalised (social) robot, the personalised robot and the sensor interface operated fully autonomously. Nevertheless, an experimenter was present during the CR sessions for all conditions, to interfere only in the case of system failures.

5.3 Experimental Criteria

– **Inclusion Criteria:** The study targeted the patients starting the outpatient phase (II) of the cardiac rehabilitation programme, which lasts 18 weeks with two sessions per week. Patients who are over 25 years old with acute myocardial infarc-

tion, percutaneous coronary intervention, coronary artery bypass graft, valve replacement, ischemic heart disease and hypertension, and ejection fraction greater than 40% were recruited. Moreover, the participants should be able to perform treadmill exercises.

- **Exclusion Criteria:** The patient-robot interface may pose limitations on the patients with visual, auditive or cognitive impairments that may impede the manipulation and correct understanding of the system, hence, such patients were excluded from the study based on their clinical records presented upon entrance to the CR programme. The patients with a different cardiovascular pathology than the aforementioned were also not considered for the study.

- **Dropout and Incomplete Criteria:** The initial duration of the CR programme was considered to be 18 weeks, where patients would attend twice per week as appointed. However, due to some of the patients missing their sessions, this initial policy resulted in a shorter CR duration for the patients (23-33 sessions). Correspondingly, the study policy was reviewed in 2018 to last 36 sessions, in order to improve the programme offered to the patients. Thus, we define a *drop-out* from the study to be the case when the patient does not attend three sessions in a row without a justification. The patient, in that case, is dropped from the study, but may continue the CR programme without the robot or tablet interface. On the other hand, if the patient could not complete the programme due to a critical health condition, funding (e.g., health insurance coverage) or COVID-19 outbreak, their CR programme is considered *incomplete*, since these reasons are beyond their control.

5.4 Participants and Demographic Data

43 patients were recruited for the study: 15 patients for *control* and *social robot* conditions, and 13 for the *personalised robot*. However, due to dropouts and incomplete therapies due to critical health conditions, funding or the COVID-19 outbreak in March 2020, only 26 patients could complete the study. The demographic data of the patients that actively participated in the rehabilitation and completed the outpatient phase are presented in Table 1.

Of all patients present at a group-based CR session, only one patient was a participant of the study. This decision was made to prevent patients in the CR condition from meeting each other, which could potentially influence their perception of the robot. However, this places additional restrictions on the scheduling, hence, only 3 to 5 patients enrolled in the study could attend a CR session per day.

5.5 Measures

To evaluate the impact of personalisation in SAR for long-term CR, we developed the following measures based on the parameters taken during a conventional CR session and our hypotheses.

Table 1 Demographic data of the patients that actively participated and completed the outpatient phase of the CR programme within the study.

	Control	Social Robot	Personalised Robot
Participants	9	11	6
Gender	9 males	10 males, 1 female	6 males
Age (years), mean (SD)	56.6 (7.8)	55.7 (11.2)	60.3 (6.5)
Age range	44-70	43-80	56-69
Body Mass Index, mean (SD)	26.2 (2.6)	29.2 (3.9)	25.0 (2.1)
- Obese		54.5%	
- Overweight	66.7%	36.4%	50%
- Healthy weight	33.3%	9.1%	50%

5.5.1 Physiological Progress

The patient's physiological progress is assessed using the variables measured by the sensor interface.

Cardiovascular parameters: These parameters reflect the patient's cardiovascular performance during the exercise performed in cardiac rehabilitation. Primarily, two measurements are analysed : (i) Average heart rate during the training phase (*THR*), and (ii) *recovery heart rate* (*RHR*), which represents the difference between the heart rate one minute after ending the training phase of the exercise, and the *THR*. *RHR* is normalised (*RHR$_{normalised}$*) with the patient's initial resting heart rate measured at the beginning of the session (*IHR*) to allow comparison between patients. Equation 1 shows the calculation for the *RHR*.

$$RHR = THR - HR_{post-training}$$
$$RHR_{normalised} = RHR/IHR$$
(1)

Gait Spatiotemporal parameters: Measuring the gait parameters is important to track the patient's performance during the exercise. The main components of analysis during the gait can be classified in distance (spatial) measurements and time (temporal parameters). Within the CR programme, three of these variables are assessed: (i) *cadence*, which represents the total number of full cycles taken within a given period (Thompson, 2002), (ii) the *step length* that describes the distance between the point of initial contact of one foot and the initial contact of the opposite foot (Thompson, 2002), and (iii) the patient's *gait speed*, which also represents the treadmill's speed and is used to measure exercise intensity.

Cervical Posture: Because the CR sessions are performed on a treadmill, a healthy posture is essential to avoid dizziness, falls and achieve a correct gait performance (Martin and McConahay, 1972). Thus, we measure the *cervical posture* (i.e., head inclination) using the camera of the tablet located in front of the patient through a head

gaze estimator (Lemaignan et al., 2016). The output data acquired using the estimator correspond to a binary value (e.g., 'looking straight ahead', 'not looking straight ahead'), which was used for immediate feedback and was not recorded for analysis.

Exercise Intensity parameters: To measure the exercise intensity, the following parameters were acquired and measured, in addition to the treadmill speed: (i) The *treadmill inclination*, that was measured by an IMU (MPU9150) located on the treadmill's floor (the values vary between 0 and 5 degrees angle), and (ii) the patient's perceived exertion, as measured using the self-reported Borg Scale (Borg, 1998). The Borg Scale assesses in a subjective manner the exertion and intensity perceived by a patient during the exercise (Aamot et al., 2014). At Fundación Cardioinfantil-Instituto de Cardiología, the Borg scale varies between 6 and 20 (6 corresponds to a very low level of the perceived exertion, 20 corresponds to a very high level of exertion). The clinicians consider values between 6 and 13 as a safe (healthy) perceived exertion level.

Warnings and Alerts Count: As mentioned in Section 4.2, the social robot provides different types of feedback. As an additional indicator of the patient's physiological progress, *call medical staff* alerts and *high heart rate* warnings during the session were counted.

5.5.2 Long-term perception of the robot

The Unified Theory of Acceptance and the Use of Technology (UTAUT) (Venkatesh et al., 2003) questionnaire and its extension the Almere model (Heerink et al., 2010) are commonly used to evaluate key aspects of a socially assistive therapy through several concepts, such as perceived utility, trust, and adaptivity. We previously adapted UTAUT and the Almere model for a CR programme with a robot, and applied it to 8 patients that completed the *social robot* condition, and a *baseline group* of 20 patients in their early outpatient or maintenance phase, without any prior experience with robots or our system (Casas et al., 2019). The *baseline group* served as a baseline perception of the robot and our system, hence, we did not include the patients from the *control* condition (the interface only condition) to avoid biasing the results with their expectations and perceptions of the system. A debriefing was organised for the *baseline group* about SAR, its potential benefits and the parameters measured by the system, in addition to a video presentation of the *social robot* condition. The results showed that the *social robot* improved the expectations and had a significant increase in the patients' perceived trust, utility, usefulness and ease of use. However, the patients and clinicians highlighted that the robot needs to have more social skills, such as personalised feedback, reminders and physical activity updates, to enhance the interaction and improve compliance. At the time of the work, the study was not completed and the *personalised robot* condition has not yet started. In this work, we compare the perceptions of all the patients that completed the *social robot* condition, and evaluate how personalisation changes these perceptions. Moreover, we developed additional questions to evaluate the personalisation features, as shown in Table 10 in Appendix A. The *social* and *personalised robot* conditions are compared

to the *baseline group* based on the same (non-personalised) questions, whereas the personalisation questions are analysed separately.

Moreover, to evaluate whether personalisation helps build a relationship with the robot and how this is affected over the long term, we applied the Working Alliance Inventory (WAI) (Horvath and Greenberg, 1989) to the patients in the *personalised robot* condition. WAI is a 36-item self-report instrument based on Bordin's pantheoretical tripartite conceptualisation, i.e., *Bond*, *Task* and *Goal*. This questionnaire was used in long-term social robotics studies to evaluate the perceived task performance and sociability of a robot (Bickmore and Picard, 2005b; Hoffman and Breazeal, 2010; Kidd and Breazeal, 2008). The *Bond* construct measures the degree of trust and familiarity between the robot and the patient (e.g., "My relationship with the robot is very important to me"). The *Task* construct evaluates the degree to which the robot and the patient agree on therapeutic tasks (e.g., "The things that the robot is requesting from me do not make sense"). The *Goal* construct aims to measure the degree to which the robot and the patient agree on the goals of the CR programme (e.g., "The robot perceives accurately what my goals are"). WAI uses negative (e.g., "I disagree with the robot about what I ought to get out of therapy.") and positive (e.g., "The robot and I are in agreement on what is important for me to work on.") formulations to limit the bias in the results. We adapted WAI for cardiac rehabilitation to analyse the long-term perception, as presented in Table 11 in Appendix A. WAI was applied at the middle of the CR programme (18 sessions) and at the end of the programme (36 sessions).

5.5.3 Video Analysis

One of the most common measurements used in SAR is the analysis of videos (Sabanovic and Simmons, 2006; Anzalone et al., 2015; Leite et al., 2012). Initially, video recordings were not considered for the *control* or *social robot* conditions as we did not expect to observe changes that would require video analysis, and due to the lack of available resources. However, during these conditions, we observed a change of behaviour towards the robot throughout the programme in the *social robot* condition, which prompted the necessity to analyse the behaviour in detail. Correspondingly, the consent forms were modified to include video recording for the *personalised robot* condition, which started the other two conditions. The sessions were recorded with a GoPro (GoPro, Inc., USA) camera installed in the CR service.

Gaze is an important factor in human-human interaction and human-robot interaction (Ruhland et al., 2015). Most of the work in HRI has focused on generating meaningful robot gaze (Mutlu et al., 2009; Mwangi et al., 2018; Admoni and Scassellati, 2017), however, other studies also explored the importance of human gaze in HRI, how to measure it and interpret it (Broz et al., 2012; Lemaignan et al., 2016; Oertel et al., 2020). We draw from these works to interpret human gaze in our study and use it as a metric for engagement. Moreover, non-verbal emotional responses, such as gestures and facial expressions, and verbal social interactions with the robot are other methods for detecting engagement (Clave et al., 2016; Oertel et al., 2020). In addition, analysing patients compliance can help determine the effectiveness of the socially assistive robot in achieving good task performance and engagement in long-

term therapy (Matarić et al., 2007; McColl and Nejat, 2013; Fasola and Matarić, 2013). Correspondingly, the video analysis was made by two independent coders based on the following interactions: (i) *gaze* of the patient to the robot, (ii) *social interaction* of the patient's verbal and non-verbal (e.g., positive or negative expressions, gestures) responses to the robot, and talking about the robot to other patients, (iii) *medical staff interaction* with the robot, such as the doctor touching the head of the robot to suppress the *call medical staff* alert, or interacting with it verbally and non-verbally beyond the requirements of the task, and (iv) patient compliance to the cervical posture request. Prior to the analysis, a coding session was performed with the coders to unify the measurement method. All the variables were coded as binary (e.g., 1: gaze triggered or 0: no gaze behaviour) and comments were added to specify the nature of the event (e.g., when a social interaction was triggered, the event was coded, and a comment was used to describe the situation). We set a 11.8% overlap (randomly selected 24 sessions corresponding to 8.95 hours) for the coding data to verify the validation of the coding results, which is sufficient to establish inter-rater reliability (O'Connor and Joffe, 2020).

5.6 Statistical Analysis

The patient's performance in a session is affected by the exercise intensity, as well as external factors, such as illness and tiredness prior to the session. Thus, to decrease the intrasubject variability, the data is analysed within six *stages* (i.e., 6 sessions per stage), as suggested by the medical staff at Fundación Cardioinfantil-Instituto de Cardiología.

5.6.1 Numerical Data

Our study is a two-way mixed design, that is, it contains repeated measures for different groups. However, the data for physiological progress, cervical posture corrections and the interactions with the robot are not normally distributed (Shapiro-Wilk test gives $p < .001$ on residuals, and the visual inspection of the residuals show a large diversion from linear reference lines). Moreover, the homogeneity of variances assumption of ANOVA is violated ($p < .05$ in Levene's test and Box's M-test) in almost all the cases, except for exertion levels and social interactions. Hence, ANOVA cannot be applied. The non-parametric test that corresponds to a repeated-measures ANOVA is a Friedman test, however, it requires a complete block design, whereas, the group sizes are not equal (unbalanced data) due to the dropouts and the incomplete CR programmes. In addition, due to the change in the experimental criteria, some of the patients completed the CR programme earlier, and the sensor failures within some of the sessions caused incomplete (missing) data. Thus, we apply Johansen's (Johansen, 1980) general formulation of Welch (Welch, 1938)-James (James, 1951)'s statistic with Approximate Degrees of Freedom (Villacorta, 2017; Welch, 1951; Keselman et al., 2003), which is suitable for applying to repeated measures and two-way mixed designs. We evaluate the differences between the stages and the conditions using pairwise tests, with Hochberg correction for multiple comparisons and Least-Squares

Estimators (i.e., without trimming), which are default parameters of the implementation (Villacorta, 2017)[3]. In order to evaluate the consistency and the magnitude of a particular phenomenon across different studies in the literature, effect sizes are reported for pairwise tests based on Glass's delta (Glass et al., 1981; Keselman et al., 2003; Villacorta, 2017). This measure does not classify effect sizes, such as 'small', 'medium' and 'large', in contrast to Cohen (1988), because the practical importance of an effect depends on the context of the applications, such as relative costs and benefits, and a small effect size can make a substantial difference. A negative effect size denotes a decrease in the mean between group 1 (e.g., *social robot*) and group 2 (e.g., *personalised robot*), whereas, a positive effect size denotes an increase. Inter-rater reliability agreement on the video data is measured by Cohen's kappa (κ) (Cohen, 1960) and interpreted according to McHugh (2012). Furthermore, the McNemar test, which enables comparing two classification algorithms that are run only once (Dietterich, 1998), is applied to compare MMIBN to NAOqi face recognition. A more detailed analysis of user recognition in comparison to a state-of-the-art open world recognition algorithm (Rudd et al., 2018) is available in (Irfan et al., 2021).

5.6.2 Ordinal Data

Likert scales are ordinal, hence non-parametric tests should be applied to analyse the questionnaires (Jamieson, 2004). Correspondingly, Wilcoxon signed-rank test is applied on WAI results with Bonferroni correction, because the same test is applied to the patients twice (i.e., at the middle of their CR programme and at the end). Mann-Whitney U-test is applied for UTAUT for analysing the significant differences between the conditions (i.e., independent samples).

6 Results

As previously stated, this work focuses on the impact of the personalisation of the robot for the CR programme. Correspondingly, the results are analysed in that perspective, comparing the effects of the *personalised robot* to that of the *social robot*, as well as to the conventional CR programme, through various measures described earlier.

6.1 Adherence

While 43 patients participated in the study within the 2.5 year study duration, corresponding to 1050 sessions, due to the reasons beyond the control of the patients –such as funding, medical condition, and the outbreak of COVID-19– 8 patients could not complete the CR programme, and an additional 9 patients dropped out of the programme. The *control* condition was completed in January 2019, however, the *social* and *personalised robot* conditions were halted due to the COVID-19 pandemic in March 2020, hence 6 patients in the *personalised robot* condition and one patient
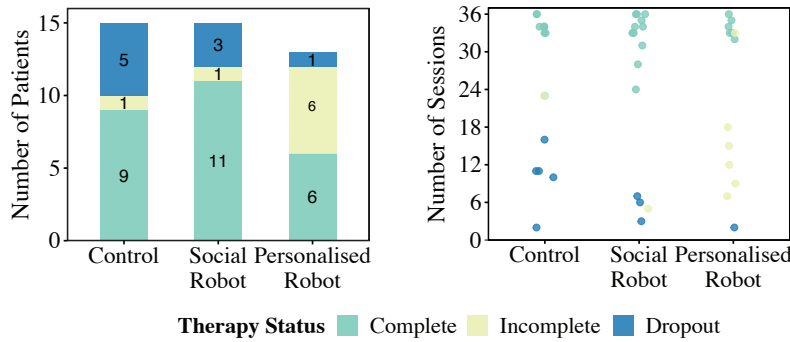
---

[3] `https://cran.r-project.org/web/packages/welchADF/index.html`

Fig. 6 The CR programme status of the users in the *control*, *social robot* and *personalised robot* conditions: 'complete' refers to the completed cardiac rehabilitation programme as determined by the clinicians; 'incomplete' is when patients need to stop the programme due to reasons beyond their control (e.g., funding, medical condition, the outbreak of COVID-19), and 'dropout' refers to not attending 3 sessions in a row without a justification.

in the *social robot* condition could not complete the programme. Thus, we do not have conclusive evidence on adherence to validate our initial prediction (P_{1a}) for the *personalised robot* condition. However, the attended sessions per condition in Fig. 6 shows that the dropouts occur at earlier stages of the CR programme mostly in the *control* condition, which could indicate a higher tendency to continue the programme with the presence of a (social or personalised) robot.

While the intended duration of the CR programme is 18 weeks (4.5 months) with sessions twice per week, patients who attend the conventional CR sessions take on average 5.7 months to finish the outpatient phase of the programme, as previously highlighted in Section 4.3.3. This duration is decreased in the *control* condition to 4.7 months, which could be due to being part of a study. Nonetheless, both the patients assisted by the *social robot* and the patients assisted by the *personalised robot* finished their CR programme earlier within 4.6 months on average. Although the difference between conditions is small, the findings suggest that SAR could encourage patients to attend more actively to the sessions. Patients performing in the *social robot* and *personalised robot* conditions were closer to the intended duration, which have multiple benefits, not only in their cardiovascular response, but also for their rehabilitation process, such as reducing the risk of a new cardiovascular event, and faster initiation of the *maintenance* phase (III) of the CR programme to acquire more independence, and reinforce the results obtained during the *outpatient* phase (II).

6.2 Physiological Progress

As mentioned in the previous sections, the *recovery heart rate* (*RHR*) and the *training heart rate* (*THR*) are the most important physiological parameters of the CR programme that determine a patient's health progress. The increase in *RHR* signals an improvement of the patient's cardiovascular functioning and healthy recovery. As previously described in Section 5.5.1, we analyse the normalised *RHR* to
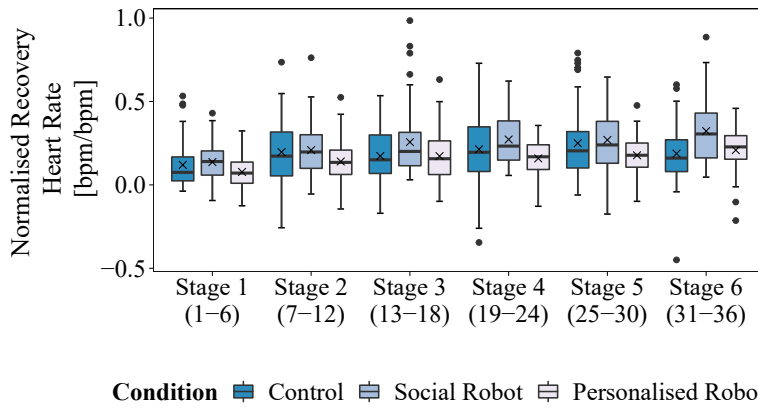
Fig. 7 Normalised recovery heart rate ($RHR_{normalised}$) throughout the cardiac rehabilitation programme for *control*, *social robot* and *personalised robot* conditions. Recovery heart rate is normalised with the initial resting heart rate on each session. The mean $RHR_{normalised}$ per stage is marked with X.

reduce the subjectivity of the measurements that change between the patients and increase the homogeneity. Fig. 7 shows how this parameter changes throughout the CR programme for the patients in all conditions. As was expected, the normalised *RHR* is significantly different between the stages ($T_{WJ}(5, 305) = 14.36, p < .001$). Table 2 shows the percentage of increments of the normalised *recovery heart rate* and significance analysis between the initial stage and subsequent stages. The results show that the increments are generally greater and more rapid in the *social robot* condition in comparison to other conditions, in contrast with our prediction P_{2a}. Nevertheless, the patients assisted by the *personalised robot* present significant differences starting from stage 5, demonstrating an improvement of the *RHR* and a successfully CR. The overall comparison between the conditions also exhibit significant differences ($T_{WJ}(2, 359) = 19.62, p < .001$). In accordance with this result, the pairwise comparison between conditions present differences, (*control-social robot*: $p = .001, \delta = 1.83$, *control-personalised robot*: $p = .02, \delta = -1.29$, and *so-*

Table 2 Welch-James ADF results (p-value and effect size in parentheses) of comparisons between stages per condition for normalised *recovery heart rate* ($RHR_{normalised}$). Significant differences ($p < .05$) are highlighted in bold. Consecutive stages and other pairwise comparisons do not exhibit any significant differences.

	Control	Increment %	Social Robot	Increment %	Personalised Robot	Increment %
Stage 1/ Stage 2	.27 *(0.47)*	49.35	.052 *(0.53)*	54.89	.51 *(0.51)*	71.08
Stage 1/ Stage 3	.72 *(0.37)*	34.24	**.002 *(0.75)***	82.43	.10 *(0.72)*	64.83
Stage 1/ Stage 4	.11 *(0.55)*	48.55	**p < .001 *(0.99)***	98.96	.10 *(0.73)*	55.01
Stage 1/ Stage 5	**.01 *(0.75)***	92.44	**p < .001 *(0.90)***	96.93	**.02 *(0.84)***	74.74
Stage 1/ Stage 6	.85 *(0.40)*	68.24	**p < .001 *(1.13)***	135.82	**.005 *(1.02)***	109.98

Table 3 Welch-James ADF results (p-value and effect size in parentheses) of comparisons between conditions per stage for normalised *recovery heart rate* ($RHR_{normalised}$). Significant differences ($p < .05$) are highlighted in bold. Consecutive stages and other pairwise comparisons do not exhibit any significant differences.

	Control - Social Robot	Control - Personalised Robot	Social Robot - Personalised Robot
Stage 1	.48 *(0.13)*	.25 *(-0.36)*	.07 *(-0.54)*
Stage 2	.72 *(0.06)*	.24 *(-0.34)*	.09 *(-0.46)*
Stage 3	.05 *(0.46)*	.99 *(-0.003)*	.07 *(-0.46)*
Stage 4	.14 *(0.32)*	.14 *(-0.33)*	**p < .001 *(-0.81)***
Stage 5	.63 *(0.10)*	.13 *(-0.42)*	**.02 *(-0.59)***
Stage 6	**.03 *(0.64)***	.66 *(0.12)*	**.03 *(-0.63)***

cial-personalised robot: $p < .001, \delta = -3.47$), elucidating an effect of the robot over the conventional CR sessions and the personalisation features (when comparing the robot-assisted sessions). As mentioned before, the patients of the *social robot* condition presented a greater increment, however, as it can be seen in Fig. 7, the distribution is more symmetric (as depicted with a median that is in the centre of the distribution) for the *personalised robot* condition, which is a positive finding signifying that the patients tend to maintain a pattern in their *RHR*. In contrast, the comparison between conditions per stage (Table 3) shows that most of the differences occurred between the *social robot* and *personalised robot*, in particular for stages 4, 5 and 6. The corresponding reason could be due to the high-intensity training applied to the patients in the *personalised robot* condition, as explained below, which might had an adverse effect on the *RHR*.

Fig. 8a shows the progress of the *THR* throughout the programme for all conditions. The comparison between the conditions does not present differences for the *THR*, neither in an overall approach ($T_{WJ}(2,402) = 1.05, p = .35$), nor for the comparison of conditions per stage ($T_{WJ}(10,275) = 1.03, p = .42$). On the other hand, the comparison between the stages shows significant differences ($T_{WJ}(5,290) = 9.2, p < .001$), in correlation with the expected behaviour during the cardiovascular rehabilitation programme. The subsequent stages present significant differences from the initial stage (Table 4) for the conditions assisted by a robot, showing that the patients assisted by the *social* and *personalised robot* improved their cardiovascular functioning. In the case of the *personalised robot*, most of the stages present a greater increment than the *social robot* condition. This outcome can be due to the physical activity intensity determined by the *treadmill speed* (as measured by the gait speed of the patient) and inclination, which was higher for the *personalised robot* (Fig. 8b and c).

For the *gait speed*, the statistical analysis shows that there are differences between the stages ($T_{WJ}(5,330) = 17.47, p < .001$) corresponding to the expected increase in the *treadmill speed* within the CR programme. However, the analysis performed for stages within each condition (Table 5) shows that the *gait speed* for the *personalised robot* only presents differences between stage 1 and 4 and stage 1 and 6, whereas the differences are significant starting from stage 2 in the other conditions. This result can suggest that the *treadmill speed* for the *personalised robot* condition were
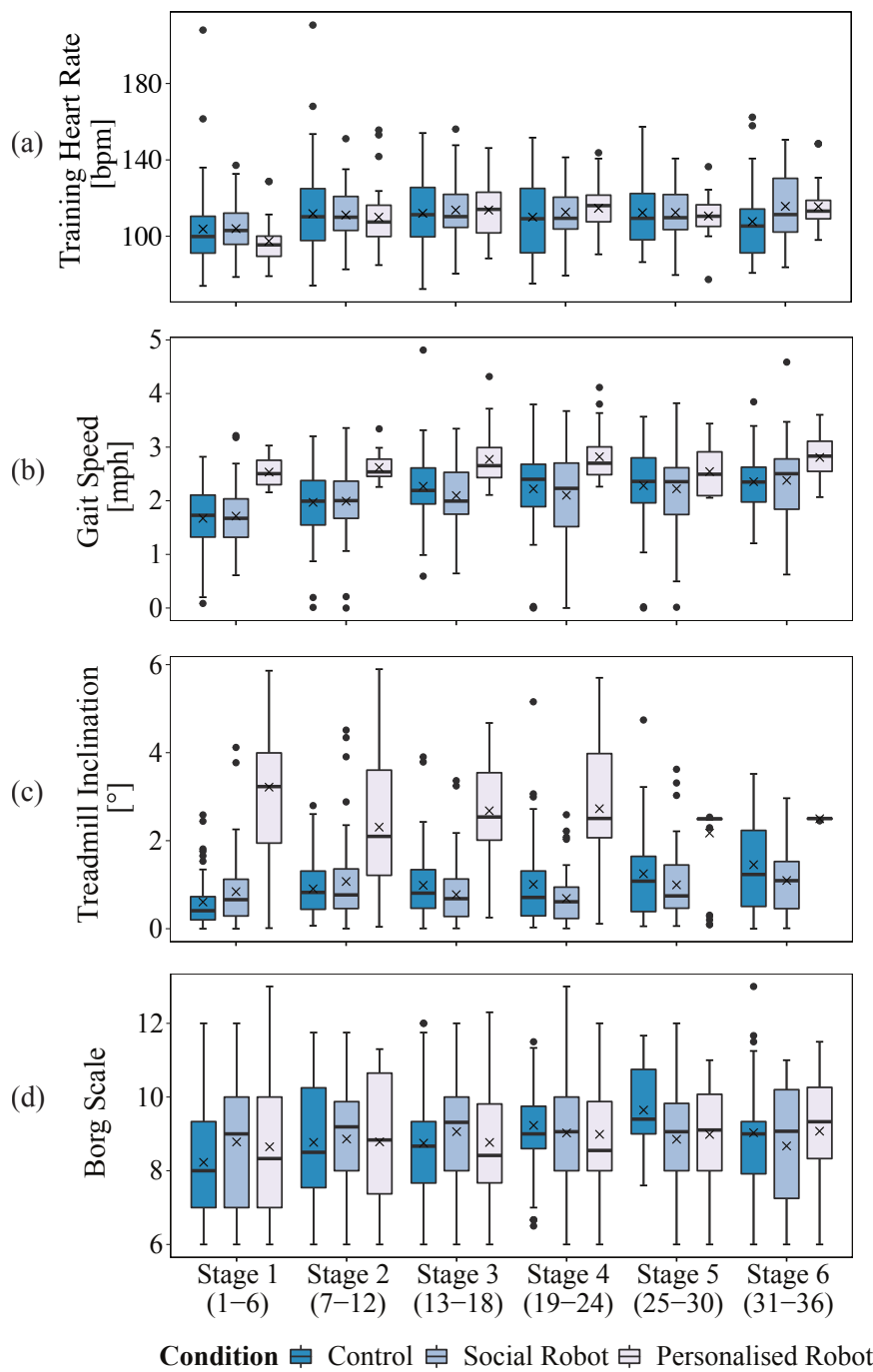
Fig. 8 Physiological parameters of the patients in all conditions: (a) *training heart rate* (*THR*), (b) *gait speed*, (c) *treadmill inclination*, (d) exertion level of the patients based on *Borg Scale* (range 6-20). The mean value per stage is marked with X.

Table 4 Welch-James ADF results (p-value and effect size in parentheses) of comparisons between stages per condition for *training heart rate* (*THR*). Significant differences ($p < .05$) are highlighted in bold. Consecutive stages and other pairwise comparisons do not exhibit any significant differences.

	Control	Increment %	Social Robot	Increment %	Personalised Robot	Increment %
Stage 1/ Stage 2	.97 (0.34)	6.10	**.03 (0.55)**	8.72	**.008 (0.89)**	7.78
Stage 1/ Stage 3	.84 (0.39)	5.88	**.002 (0.70)**	11.47	**p < .001 (1.28)**	11.90
Stage 1/ Stage 4	.97 (0.29)	4.04	**.006 (0.64)**	10.47	**p < .001 (1.47)**	12.06
Stage 1/ Stage 5	.69 (0.42)	6.94	**.01 (0.63)**	10.71	**p < .001 (1.23)**	8.50
Stage 1/ Stage 6	.97 (0.18)	4.50	**.01 (0.77)**	13.21	**p < .001 (1.48)**	15.43

more homogeneous across time with slight differences in increments. The comparison between the conditions show significant differences, for the overall perspective ($T_{WJ}(2, 439) = 110.24, p < .001$) and considering the stages ($T_{WJ}(10, 298) = 2.39, p = .01$). As it can be seen in the Fig. 8b, the *gait speed* was consistently higher for the *personalised robot* condition, which is confirmed with the significant differences found between the conditions (*control* and *personalised robot*: $T_{WJ}(1, 386) = 125.87, p < .001, \delta = 6.33$, *social robot* and *personalised robot*: $T_{WJ}(1, 415) = 163.36, p < .001, \delta = 6.75$). Similarly, the comparison between conditions per stage (Table 6) shows that there are significant differences between the *personalised robot* with the *control* and *social robot* conditions.

In the case of the *treadmill inclination*, there are no significant differences between the stages ($T_{WJ}(5, 166) = 1.20, p = .31$). We observed that during the study, that the healthcare staff do not drastically change the inclination to minimise patient exertion and assure a safe rehabilitation programme. However, the outcomes of the Welch-James ADF test shows that there are significant differences between the conditions overall ($T_{WJ}(2, 280) = 156.51, p < .001$) and when considering the 'interaction' between conditions and stages ($T_{WJ}(10, 224) = 3.93, p < .001$). This difference is due

Table 5 Welch-James ADF results (p-value and effect size in parentheses) of comparisons between stages per condition for *gait speed*. Significant differences ($p < .05$) are highlighted in bold. Consecutive stages and other pairwise comparisons do not exhibit any significant differences.

	Control	Increment %	Social Robot	Increment %	Personalised Robot	Increment %
Stage 1/ Stage 2	.17 (0.48)	22.66	.12 (0.46)	16.18	.91 (0.36)	3.20
Stage 1/ Stage 3	**p < .001 (0.96)**	44.06	**.003 (0.66)**	22.64	.15 (0.65)	6.66
Stage 1/ Stage 4	**.002 (0.79)**	41.66	**.02 (0.57)**	22.98	**.04 (0.82)**	8.30
Stage 1/ Stage 5	**p < .001 (0.89)**	42.56	**.001 (0.78)**	24.36	.91 (0.03)	3.23
Stage 1/ Stage 6	**p < .001 (1.20)**	45.58	**p < .001 (1.04)**	41.49	**.04 (0.83)**	12.00

Table 6 Welch-James ADF results (p-value and effect size in parentheses) of comparisons between conditions per stage for *gait speed*. Significant differences ($p < .05$) are highlighted in bold. Consecutive stages and other pairwise comparisons do not exhibit any significant differences.

	Control - Social Robot	Control - Personalised Robot	Social Robot - Personalised Robot
Stage 1	.69 (0.08)	**p < .001 (2.09)**	**p < .001 (2.07)**
Stage 2	.86 (0.03)	**p < .001 (1.45)**	**p < .001 (1.39)**
Stage 3	.15 (-0.27)	**p < .001 (0.88)**	**p < .001(1.24)**
Stage 4	.40 (-0.15)	**p < .001 (0.95)**	**p < .001 (1.13)**
Stage 5	.71 (-0.08)	.12 (0.43)	**.03 (0.55)**
Stage 6	.87 (0.04)	**.002 (0.94)**	**.006 (0.75)**

to the higher *treadmill inclination* applied in the *personalised robot* condition: *control* and *personalised robot* ($T_{WJ}(1, 220) = 232.87, p < .001, \delta = 9.67$), *social* and *personalised robot* ($T_{WJ}(1, 192) = 303.04, p < .001, \delta = 10.89$), *control* and *social robot* ($T_{WJ}(1, 347) = 3.07, p = .08, \delta = -0.90$). These differences can be also seen in Fig. 8c and Table 7 (the analysis between stages), where the inclination in the *personalised robot* group is significantly different from the other conditions.

The differences in the *gait speed* and *treadmill inclination* indicate that high-intensity training is applied for the *personalised robot* condition. This type of training did not have a negative effect on the *training heart rate* because the medical team could intervene when the value reached a critical level based on the robot's alerts. High-intensity training might have resulted due to the following reasons: (i) clinicians trusted the continuous monitoring and capabilities of the robot more over time, thus applying it in the *personalised robot* condition, which started two years after the *social robot* condition, (ii) clinicians relied on the progress feedback of the *personalised robot* to adjust the session intensity, or (iii) the patients in the *personalised robot* condition had a better cardiovascular functioning initially, hence further study is necessary to confirm the findings and reveal the underlying reasons, which could shed a more clear light on how personalisation affects cardiovascular performance and explain why P_{2a} was not validated in this study. Nonetheless, the interviews with the clinicians (Céspedes et al., 2021) provide support for trust in continuous monitor-

Table 7 Welch-James ADF results (p-value and effect size in parentheses) of comparisons between conditions per stage for *treadmill inclination*. Significant differences ($p < .05$) are highlighted in bold. Consecutive stages and other pairwise comparisons do not exhibit any significant differences.

	Control - Social Robot	Control - Personalised Robot	Social Robot - Personalised Robot
Stage 1	.09 (0.33)	**p < .001 (2.43)**	**p < .001 (2.01)**
Stage 2	.28 (0.21)	**p < .001 (1.32)**	**p < .001 (1.00)**
Stage 3	.14 (-0.28)	**p < .001 (1.72)**	**p < .001 (2.06)**
Stage 4	**.049 (-0.41)**	**p < .001 (1.38)**	**p < .001 (2.00)**
Stage 5	.17 (-0.28)	**p < .001 (1.03)**	**p < .001 (1.45)**
Stage 6	.13 (-0.38)	**p < .001 (1.85)**	**p < .001 (3.45)**

ing, and correlates with their initial perspectives on applying high-intensity training based on this feature (Casas et al., 2019).

There are significant differences for the *Borg scale* between the stages for all conditions ($T_{WJ}(5, 255) = 2.34, p = .04$), which indicate that the exertion level changes throughout the rehabilitation due to the physical activity intensity. However, Fig. 8d shows that the sessions mostly remained within the safe exertion level (between 6 and 13) of the patients. On the other hand, the comparison between the conditions ($T_{WJ}(2, 377) = 0.18, p = .84$) and the 'interaction' of the stage with the conditions ($T_{WJ}(10, 269) = 1.67, p = .09$) did not exhibit significant differences. Nonetheless, a significant difference was found in stage 5, between *control* and *social robot* conditions ($p = .01, \delta = -0.59$) and *control* and *personalised robot* ($p = .04, \delta = -0.53$). Furthermore, there are only significant differences for the *control* condition between the initial stage and stage 4 ($p = .004, \delta = 0.73$), and stage 5 ($p < .001, \delta = 1.12$), indicating that the *control* group perceived a higher level of exertion in these stages. While these differences occurred only in the *control* group, the patients assisted by the robot in both scenarios maintained a healthy exertion level throughout the CR programme, despite having a higher intensity training in the *personalised robot* group.

The warning and critical heart rates detected by the robot shed a different light on the story (Fig. 9). For instance, the *high HR warning* significantly differs between stages ($T_{WJ}(5, 227) = 4.79, p < .001$) due to the physical intensity changes. However, only the *personalised robot* condition presents significant differences between the initial stage and the subsequent stages (except for stage 5). Comparison between the robot conditions presents differences for the overall analysis ($T_{WJ}(1, 408) = 5.53, p = .02$). However, there is only a significant difference in stage 1 ($p = .003, \delta = -0.71$). Similar to the *high HR warning*, the *call medical staff* alert significantly differs between the stages ($T_{WJ}(5, 119) = 3.38, p = .007$), supporting that the heart rate changes across the CR programme due to the physical activity intensity, and may reach critical values. Our previous work showed that these critical alerts could be crucial in promptly detecting any complications and facilitate fast intervention by the
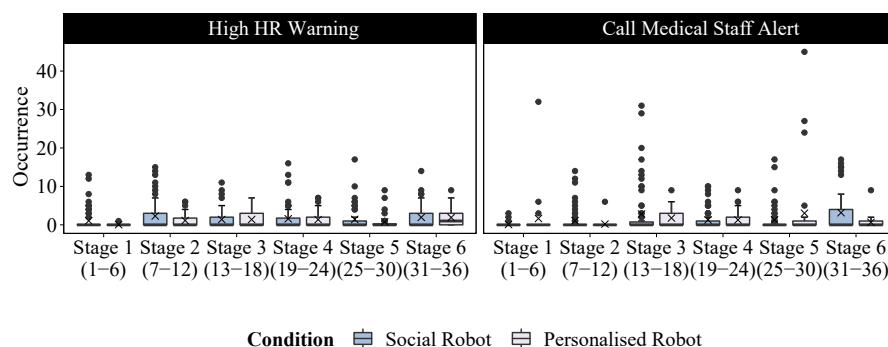
Fig. 9 Number of high heart rate warnings and critical heart rate (*call medical staff*) alerts of the patients throughout the CR programme. The results show that in contrast to the low perceived exertion levels (Borg scale), warning and critical heart rate values may arise in the sessions throughout the programme. The mean value per stage is marked with X.

medical staff for life-saving measures (Irfan et al., 2020a). In particular, the analysis between the stages based on the condition shows significant differences mostly in the *social robot* (between the initial stage and stages 3, 4 and 6) than the *personalised robot*, where the significant differences were only observed between stages 2 and 3. There were no significant differences between the conditions on the number of alerts received ($T_{WJ}(1, 132) = 0.20, p = .66$), however, in stage 2 ($p = .03, \delta = -0.44$) and stage 6 ($p = .01, \delta = -0.70$) the patients in the *personalised robot* condition had a lower number of alerts, as can be observed in Fig. 9. The lower number of alerts despite the higher intensity training could indicate a positive effect of personalisation on increasing patients' cardiovascular functioning, thus, validating P_{2b}.

6.3 Interaction with the Robot

As previously highlighted, a correct *cervical posture* is important for the safety of the patient during exercise on a treadmill, to prevent dizziness and falls. The comparison between stages shows significant differences ($T_{WJ}(5, 209) = 3.16, p = .009$), correlating with the expected posture behaviour of the patients depending on the physical activity intensity. On the other hand, no significant differences were found between the stages per condition, in addition to a lack of interaction between stages and condition ($T_{WJ}(5, 209) = 1.16, p = .33$). However, significant differences were found between the *social robot* and *personalised robot* conditions ($T_{WJ}(1, 421) = 58.24, p < .001$), as detailed in Table 8. Fig. 10 shows that the posture corrections were lower for the patients assisted by the *personalised robot*. The underlying reason could be the progress feedback given by the *personalised robot* to the user in an individualised manner. For instance, two events are highly relevant to influence a patient's *cervical posture*: (i) the feedback provided at the end of the sessions could positively affect the patient's intrinsic motivation to improve their next sessions (which would validate P_{1b}), and (ii) including the patient's name as part of the feedback could improve the patient's perception and reaction over this type of feedback. However, considering that the number of posture correction requests was lower for the patients in the *personalised robot* condition starting from the first stage, the result may also be due to the differing characteristics of the patients, hence, further study with a larger population is necessary to confirm the effects.

Table 8 Welch-James ADF results (p-value and effect size in parentheses) for *cervical posture* correction in the *social robot* and *personalised robot* conditions, and the patients' compliance in the *personalised robot* condition. Significant differences ($p < .05$) are highlighted in bold. Consecutive stages and other pairwise comparisons do not exhibit any significant differences.

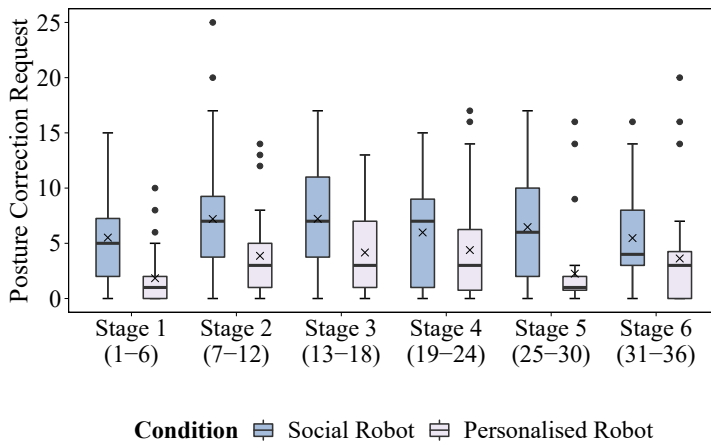|	Social Robot - Personalised Robot
Stage 1	**$p < .001$(-1.04)**
Stage 2	**.001 (-0.71)**
Stage 3	**.001 (-0.70)**
Stage 4	.10 (-0.35)
Stage 5	**$p < .001$ (-0.95)**
Stage 6	.10 (-0.41)

Fig. 10 The number of cervical posture correction requests by the *social robot* and the *personalised robot* conditions. The results show that the corrections were significantly less in the *personalised robot* condition. The mean value per stage is marked with X.

The patients in the *personalised robot* condition complied well with the requests of the robot, i.e., no significant differences were observed between requests and patient posture corrections ($T_{WJ}(1, 107) = 2.81, p = .10$), as visible in Fig. 11, with a moderate agreement between raters of the video analysis (Cohen's $\kappa = 0.665, z = 8.87, p < .001$). While this compliance does not significantly differ with stage in general ($T_{WJ}(5, 71) = 0.87, p = .51$), the in-depth analysis showed that the patients' corrections significantly differed from the robot requests in the last stage of the CR programme ($p = .03, \delta = -0.77$), which indicate that the patients mostly maintained their compliance throughout the long-term rehabilitation, validating P_{3a}, only to de-
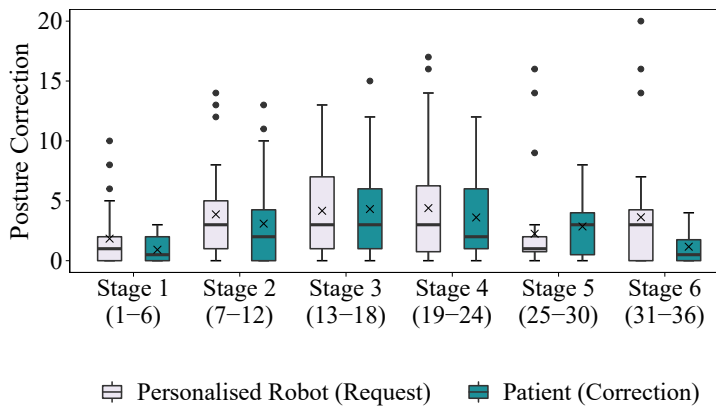

Fig. 11 The number of cervical posture correction requests by the *personalised robot* and the corresponding patients' compliance to the requests. The results show that the patients complied well to the *personalised robot* throughout the CR programme, except on the final stage. The mean value per stage is marked with X.
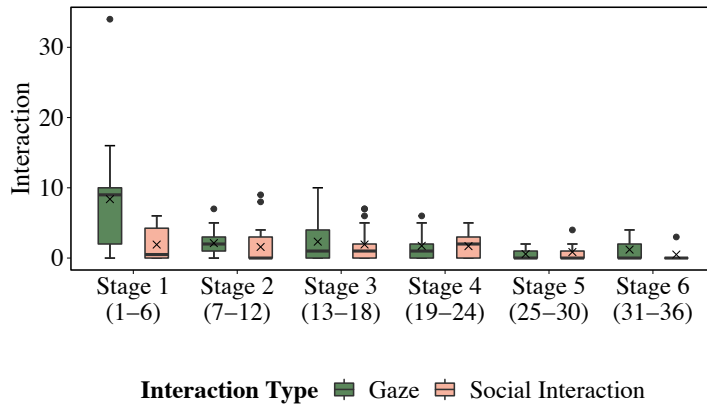
Fig. 12 Gaze and social interaction of the patients with the robot over the duration of the CR programme. While gaze decreased over time, social interaction was maintained throughout the long-term rehabilitation. The mean value per stage is marked with X.

crease towards the end of the programme. The patients' corrections are higher in stage 3 ($p = .004, \delta = 1.28$) and stage 4 ($p = .03, \delta = 1.16$) in comparison to the initial stage, which may have resulted from the physical intensity of the exercise. During the sessions, the experimenters observed that the patients had more difficulty in achieving a straight posture when the intensity increases were higher. We could not analyse the compliance for the *social robot* condition, due to the lack of video data or recorded gaze direction.

Fig. 12 shows the *gaze* and *social interaction* for the *personalised robot* scenario, as obtained from the video analysis. There is a moderate agreement between raters for gaze (Cohen's $\kappa = 0.712, z = 9.55, p < .001$) and a minimal agreement for social interactions ($\kappa = 0.319, z = 3.54, p < .001$), which could be due to the subtle cues in facial expressions that may go unnoticed. While *gaze* generally decreased over the duration of the CR programme ($T_{WJ}(5, 26) = 4.43, p = .005$), in contrast with P_{3b}, a significant difference only exists between stage 2 and 5 ($p = .02, \delta = -1.24$). Although we expected that *gaze* would be maintained over time due to the personalisation features, either a decrease in the novelty effect or an increase in the physical activity intensity may have affected the gaze. The experimenters observed that as the intensity is increasing during the sessions, focusing on the robot becomes more challenging. On the other hand, *social interaction* was maintained throughout the programme ($T_{WJ}(5, 27) = 1.31, p = .29$) validating P_{3c}, and social interactions occurred often in response to the personalised behaviours of the robot (such as progress feedback or correct user recognition), which indicate that the personalisation features help maintain the social interaction over the long-term rehabilitation.

The patients interacted with the *personalised robot* in a variety of ways, such as talking to the robot, smiling at being recognised or giving a negative response for an incorrect recognition, smiling after motivational feedback, or saying "Bye!" to the robot on their last session. They also talked to other patients about the benefits of the robot during the session. The video recordings showed that the medical staff

promptly responded to the robot's requests for assistance, and also interacted with the robot repeatedly in several sessions, such as playfully trying to capture its gaze, thanking the robot, joking with it, or talking about its benefits to other medical staff, which often elicited positive reactions (e.g., smile, laugh, nod) during the interaction from the patients in the study and those around.

6.4 Perception of the Robot

Fig. 13 and Table 9 presents the Unified Theory of Acceptance and the Use of Technology (UTAUT) questionnaire results and the significant differences between the conditions. The perceptions of the patients that completed the CR programme with the *social robot* are significantly more positive than the expectations of the *baseline group*, in terms of perceived usefulness, utility, ease of use, and trust, in agreement with the preliminary results (Casas et al., 2019). On the one hand, the patients perceived the *personalised robot* significantly safer and trusted it more than the *baseline group*. On the other hand, the perceived utility of the *personalised robot* was significantly less than the *social robot*, in contrast with P_{4a}. We believe this may be due to the user recognition and recall problems that we experienced within the sessions, which may have caused negative experiences (Hancock et al., 2011). Only 38% of the known users were correctly recognised by user recognition, and 44% of the new users were correctly detected. The poor performance was due to the failures arising from face recognition, which identified most users as new, only correctly identifying 35% of known users, significantly less than the multi-modal user recognition model, MMIBN ($p = .01$). Nonetheless, both the *social robot* and the *personalised robot* conditions improved the expectations about the robot and the system. Moreover, the responses to the questions developed for the *personalised robot* condition (Table 10) showed that the patients highly enjoyed the personalisation features (perceived enjoyment (PE): $Mdn = 5$ on a scale from 1 to 5), especially because the robot used their name, recognised them and tracked their progress, validating P_{4b}. Moreover, the usefulness of the *personalised robot*, such as feeling engaged with the programme and motivation to come to the sessions, was highly positive (perceived usefulness, $Mdn = 5$), validating P_{1b}. Usefulness and enjoyment are important factors for long-

Table 9 Mann-Whitney U-test results for the Unified Theory of Acceptance and the Use of Technology (UTAUT) questionnaire (non-personalised questions) for the *baseline group*, the *social robot* and *personalised robot* conditions. The significant differences ($p < .05$) are highlighted in bold.

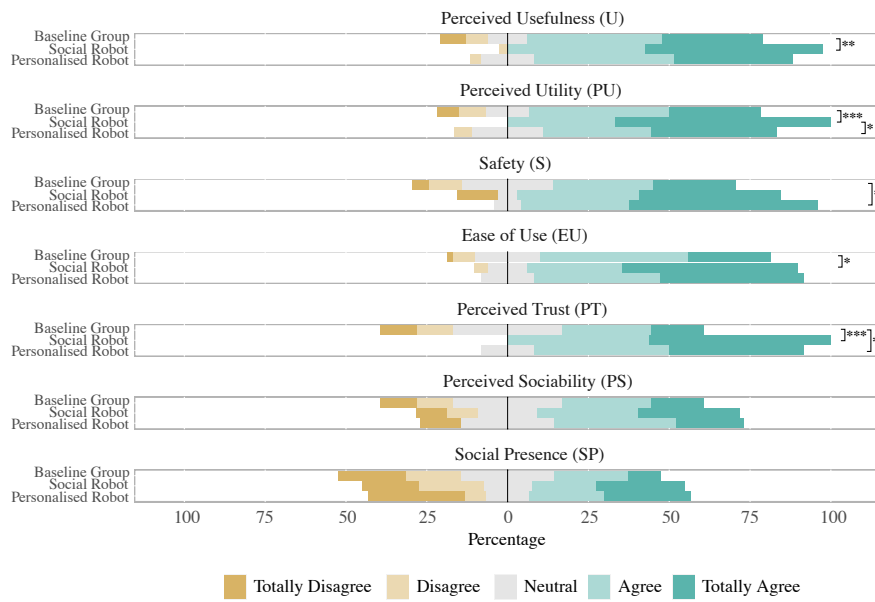| Construct	Baseline Group - Social Robot	Baseline Group - Personalised Robot	Social Robot - Personalised Robot
Perceived Usefulness (U)	**.002**	.35	.07
Perceived Utility (PU)	**$p < .001$**	.49	**.04**
Safety (S)	.22	**.02**	.28
Ease of Use (EU)	**.03**	.13	.70
Perceived Trust (PT)	**$p < .001$**	**.03**	.17
Perceived Sociability (PS)	.10	.26	.65
Social Presence (SP)	.17	.34	.78

Fig. 13 Unified Theory of Acceptance and the Use of Technology (UTAUT) questionnaire results for the *baseline group*, the *social robot* and *personalised robot* conditions. The patients in the baseline group did not have prior experience of the system, and completed the questionnaire after the debriefing and video demonstrations of the *social robot*. The patients in the robot conditions completed the questionnaire after their last session of the outpatient phase of the cardiac rehabilitation programme (i.e., after completing the study). Significant differences are denoted with $p < .05$:*, $p < .01$:**, $p < .001$:***, as presented in Table 9.

term acceptance of the robot (de Graaf et al., 2016). In addition, most patients felt attached to the robot at the end of the programme ($Mdn = 4.5$).

The additional feedback (through open questions) of the patients in the *personalised robot* condition was similar to that of previous findings of the *social robot* condition (Casas et al., 2019). All patients recommended the system for future patients, and commented on its usefulness, personalisation and effects on user motivation, such as:

– *"The cardiac rehabilitation with the robot will help you to recover as quickly as possible, and you will be able to progress by being linked to the robot."* - P6
– *"I feel confident in doing the rehabilitation with the robot, because I know that it is personalised and constantly monitoring my performance and progress."* - P6
– *"I really like the idea of the robot, as he was constantly monitoring. Also, I think the corrections the robot made are good, it keeps me focused on the therapy."* - P5
– *"It feels more comfortable being on a treadmill with the robot and because the robot is more aware of the patient."* - P5
– *"Working with the robot motivates me."* - P1, P6
– *"Working with the robot makes me feel happy."* - P1

– *"The robot interacts in a positive way with me, it helps me along with the medical staff, and it is also a good tool for them. I would not change anything (about the robot)." - P3*
– *"I would recommend the robot, it is a great help." - P2*

Nonetheless, the patients noted the need for improving the robustness of the user recognition and sensors, and decreasing the repetitiveness of the robot phrases, which was also mentioned in the previous study with the *social robot* (Casas et al., 2019). In addition, one patient found the appearance and the sound of the robot to be childish. Furthermore, because the progress feedback addressed the difficulties experienced in the sessions, some of the patients had concerns that they were not recovering well.

Through the WAI questionnaire (Table 11), we can analyse how the patients' overall perception of the robot changed over the long-term CR programme. Fig. 14 shows the patients' responses at the middle of the programme and at the final session. Wilcoxon signed-rank test shows that there is a significant improvement between the perceived goal construct in the positive formulation ($p = .003$, $V = 42$) from the middle of the CR programme to the final session. No significant differences are found in the other constructs between the tests ($p > .05$). The positive formulation (on the right of the scale) of *Bond*, *Goal* and *Task* show that the robot and the CR programme were generally positively perceived, and the patients maintained their bond with the robot over the duration of the programme, validating P_{3d}. Furthermore, the patients generally disagree with the negatively formulated questions (on the left of the scale), such as "I feel uncomfortable with the robot.", indicating that despite the negative user experience due to sensor and recognition failures, patients perceived the robot highly positively, which may support that personalisation mitigates the negative user experience, similar to (Irfan et al., 2020b). Another possibility is that the benefits patients felt from the intervention mitigated their negative feelings about the errors.
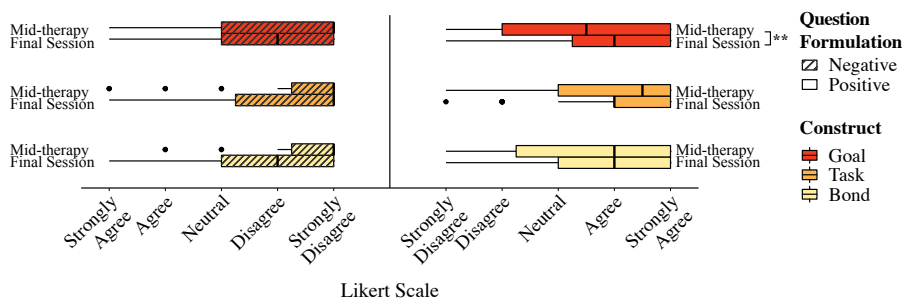

Fig. 14 Working Alliance Inventory responses for the *personalised robot* condition, evaluated at the middle of the CR programme (18th session) and the final session. The results suggest that the patients' positive perception of the robot was maintained throughout the programme. A significant improvement was achieved for the perceived goal construct in positive formulation ($p = .003$, $V = 42$).

7 Discussion

7.1 Hypotheses and Predictions

Previous studies in SAR (in Section 2) highlighted the importance of encouraging feedback and continuous monitoring during repetitive exercise for healthcare programmes to increase motivation and enhance task performance. Other research has shown the value of personalisation for long-term HRI to improve users' motivation, task performance, engagement and perceptions. Correspondingly, this work aimed to address how these previous findings translate to long-term cardiac rehabilitation programmes, through the use of fully autonomous generic and personalised socially assistive robots in a clinical environment with non-expert users, that is, the patients and the clinicians.

Our study showed partial support for **H1** (*Personalisation will improve patient motivation and adherence to the CR programme*): adherence (P_{1a}) could not be evaluated properly, but an increase in motivation for the programme (P_{1b}) was validated. **H2** (*Personalisation will improve the cardiovascular performance of the patients*) was also partially supported: while a lower gain of normalised recovery heart rate was found with the personalised robot, thus, not validating P_{2a} probably due to an imbalance of participants initial health levels or the higher training intensity, the lower number of alerts (P_{2b}) was validated. **H3** (*Interaction with the personalised robot will be maintained throughout the long-term programme*) was mostly supported by our study: except for an initial decrease in gaze (in contrast to P_{3b}), patients complied with the robot (P_{3a}), and the social interaction (P_{3c}) and bond (P_{3d}) with the robot were maintained throughout the long-term programme, validating these predictions. Finally, **H4** (*Personalisation will improve patients' perceptions of the robot*) was partially supported by our results: while the personalised robot was rated lower in utility (in contrast to P_{4a}) due to recognition and recall problems, personalised features were acknowledged and enjoyed (validating P_{4b}).

Overall, our long-term study showed that personalisation presents promises for SAR in cardiac rehabilitation programmes. We draw on these findings and highlight our key takes and suggestions based on the drawbacks and limitations, to benefit future researchers exploring SAR for long-term interactions in the real world.

7.2 Benefits of Personalisation

The main challenges of cardiac rehabilitation are providing close monitoring of patients within the group sessions and assuring adherence to the long-term programme to ensure that the patients recover fully and retain healthy habits. Our study showed that robots motivate patients to continue the programme and finish the programme earlier. Moreover, patients assisted by the *personalised robot* acknowledged that the personalisation features, such as the progress feedback and adherence tracking, encourage them to come to the CR sessions.

The goal of the outpatient phase of the cardiac rehabilitation programmes is improving cardiovascular functioning through structured exercises that progressively

increases in intensity to reduce the risk of suffering recurrent events and accelerating recovery. Our results indicate that the patients assisted by a (generic or personalised) robot achieved a significant improvement in their training and recovery heart rate and, thus, cardiovascular functioning. Moreover, the clinicians trusted the continuous monitoring aspects of the robots (Casas et al., 2019; Céspedes et al., 2021), which may have reinforced applying high-intensity training throughout the CR programme for the *personalised robot*, causing a greater increment in the patients' training heart rate. The high-intensity training did not increase the perceived exertion level, and the patients experienced a lower frequency in reaching the critical heart rates in comparison to the patients in the *social robot* condition, further supporting the improvement in patients' cardiovascular functioning. Nonetheless, the adoption of the technology by the medical staff and their immediate intervention in critical cases played an important role in achieving this improvement (Irfan et al., 2020a).

Throughout the CR programme, the patients assisted by the *personalised robot* maintained a better posture than the patients assisted by the *social robot* despite a higher exercise intensity. Moreover, the patients mostly complied with the *personalised robot*'s requests throughout the long-term rehabilitation. These findings show the importance of personalisation through addressing the person by their name and progress feedback such that the patients are more motivated to maintain and improve their good posture. Personalisation features also elicited gaze and social interaction with the robot, such as smiling when the robot addressed them with their name or upon correct user recognition, thanking the robot or talking to it after receiving personalised feedback. While gaze decreased after the initial session, which could be due to the fading novelty effect or the increasing exercise intensity requiring the patients to focus on the exercise, the social interaction with the personalised robot was maintained throughout the long-term programme, as intended. Future work can examine the behaviours of the patients assisted by the *social robot* through video analysis to compare the benefits of personalisation more thoroughly. However, the presence of a camera and videotaping may lead to a change of behaviour when observed (Irfan et al., 2018a), known as the Hawthorne effect (Roethlisberger et al., 1939), as well as selection bias on willingness to participate in the experiment, which could have affected the interaction of patients with the *personalised robot*, but this might have been offset during the long-term programme. Recording the gaze direction and duration through the tablet could be an alternative approach in detecting engagement (Oertel et al., 2020) that could eliminate such confounds.

Both the *social robot* and the *personalised robot* met the expectations about the system. The patients commended the use of both robots, and expressed feeling more secure due to continuous monitoring and immediate feedback. In addition, since the patients participated in group sessions, other patients that were not part of the study were able to observe its benefits, and several of them declared interest in working with a robot in their CR programme. Personalisation features were highly enjoyed by the patients, and the progress and adherence feedback were reported to be useful by the patients, suggesting improvement in motivation and increased adherence to the programme, as initially aimed. While the patients felt initially sceptical towards the robot's role due to their lack of prior experience (as observed by the experimenters), the clinicians noted that their trust in the measurements and its feedback increased

over time. These observations were supported by the WAI questionnaire results in which the perceived goal was significantly improved over time. Moreover, patients maintained their bond with the *personalised robot* and their perceptions over the duration of the programme. These findings are in line with the clinicians' perceptions that the personalisation of the robot improves the quality of the interaction (Céspedes et al., 2021).

Continuous monitoring and immediate feedback aspects of both robots were highly appreciated by the patients and the clinicians in terms of the awareness of patient performance within the session. Additionally, progress tracking throughout the programme by the *personalised robot* helped inform both the medical staff and the patients, which increased the awareness of the medical staff further to detect complications (Irfan et al., 2020a; Céspedes et al., 2021), and provided knowledge of their recovery to the patients, thereby, improving their perception of the CR programme and their motivation. While all of these aspects could be provided without the presence of a robot, i.e., through verbal feedback from the tablet interface, previous studies showed the added benefits of embodiment in improving compliance, likeability, social engagement, adherence and task performance, as discussed in Section 2. However, having a robot in a rehabilitation programme is not cost efficient, in terms of the initial investment and maintenance requirements, thus, future work can investigate whether the lack of a robot could reproduce the benefits found in this study.

7.3 Drawbacks of Personalisation and Suggestions for Future Work

Long-term studies are labour and time intensive, thus, it is challenging to recruit subjects willing to participate, especially for a rehabilitation programme with a novel system, as the patients may be sceptical towards the approach (Casas et al., 2019). Moreover, dropouts and the incomplete rehabilitation experienced due to unforeseen reasons such as funding, the outbreak of COVID-19, and other medical conditions, caused a relatively limited number of patients with skewed characteristics (i.e., gender, age and obesity) that may have affected the overall results. Furthermore, the patients were progressively recruited over 2.5 years and the *personalised robot* condition was started two years after the initial two conditions, which could have affected the initial technological acceptance and perceptions of patients towards robots (Pino et al., 2015). Nevertheless, our statistical analyses and the effect sizes indicate that the patients assisted by the *personalised robot* improved their cardiovascular functioning in terms of their training heart rate and endurance (evident by low exertion levels and lower critical heart rate alerts) with the high-intensity training, along with a better cervical posture, in comparison to the *social robot*. However, personalisation was found to have drawbacks.

While the *personalised robot* led to a higher increment from the patients' initial recovery heart rate than the conventional CR programme, which is one of the primary parameters that reflect the patient's cardiovascular performance, the *social robot* presented a greater and more rapid improvement. However, this result could be due to the higher intensity training applied to the patients assisted by the *personalised*

robot, hence, a further balanced (i.e., same level of exercise intensity) experiment is necessary to confirm the findings.

Relying on full autonomy in a real-world environment brought about challenges such as malfunctioning in the sensors and the robot, as well as connection problems with the tablet interface, which may have caused negative perceptions of the robot, in addition to missing data. Especially relying on a user recognition system for personalisation, decreased the perceived usefulness, utility and trust to the robot, and the low reliability of the user recognition system was remarked by the patients. While the multi-modal incremental Bayesian network (MMIBN) with online learning is the only method that supports sequential and incremental recognition of previous and new users without requiring pre-training for real-time HRI, user recognition can be improved by using identifiers with lower noise, integrating additional non-intrusive modalities (e.g., voice, facial marks, eye colour) or by using other online learning methods. Another option is to remove user recognition, e.g., through requesting the patient's name from the tablet interface either from the patient or the clinician, however, the questionnaires and video analysis showed that the patients enjoyed being recognised, and it could also decrease the naturalness of the interaction. Moreover, the acquisition of more robust medical sensors is necessary to ensure the reliability of the data. Nonetheless, the highly positive perceptions of the patients indicate that negative user experiences can be overcome with the added benefits of the robot.

Patients and clinicians remarked that both robots were repetitive, thus, novel and a larger variety of robot responses should be added to improve the sociability and social presence of the robot, especially in long-term interactions. A future research direction can be to adapt these responses based on the patient's sensory values, and adapt progress feedback to reflect the overall (or stage) progress instead of comparing with the previous session, to keep the interactions engaging and interesting in the long-term (Matarić and Scassellati, 2016).

While speech and emotion recognition can be integrated with the robot to improve the naturalness and adaptivity of the interaction, these may not reliable in a noisy real-world environment. For instance, Fundación Cardioinfantil-Instituto de Cardiología plays loud music to motivate the patients in their exercise, which could cause failures in speech recognition. Thus, gesture recognition could be a better alternative. Emotions can be misleading due to the patients' exertion levels, hence, extracting fatigue levels from cameras and wearable sensors could be more reliable and provide a complementary measure to the Borg scale for detecting when to provide motivation and additional alerts (Pinto et al., 2020).

The appearance and voice of the (NAO) robot were found to be childish by one of the patients, which could have affected the underlying perceptions of the patients (Goetz et al., 2003; Pino et al., 2015). Future work can use other taller and mobile platforms, such as a Pepper robot (SoftBank Robotics Europe, France), to also address other patients in the group session. Moreover, other robotic platforms with facial expressions, such as Nexi (MIT Media Lab, USA), can be used to improve the sociability of the robot (Johanson et al., 2020).

While the architecture of our patient-robot interface and personalisation features were designed in collaboration with medical specialists, this study showed the various additional features that can be changed (e.g., progress feedback structure) or

added to the system based on the needs of the patients, highlighting the importance of *co-design* (or participatory design). Co-design is the process where users (stakeholders) are involved in designing a product from the idea generation stage (Sanders and Stappers, 2008). Involving patients and medical specialists in the design of a socially assistive system not only allows adapting to their (changing) needs, but also enabling them to understand the limitations of the system and actively contribute to the design of their care (Bate and Robert, 2007; Šabanović, 2010).

8 Conclusion

This paper presented a personalised socially assistive robot for the outpatient phase of a long-term cardiac rehabilitation programme. The aim of the socially assistive robot is to improve motivation and adherence to the programme. Personalisation features, such as recognising patients, addressing by their name, tracking their attendance and providing progress feedback, were developed to improve and maintain motivation over the long-term interaction. Three conditions were designed to evaluate the impact of socially assistive robotics and personalisation on conventional cardiac rehabilitation, labelled as *social robot*, *personalised robot* and *control*, respectively. 43 patients were recruited for the study, however, due to the dropouts and the external unforeseen reasons, such as funding, the outbreak of COVID-19, and other medical conditions, 26 patients (9 for *control*, 11 for *social robot*, 6 for *personalised robot*) actively participated and completed the outpatient phase of cardiac rehabilitation and the study.

The *social* and *personalised robots* were found to improve cardiovascular functioning in comparison to a conventional cardiac rehabilitation programme. Furthermore, the patients assisted by a robot completed their rehabilitation in a shorter duration, suggesting benefits on adherence. Moreover, the perceptions of the patients and the clinicians improved in comparison to expectations, and they were recommended for future use by both groups. The continuous monitoring of the robot enabled prompt detection of critical conditions, which may have reinforced the trust of the clinicians in the robot and thus, providing high-intensity training in the *personalised robot* condition. This resulted in a higher training heart rate without an adverse effect on the endurance (low perceived exertion levels and critical heart rates), however, a lower and slower improvement in the recovery heart rate in comparison to the patients assisted by a *social robot*. Moreover, relying on a fully autonomous robot for personalisation in long-term rehabilitation brought along sensor and user recognition failures, which decreased the perceived utility of the robot. On the other hand, personalised features often elicited gaze and social interaction, facilitated a bond with the user, and were highly positively perceived and this perception was maintained in the long-term, suggesting that various benefits of personalisation can overcome its drawbacks, supporting the potential for improving the conventional cardiac rehabilitation programmes and the long-term interaction with a socially assistive robot.

Funding

Conflicts of interest

The authors have no conflicts of interest to declare that are relevant to the content of this article.

Availability of data and material

The data consists of patients' medical records, hence, they are not made publicly available due to their sensitive nature.

Code availability

The multi-modal incremental Bayesian network (Irfan et al., 2018b, 2021) for user recognition is available online[4]. However, other parts of the patient-robot interface are not publicly available.

Author Contributions

B.I. developed the personalisation features of the robot. N.C. performed the clinical assessment and data processing. B.I. and N.C. performed the statistical analysis. N.C. and J.C. developed the sensor interface. B.I. and E.S. contributed to the development of the patient-robot interface for cardiac rehabilitation. B.I. and N.C. led the manuscript writing. M.R. and L.G. are the medical specialists that managed the experimental protocol. M.M. developed the experimental protocol. E.S., T.B., M.M., J.C. and C.C. were involved in the revisions and corrections of the manuscript.

---

[4] `https://github.com/birfan/MultimodalRecognition`

Ethics approval

This study was approved by the Ethics Committee of the Clinic Fundación Cardio-Infantil Instituto de Cardiología located in Bogotá, Colombia.

Consent to participate

The participants gave their written consent to take part in the study. The patients in the personalised robot condition also gave their written consent for video recording.

A Questionnaires

Table 10 Additional questions developed for the Unified Theory of Acceptance and the Use of Technology (UTAUT) questionnaire to evaluate the perceptions of the patients specific to the *personalised robot* condition, in terms of perceived usefulness (U), perceived utility (PU), perceived enjoyment (PE), perceived adaptivity (PA), perceived sociability (PS), social presence (SP) and attitude (A).

Construct	Question
U	I feel encouraged to come to the sessions. I feel engaged in the therapy. I feel that the robot helped me progress in my therapy. I feel encouraged about my therapy when the robot comments on my session performance.
PU	The robot recognises me correctly. The robot remembers my previous sessions correctly. The robot tracks my session performance correctly.
PE	I am pleased that the robot recognises me. I am pleased that the robot uses my name. I am pleased to hear about my therapy progress. I am pleased that the robot remembers me. I am pleased to work with the robot.
PA	I feel that the robot personalises its interaction.
PS	I feel that the robot knows me well.
SP	I feel that the robot has a personality. I feel compelled to come to the sessions because the robot comments on my absence.
A	I feel attached to the robot.

Table 11 Adapted Working Alliance Inventory (WAI) that measures the long-term perception of the robot within the *personalised robot* condition, with *Bond*, *Task* and *Goal* constructs.

Construct	Formulation	Question
Bond	Positive	The robot and I understand each other. I believe the robot likes me. I believe the robot is genuinely concerned for my welfare. The robot and I respect each other. I am confident in the robot's ability to help me. I feel that the robot appreciates me. The robot and I trust one and other. My relationship with the robot is very important to me. I feel the robot cares about me even when I do things that the robot does not understand me.
	Negative	I feel uncomfortable with the robot. I feel the robot is not totally honest about its feelings toward me. I have the feeling that if I say or do the wrong things the robot will stop working with me.
Task	Positive	The robot and I agree about the things I will need to do in the therapy to help improve my situation. What I am doing in the therapy gives me new ways of looking at my problem. I am clear on what my responsibilities are in therapy. I feel that the things I do in therapy will help me to accomplish the changes that I want. I am clear as to what the robot wants me to do in these sessions. The robot and I are in agreement on what is important for me to work on. I believe the way that the robot and I are working in my problem is correct.
	Negative	I find what I am doing in therapy confusing. I believe the time robot and I are spending together is not spent efficiently. I find that the robot tasks during the therapy are unrelated to my concerns. I am frustrated by the things I am doing in therapy. The things that robot is requesting from me do not make sense.
Goal	Positive	The robot perceives accurately what my goals are. I wish that the robot could configure the therapy according the purpose of our session. The goals of these session are important to me. The robot and I are working towards mutually agreed upon goals. As a result of these session I am clearer as to how I might be able to change. The robot and I collaborate on setting goals for my therapy. The robot and I established a good understanding of the kind of changes that would be good for me.
	Negative	I am worried about the outcome of these sessions. I disagree with the robot about what I ought to get out of therapy. The robot does not understand what I am trying to accomplish in therapy. The robot and I have different ideas on what my problems are. I do not know what to expect as the result of my therapy.

References

Aamot IL, Forbord SH, Karlsen T, Støylen A (2014) Does rating of perceived exertion result in target exercise intensity during interval training in cardiac rehabilitation? A study of the Borg scale versus a heart rate monitor. Journal of Science and Medicine in Sport 17(5):541–545, DOI 10.1016/j.jsams.2013.07.019

Admoni H, Scassellati B (2017) Social eye gaze in human-robot interaction: a review. Journal of Human-Robot Interaction 6(1):25–63

Ahmad MI, Mubin O, Orlando J (2017) A systematic review of adaptivity in human-robot interaction. Multimodal Technologies and Interaction 3(14)

Ahmad MI, Mubin O, Shahid S, Orlando J (2019) Robot's adaptive emotional feedback sustains children's social engagement and promotes their vocabulary learning: a long-term child-robot interaction study. Adaptive Behavior 27(4):24–266, DOI 10.1177/1059712319844182

Andriella A, Torras C, Alenyá G (2020) Short-term human–robot interaction adaptability in real-world environments. International Journal of Social Robotics 12:639–657, DOI 10.1007/s12369-019-00606-y

Anzalone SM, Boucenna S, Ivaldi S, Chetouani M (2015) Evaluating the Engagement with Social Robots. International Journal of Social Robotics 7(4):465–478, DOI 10.1007/s12369-015-0298-7

Bainbridge WA, Hart J, Kim ES, Scassellati B (2008) The effect of presence on human-robot interaction. In: RO-MAN 2008 - The 17th IEEE International Symposium on Robot and Human Interactive Communication, pp 701–706, DOI 10.1109/ROMAN.2008.4600749

Bate P, Robert G (2007) Bringing User Experience to Healthcare Improvement: The Concepts, Methods and Practices of Experience-Based Design. Radcliffe Publishing, Oxford

Belpaeme T, Baxter P, Read R, Wood R, Cuayáhuitl H, Kiefer B, Racioppa S, Kruijff-Korbayová I, Athanasopoulos G, Enescu V, Looije RM, Neerincx MA, Demiris Y, Ros-Espinoza R, Beck A, Cañamero L, Hiolle A, Lewis M, Baroni I, Nalin M, Cosi P, Paci G, Tesser F, Sommavilla G, Humbert R (2013) Multimodal child-robot interaction: Building social bonds. J Hum-Robot Interact 1(2):33–53

Bethell H, Lewin R, Dalal H (2009) Cardiac rehabilitation in the united kingdom. Heart 95(4):271–275, DOI 10.1136/hrt.2007.134338

Bickmore TW, Picard RW (2005a) Establishing and maintaining long-term human-computer relationships. ACM Trans Comput-Hum Interact 12(2):293–327, DOI 10.1145/1067860.1067867

Bickmore TW, Picard RW (2005b) Establishing and maintaining long-term human-computer relationships. ACM Transactions on Computer-Human Interaction 12(2):293–327, DOI 10.1145/1067860.1067867

Blanson Henkemans OA, Bierman BP, Janssen J, Neerincx MA, Looije R, van der Bosch H, van der Giessen JA (2013) Using a robot to personalise health education for children with diabetes type 1: A pilot study. Patient Education and Counseling 92(2):174–181, DOI 10.1016/j.pec.2013.04.012

Borg G (1998) Borg's perceived exertion and pain scales. Human Kinetics

Breazeal C (2011) Social robots for health applications. In: 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp 5368–5371, DOI 10.1109/IEMBS.2011.6091328

Broadbent E, Garret J, Jepsen N, Ogilvie V, Seok Ahn H, Robinson H (2018) Using Robots at Home to Support Patients With Chronic Obstructive Pulmonary Disease: Pilot Randomized Controlled Trial. J Med Internet Res 20(2):e8640, DOI 10.2196/JMIR.8640, URL https://www.jmir.org/2018/2/e45

Broz F, Lehmann H, Nakano Y, Mutlu B (2012) Gaze in hri: from modeling to communication. In: Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction, pp 491–492

Campos J, Kennedy J, Lehman JF (2018) Challenges in exploiting conversational memory in human-agent interaction. In: International Conference on Autonomous Agents and Multiagent Systems

Cao HL, Esteban PG, Bartlett M, Baxter P, Belpaeme T, Billing E, Cai H, Coeckelbergh M, Costescu C, David D, et al. (2019) Robot-enhanced therapy: Development and validation of supervised autonomous robotic system for autism spectrum disorders therapy. IEEE robotics & automation magazine 26(2):49–58

Carlson J, Johnson J, Franklin B, VanderLaan R (2000) Program participation, exercise adherence, cardiovascular outcomes, and program cost of traditional versus modified cardiac rehabilitation. Am J Cardiol 86(1):17–23, DOI 10.1016/s0002-9149(00)00822-5

Casas J, Gomez NC, Senft E, Irfan B, Gutiérrez LF, Rincón M, Múnera M, Belpaeme T, Cifuentes CA (2018) Architecture for a social assistive robot in cardiac rehabilitation. In: 2018 IEEE 2nd Colombian

Conference on Robotics and Automation (CCRA), pp 1–6, DOI 10.1109/CCRA.2018.8588133

Casas J, Irfan B, Senft E, Gutiérrez L, Rincon-Roncancio M, Munera M, Belpaeme T, Cifuentes CA (2018a) Social assistive robot for cardiac rehabilitation: A pilot study with patients with angioplasty. In: Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction, ACM, pp 79–80, DOI 10.1145/3173386.3177052

Casas J, Irfan B, Senft E, Gutiérrez L, Rincon-Roncancio M, Munera M, Belpaeme T, Cifuentes CA (2018b) Towards a sar system for personalized cardiac rehabilitation: A patient with pci. In: 2018 ACM/IEEE International Conference on Human-Robot Interaction Personal Robots for Exercising and Coaching workshop, ACM

Casas J, Céspedes N, Cifuentes C, Gutierrez LF, Rincón-Roncancio M, Múnera M (2019) Expectation vs. reality: Attitudes towards a socially assistive robot in cardiac rehabilitation. Appl Sci 9:4651, DOI 10.3390/app9214651

Casas J, Senft E, Gutiérrez L, Rincón-Roncancio M, Múnera M, Belpaeme T, Cifuentes CA (2020) Social assistive robots: Assessing the impact of a training assistant robot in cardiac rehabilitation. International Journal of Social Robotics DOI 10.1007/s12369-020-00708-y

Castellano G, Aylett R, Dautenhahn K, Paiva A, McOwan PW, Ho S (2008) Long-term affect sensitive and socially interactive companions. In: 4th International Workshop on Human-Computer Conversation

Céspedes N, Irfan B, Senft E, Cifuentes CA, Gutierrez LF, Rincon-Roncancio M, Belpaeme T, Munera M (2021) A socially assistive robot for long-term cardiac rehabilitation in the real world. Frontiers in Neurorobotics DOI 10.3389/fnbot.2021.633248

Churamani N, Anton P, Brügger M, Flieundefinedwasser E, Hummel T, Mayer J, Mustafa W, Ng HG, Nguyen TLC, Nguyen Q, et al (2017) The impact of personalisation on human-robot interaction in learning scenarios. In: Proceedings of the 5th International Conference on Human Agent Interaction, Association for Computing Machinery, New York, NY, USA, pp 171–180, DOI 10.1145/3125739.3125756

Clabaugh C, Mahajan K, Jain S, Pakkar R, Becerra D, Shi Z, Deng E, Lee R, Ragusa G, Matarić M (2019) Long-term personalization of an in-home socially assistive robot for children with autism spectrum disorders. Frontiers in Robotics and AI 6:110, DOI 10.3389/frobt.2019.00110

Clave C, Cafaro A, Campano S, Pelachaud C (2016) Fostering user engagement in face-to-face human-agent interactions: A survey. Toward Robotic Socially Believable Behaving Systems II:93–120

Cohen J (1960) A coefficient of agreement for nominal scales. Educational and Psychological Measurement 20(1):37–46, DOI 10.1177/001316446002000104

Cohen J (1988) Statistical Power Analysis for the Behavioral Sciences. Routledge Academic, New York, NY

Coninx A, Baxter P, Oleari E, Bellini S, Bierman B, Blanson Henkemans O, Cañamero L, Cosi P, Enescu V, Ros Espinoza R, Hiolle A, Humbert R, Kiefer B, Kruijff-Korbayovà I, Looije R, Mosconi M, Neerincx M, Paci G, Patsis G, Pozzi C, Sacchitelli F, Sahli H, Sanna A, Sommavilla G, Tesser F, Demiris Y, Belpaeme T (2015) Towards Long-Term Social Child-Robot Interaction: Using Multi-Activity Switching to Engage Young Users. Journal of Human-Robot Interaction 5(1):32, DOI 10.5898/jhri.5.1.coninx

Dantcheva A, Elia P, Ross A (2016) What else does your biometric data reveal? a survey on soft biometrics. IEEE Transactions on Information Forensics and Security 11(3):441–467, DOI 10.1109/TIFS.2015.2480381

Deng E, Mutlu B, Mataric MJ (2019) Embodiment in socially interactive robots. Foundations and Trends in Robotics 7(4):251–356, DOI 10.1561/2300000056

Dietterich TG (1998) Approximate statistical tests for comparing supervised classification learning algorithms. Neural Computation 10(7):1895–1923, DOI 10.1162/089976698300017197

Duffy BR, Rooney CFB, Hare GMPO, Donoghue RPSO (1999) What is a Social Robot ? In: 10th Irish Conference on Artificial Intelligence & Cognitive Science, pp 1–3

Fasola J, Mataric MJ (2012) Using socially assistive human–robot interaction to motivate physical exercise for older adults. Proceedings of the IEEE 100(8):2512–2526, DOI 10.1109/JPROC.2012.2200539

Fasola J, Matarić MJ (2013) A socially assistive robot exercise coach for the elderly. J Hum-Robot Interact 2(2):3–32

Feil-Seifer D, Matarić MJ (2005) Defining socially assistive robotics. In: 9th International Conference on Rehabilitation Robotics (ICORR) 2005, pp 465–468, DOI 10.1109/ICORR.2005.1501143

Feingold Polak R, Tzedek SL (2020) Social Robot for Rehabilitation: Expert Clinicians and Post-Stroke Patients' Evaluation Following a Long-Term Intervention, Association for Computing Machinery, New York, NY, USA, p 151–160. URL https://doi.org/10.1145/3319502.3374797

Feng Y, Barakova EI, Yu S, Hu J, Matthias Rauterberg GW (2020) Effects of the level of interactivity of a social robot and the response of the augmented reality display in contextual interactions of people with dementia. Sensors (Switzerland) 20(13):1–12, DOI 10.3390/s20133771

Fernaeus Y, Håkansson M, Jacobsson M, Ljungblad S (2010) How do you play with a robotic toy animal? a long-term study of pleo. In: Proceedings of the 9th International Conference on Interaction Design and Children, Association for Computing Machinery, New York, NY, USA, IDC '10, pp 39–48, DOI 10.1145/1810543.1810549

Glas DF, Wada K, Shiomi M, Kanda T, Ishiguro H, Hagita N (2017) Personal greetings: Personalizing robot utterances based on novelty of observed behavior. International Journal of Social Robotics 9(2):181–198

Glass GV, McGaw B, Smith ML (1981) Meta-analysis insocial research. Sage, Newbury Park, CA

Gockley R, Bruce A, Forlizzi J, Michalowski M, Mundell A, Rosenthal S, Sellner B, Simmons R, Snipes K, Schultz AC, Jue Wang (2005) Designing robots for long-term social interaction. In: 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp 1338–1343, DOI 10.1109/IROS.2005.1545303

Goetz J, Kiesler S, Powers A (2003) Matching robot appearance and behavior to tasks to improve human-robot cooperation. In: The 12th IEEE International Workshop on Robot and Human Interactive Communication, 2003. Proceedings. ROMAN 2003., pp 55–60, DOI 10.1109/ROMAN.2003.1251796

de Graaf MMA, Ben Allouch S, van Dijk JAGM (2016) Long-term evaluation of a social robot in real homes. Interaction Studies 17(3):462–491, DOI https://doi.org/10.1075/is.17.3.08deg

Hammill BG, Curtis LH, Schulman KA, Whellan DJ (2010) Relationship between cardiac rehabilitation and long-term risks of death and myocardial infarction among elderly medicare beneficiaries. Circulation 121(1):63–70, DOI 10.1161/CIRCULATIONAHA.109.876383

Hancock PA, Billings DR, Schaefer KE, Chen JYC, de Visser EJ, Parasuraman R (2011) A meta-analysis of factors affecting trust in human-robot interaction. Human Factors 53(5):517–527, DOI 10.1177/0018720811417254, pMID: 22046724

Heerink M, Kröse B, Evers V, Wielinga B (2010) Assessing acceptance of assistive social agent technology by older adults: the almere model. Int J of Soc Robotics 2:361–375, DOI 10.1007/s12369-010-0068-5

Hemminahaus J, Kopp S (2017) Towards adaptive social behavior generation for assistive robots using reinforcement learning. In: 2017 12th ACM/IEEE International Conference on Human-Robot Interaction (HRI, pp 332–340

Ho WC, Dautenhahn K, Lim MY, Casse KD (2010) Modelling human memory in robotic companions for personalisation and long-term adaptation in hri. In: Samsonovich AV, Johannsdottir KR, Chella A, Goertzel B (eds) BICA, IOS Press, Frontiers in Artificial Intelligence and Applications, vol 221, pp 64–71

Hoffman G, Breazeal C (2010) Effects of anticipatory perceptual simulation on practiced human-robot tasks. Autonomous Robots 28(4):403–423, DOI 10.1007/s10514-009-9166-3

Horvath AO, Greenberg LS (1989) Development and validation of the working alliance inventory. Journal of Counseling Psychology 36(2):223–233, DOI 10.1037/0022-0167.36.2.223

Hung L, Liu C, Woldum E, Au-Yeung A, Berndt A, Wallsworth C, Horne N, Gregorio M, Mann J, Chaudhury H (2019) The benefits of and barriers to using a social robot paro in care settings: a scoping review. BMC Geriatrics 19(232), DOI 10.1186/s12877-019-1244-6

Irfan B, Kennedy J, Lemaignan S, Papadopoulos F, Senft E, Belpaeme T (2018a) Social psychology and human-robot interaction: An uneasy marriage. In: Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction, ACM, pp 13–20, DOI 10.1145/3173386.3173389

Irfan B, Lyubova N, Garcia Ortiz M, Belpaeme T (2018b) Multi-modal open-set person identification in hri. In: 2018 ACM/IEEE International Conference on Human-Robot Interaction Social Robots in the Wild workshop

Irfan B, Ramachandran A, Spaulding S, Glas DF, Leite I, Koay KL (2019) Personalization in long-term human-robot interaction. In: 2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI), IEEE, pp 685–686, DOI 10.1109/HRI.2019.8673076

Irfan B, Céspedes Gomez N, Casas J, Senft E, Gutiérrez LF, Rincon-Roncancio M, Munera M, Belpaeme T, Cifuentes CA (2020a) Using a personalised socially assistive robot for cardiac rehabilitation: A long-term case study. In: 29th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), IEEE, pp 124–130, DOI 10.1109/RO-MAN47096.2020.9223491

Irfan B, Hellou M, Mazel A, Belpaeme T (2020b) 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, Association for Computing Machinery, New York, NY, USA,

HRI '20, pp 272–274, DOI 10.1145/3371382.3378278

Irfan B, Ortiz MG, Lyubova N, Belpaeme T (2021) Multi-modal open world user identification. ACM Transactions on Human-Robot Interaction (THRI) 11(1), DOI 10.1145/3477963

Jain AK, Dass SC, Nandakumar K (2004) Soft biometric traits for personal recognition systems. In: International Conference on Biometric Authentication, Springer, Hong Kong, China, no. 3072 in LNCS, pp 731–738, DOI 10.1007/978-3-540-25948-0_99

James GS (1951) The comparison of several groups of observations when the ratios of the population variances are unknown. Biometrika 38(3/4):324–329, DOI 10.2307/2332578

Jamieson S (2004) Likert scales: how to (ab)use them. Medical Education 38(12):1217–1218, DOI 10.1111/j.1365-2929.2004.02012.x

Johansen S (1980) The welch-james approximation to the distribution of the residual sum of squares in a weighted linear regression. Biometrika 67(1):85–92, DOI 10.2307/2335320

Johanson DL, Ahn HS, Broadbent E (2020) Improving interactions with healthcare robots: A review of communication behaviours in social and healthcare contexts. International Journal of Social Robotics DOI 10.1007/s12369-020-00719-9

Jolly K, Taylor R, Lip G, Greenfield S, Raftery J, Mant J, Lane D, Jones M, Lee KW, Stevens A (2007) The birmingham rehabilitation uptake maximisation study (brum). home-based compared with hospital-based cardiac rehabilitation in a multi-ethnic population: cost-effectiveness and patient adherence. Health Technology Assessment 11(35)

Kanda T, Hirano T, Eaton D, Ishiguro H (2004) Interactive robots as social partners and peer tutors for children: A field trial. Hum-Comput Interact 19(1):61–84, DOI 10.1207/s15327051hci1901\%262_4

Kanda T, Sato R, Saiwaki N, Ishiguro H (2007) A two-month field trial in an elementary school for long-term human-robot interaction. IEEE Transactions on Robotics 23(5):962–971, DOI 10.1109/TRO.2007.904904

Kanda T, Shiomi M, Miyashita Z, Ishiguro H, Hagita N (2010) A communication robot in a shopping mall. IEEE Transactions on Robotics 26(5):897–913, DOI 10.1109/TRO.2010.2062550

Kang KI, Freedman S, Mataric MJ, Cunningham MJ, Lopez B (2005) A hands-off physical therapy assistance robot for cardiac patients. In: 9th International Conference on Rehabilitation Robotics, 2005 (ICORR 2005), IEEE, pp 337–340, DOI 10.1109/ICORR.2005.1501114

Kennedy J, Baxter P, Belpaeme T (2015) The robot who tried too hard: Social behaviour of a robot tutor can negatively affect child learning. In: Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction, Association for Computing Machinery, New York, NY, USA, HRI '15, pp 67–74, DOI 10.1145/2696454.2696457

Keselman HJ, Wilcox RR, LIX LM (2003) A generally robust approach to hypothesis testing in independent and correlated groups designs. Psychophysiology 40(4):586–596, DOI 10.1111/1469-8986.00060

Kidd CD, Breazeal C (2007) A robotic weight loss coach. In: Proceedings of the 22nd National Conference on Artificial Intelligence - Volume 2, AAAI Press, AAAI'07, pp 1985–1986

Kidd CD, Breazeal C (2008) Robots at home: Understanding long-term human-robot interaction. 2008 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS pp 3230–3235, DOI 10.1109/IROS.2008.4651113

Kraus W, Keteyian S (2007) Cardiac Rehabilitation. Humana Press, Totowa, NJ, USA, DOI 10.1007/978-1-59745-452-0

Lambert A, Norouzi N, Bruder G, Welch G (2020) A systematic review of ten years of research on human interaction with social robots. International Journal of Human–Computer Interaction 36(19):1804–1817, DOI 10.1080/10447318.2020.1801172

Lane GW, Noronha D, Rivera A, Craig K, Yee C, Mills B, Villanueva E (2016) Effectiveness of a social robot, "Paro," in a VA long-term care setting. Psychological services 13(3):292–299, DOI 10.1037/ser0000080

Lara JS, Casas J, Aguirre A, Munera M, Rincon-Roncancio M, Irfan B, Senft E, Belpaeme T, Cifuentes CA (2017) Human-robot sensor interface for cardiac rehabilitation. In: 2017 International Conference on Rehabilitation Robotics (ICORR), pp 1013–1018, DOI 10.1109/ICORR.2017.8009382

Lee KM, Jung Y, Kim J, Kim SR (2006) Are physically embodied social agents better than disembodied social agents?: The effects of physical embodiment, tactile interaction, and people's loneliness in human-robot interaction. International Journal of Human-Computer Studies 64(10):962–973, DOI 10.1016/j.ijhcs.2006.05.002

Leite I, Mascarenhas S, Martinho C, Prada R, Paiva A (2012) The Influence of Empathy in Human-Robot Relations. International Journal of Human-Computer Studies 71(3):250–260

Leite I, Martinho C, Paiva A (2013) Social robots for long-term interaction: A survey. International Journal of Social Robotics 5(2):291–308, DOI 10.1007/s12369-013-0178-y

Leite I, Castellano G, Pereira A, Martinho C, Paiva A (2014) Empathic robots for long-term interaction. Int J of Social Robotics 6:329–341, DOI 10.1007/s12369-014-0227-1

Leite I, Pereira A, Lehman JF (2017) Persistent memory in repeated child-robot conversations. In: Proceedings of the 2017 Conference on Interaction Design and Children, ACM, New York, NY, USA, pp 238–247

Lemaignan S, Garcia F, Jacq A, Dillenbourg P (2016) From real-time attention assessment to "with-me-ness" in human-robot interaction. In: 2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI), IEEE, pp 157–164, DOI 10.1109/HRI.2016.7451747

Li J (2015) The benefit of being physically present. Int J Hum-Comput Stud 77(C):23–37, DOI 10.1016/j.ijhcs.2015.01.001

Maclean N, Pound P (2000) A critical review of the concept of patient motivation in the literature on physical rehabilitation. Social Science & Medicine 50(4):495–506, DOI 10.1016/S0277-9536(99)00334-2

Martin CM, McConahay DR (1972) Maximal treadmill exercise electrocardiography. Correlations with coronary arteriography and cardiac hemodynamics. Circulation 46(5):956–962, DOI 10.1161/01.CIR.46.5.956

Matarić M, Scassellati B (2016) Socially assistive robotics. In: Siciliano B, Khatib O (eds) Springer Handbook of Robotics, Springer, Cham, pp 1973–1993, DOI 10.1007/978-3-319-32552-1_73

Matarić M, Eriksson J, Feil-Seifer D, Winstein CJ (2007) Socially assistive robotics for post-stroke rehabilitation. J NeuroEngineering Rehabil 4(5), DOI doi.org/10.1186/1743-0003-4-5

Matarić M, Tapus A, Winstein C, Eriksson J (2009) Socially assistive robotics for stroke and mild tbi rehabilitation. Stud Health Technol Inform 145:249–62, pMID: 19592798.

Matsumoto T, Satake S, Kanda T, Imai M, Hagita N (2012) Do you remember that shop? computational model of spatial memory for shopping companion robots. In: Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction, Association for Computing Machinery, New York, NY, USA, HRI '12, pp 447–454, DOI 10.1145/2157689.2157836

McColl D, Nejat G (2013) Meal-time with a socially assistive robot and older adults at a long-term care facility. J Hum-Robot Interact 2(1):152–171, DOI 10.5898/JHRI.2.1.McColl

McHugh ML (2012) Interrater reliability: the kappa statistic. Biochem Med (Zagreb) 22(3):276–282

Mutlu B, Osman S, Forlizzi J, Hodgins J, Kiesler S (2006) Task structure and user attributes as elements of human-robot interaction design. In: ROMAN 2006 - The 15th IEEE International Symposium on Robot and Human Interactive Communication, pp 74–79, DOI 10.1109/ROMAN.2006.314397

Mutlu B, Shiwa T, Kanda T, Ishiguro H, Hagita N (2009) Footing in human-robot conversations: how robots might shape participant roles using gaze cues. In: Proceedings of the 4th ACM/IEEE international conference on Human robot interaction, pp 61–68

Mwangi E, Barakova EI, Díaz-Boladeras M, Mallofré AC, Rauterberg M (2018) Directing attention through gaze hints improves task solving in human–humanoid interaction. International Journal of Social Robotics 10(3):343–355

O'Connor C, Joffe H (2020) Intercoder reliability in qualitative research: Debates and practical guidelines. International Journal of Qualitative Methods 19, DOI 10.1177/1609406919899220

Oertel C, Castellano G, Chetouani M, Nasir J, Obaid M, Pelachaud C, Peters C (2020) Engagement in human-agent interaction: An overview. Frontiers in Robotics and AI 7:92, DOI 10.3389/frobt.2020.00092

Okamura AM, Matarić MJ, Christensen HI (2010) Medical and health-care robotics. IEEE Robotics Automation Magazine 17(3):26–37

Pino M, Boulay M, Jouen F, Rigaud A (2015) "are we ready for robots that care for us?" attitudes and opinions of older adults toward socially assistive robots. Frontiers in Aging Neuroscience 7:141, DOI 10.3389/fnagi.2015.00141

Pinto MJ, Aguirre A, Cifuentes CA, Munera M (2020) Wearable Sensors for Monitoring Exercise and Fatigue Estimation in Rehabilitation. Eds. Internet of Medical Things: Paradigm of Wearable Devices. CRC Press Taylor & Francis Group

Prescott T, Caleb-Solly P (2017) Robotics in social care: A connected care ecosystem for independent living. UK-RAS Network Robotics & Autonomous Systems URL https://www.ukras.org/wp-content/uploads/2018/10/UK_RAS_wp_social_spread_low_res_ref.pdf

Reig S, Luria M, Forberger E, Won I, Steinfeld A, Forlizzi J, Zimmerman J (2021) Social robots in service contexts: Exploring the rewards and risks of personalization and re-embodiment. In: Designing Interactive Systems Conference 2021, Association for Computing Machinery, New York, NY, USA, DIS '21,

pp 1390–1402, DOI 10.1145/3461778.3462036

Richardson K, Coeckelbergh M, Wakunuma K, Billing E, Ziemke T, Gomez P, Vanderborght B, Belpaeme T (2018) Robot enhanced therapy for children with autism (DREAM): A social model of autism. IEEE Technology and Society Magazine 37(1):30–39

Riek LD (2017) Healthcare robotics. Commun ACM 60(11):68–78, DOI 10.1145/3127874

Roethlisberger F, Dickson W, Wright H, Company WE (1939) Management and the Worker: An Account of a Research Program Conducted by the Western Electric Company, Hawthorne Works, Chicago. Harvard University Press

Rossi S, Ferland F, Tapus A (2017) User profiling and behavioral adaptation for HRI: A survey. Pattern Recognition Letters 99:3–12, DOI https://doi.org/10.1016/j.patrec.2017.06.002, user Profiling and Behavior Adaptation for Human-Robot Interaction

Rudd EM, Jain LP, Scheirer WJ, Boult TE (2018) The extreme value machine. IEEE transactions on pattern analysis and machine intelligence 40(3):762–768

Ruhland K, Peters CE, Andrist S, Badler JB, Badler NI, Gleicher M, Mutlu B, McDonnell R (2015) A review of eye gaze in virtual agents, social robotics and hci: Behaviour generation, user interaction and perception. Computer graphics forum 34(6):299–326, DOI 10.1111/cgf.12603

Sabanovic S, Simmons R (2006) Robots in the Wild : Observing Human-Robot Social Interaction Outside the Lab. In: 9th IEEE International Workshop on Advanced Motion Control, 2006, pp 576–581, DOI 10.1109/AMC.2006.1631758

Sabelli AM, Kanda T, Hagita N (2011) A conversational robot in an elderly care center: An ethnographic study. In: Proceedings of the 6th International Conference on Human-Robot Interaction, Association for Computing Machinery, New York, NY, USA, HRI '11, pp 37–44, DOI 10.1145/1957656.1957669

Sanders EBN, Stappers PJ (2008) Co-creation and the new landscapes of design. CoDesign 4(1):5–18, DOI 10.1080/15710880701875068

Scane K, Alter D, Oh P, Brooks D (2012) Adherence to a cardiac rehabilitation home program model of care: a comparison to a well-established traditional on-site supervised program. Appl Physiol Nutr Metab 37(2):206–213, DOI 10.1139/h11-151

Scassellati B, Boccanfuso L, Huang CM, Mademtzi M, Qin M, Salomons N, Ventola P, Shic F (2018) Improving social skills in children with asd using a long-term, in-home social robot. Science Robotics 3(21), DOI 10.1126/scirobotics.aat7544

Schneider S, Kummert F (2021) Comparing robot and human guided personalization: Adaptive exercise robots are perceived as more competent and trustworthy. International Journal of Social Robotics 13:169–185, DOI 10.1007/s12369-020-00629-w

Shahsavari H, Shahriari M, Alimohammadi N (2012) Motivational factors of adherence to cardiac rehabilitation. Iranian journal of nursing and midwifery research 17(4):318–324, DOI 10.1002/14651858.CD001800.pub3

She J, Nakamura H, Makino K, Ohyama Y, Hashimoto H (2014) Selection of suitable maximum-heart-rate formulas for use with Karvonen formula to calculate exercise intensity. International Journal of Automation and Computing 12(1):62–69, DOI 10.1007/s11633-014-0824-3

Siegert RJ, Taylor WJ (2004) Theoretical aspects of goal-setting and motivation in rehabilitation. Disabil Rehabil 26(1):1–8, DOI 10.1080/09638280410001644932

Simms K, Myers C, Adams J, Hartman J, Lindsey C, Doler M, Suhr J (2007) Exercise tolerance testing in a cardiac rehabilitation setting: an exploratory study of its safety and practicality for exercise prescription and outcome data collection. Baylor University Medical Center 20(4):344–347

Suaya J, Stason W, Ades P, Normand S, Shepard D (2009) Cardiac rehabilitation and survival in older coronary patients. J Am Coll Cardiol 54(1):25–33, DOI 10.1016/j.jacc.2009.01.078

Süssenbach L, Riether N, Schneider S, Berger I, Kummert F, Lütkebohle I, Pitsch K (2014) A robot as fitness companion: Towards an interactive action-based motivation model. In: The 23rd IEEE International Symposium on Robot and Human Interactive Communication, pp 286–293, DOI 10.1109/ROMAN.2014.6926267

Tapus A (2009) Improving the quality of life of people with dementia through the use of socially assistive robots. In: 2009 Advanced Technologies for Enhanced Quality of Life, pp 81–86

Tapus A, Matarić MJ, Scassellati B (2007) Socially assistive robotics [grand challenges of robotics]. IEEE Robotics Automation Magazine 14(1):35–42

Tapus A, Ţăpuş C, Matarić MJ (2008) User-robot personality matching and assistive robot behavior adaptation for post-stroke rehabilitation therapy. Intelligent Service Robotics 1(2)

Tapus A, Tapus C, Mataric MJ (2009) The use of socially assistive robots in the design of intelligent cognitive therapies for people with dementia. In: 2009 IEEE International Conference on Rehabilitation

Robotics, pp 924–929, DOI 10.1109/ICORR.2009.5209501

Thompson D (2002) Stride Analysis. `https://ouhsc.edu/bserdac/dthompso/web/gait/knmatics/stride.htm`, [Online], Accessed 16-June-220.

Turk-Adawi K, Supervia M, Lopez-Jimenez F, Pesah E, Ding R, Britto RR, Bjarnason-Wehrens B, Derman W, Abreu A, Babu AS, Santos CA, Jong SK, Cuenza L, Yeo TJ, Scantlebury D, Andersen K, Gonzalez G, Giga V, Vulic D, Vataman E, Cliff J, Kouidi E, Yagci I, Kim C, Benaim B, Estany ER, Fernandez R, Radi B, Gaita D, Simon A, Chen SY, Roxburgh B, Martin JC, Maskhulia L, Burdiat G, Salmon R, Lomelí H, Sadeghi M, Sovova E, Hautala A, Tamuleviciute-Prasciene E, Ambrosetti M, Neubeck L, Asher E, Kemps H, Eysymontt Z, Farsky S, Hayward J, Prescott E, Dawkes S, Santibanez C, Zeballos C, Pavy B, Kiessling A, Sarrafzadegan N, Baer C, Thomas R, Hu D, Grace SL (2019) Cardiac Rehabilitation Availability and Density around the Globe. EClinicalMedicine 13:31–45, DOI 10.1016/j.eclinm.2019.06.007

Vandemeulebroucke T, Dierckx de Casterlé B, Gastmans C (2018) How do older adults experience and perceive socially assistive robots in aged care: a systematic review of qualitative evidence. Aging & Mental Health 22(2):149–167, DOI 10.1080/13607863.2017.1286455

Vasco V, Willems C, Chevalier P, De Tommaso D, Gower V, Gramatica F, Tikhanoff V, Pattacini U, Metta G, Wykowska A (2019) International conference on social robotics (icsr 2019). In: Train with Me: A Study Comparing a Socially Assistive Robot and a Virtual Agent for a Rehabilitation Task, Springer, DOI 10.1007/978-3-030-35888-4_42

Venkatesh V, Morris MG, Davis GB, Davis FD (2003) User acceptance of information technology: Toward a unified view. MIS Quarterly 27(3):425–478

Villacorta PJ (2017) The welchADF Package for Robust Hypothesis Testing in Unbalanced Multivariate Mixed Models with Heteroscedastic and Non-normal Data. The R Journal 9(2):309–328, DOI 10.32614/RJ-2017-049

Šabanović S (2010) Robots in society, society in robots. Int J of Soc Robotics 2:439–450, DOI 10.1007/s12369-010-0066-7

Wainer J, Feil-seifer DJ, Shell DA, Mataric MJ (2006) The role of physical embodiment in human-robot interaction. In: ROMAN 2006 - The 15th IEEE International Symposium on Robot and Human Interactive Communication, pp 117–122, DOI 10.1109/ROMAN.2006.314404

Welch BL (1938) The significance of the difference between two means when the population variances are unequal. Biometrika 29(3/4):350–362, DOI 10.2307/2332010

Welch BL (1951) On the comparison of several mean values: An alternative approach. Biometrika 38(3/4):330–336, DOI 10.2307/2332579

Winkle K, Caleb-Solly P, Turton A, Bremner P (2018) 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, Association for Computing Machinery, New York, NY, USA, HRI '18, pp 289–297, DOI 10.1145/3171221.3171273

World Health Organization (2011) World Report on Disability, vol 91. The World Bank

Zheng X, Glas DF, Minato T, Ishiguro H (2019) Four memory categories to support socially- appropriate conversations in long-term hri. In: 2019 ACM/IEEE International Conference on Human-Robot Interaction Personalization in Long-Term Human-Robot Interaction (PLOT-HRI19) workshop, Daegu, South Korea

Bahar Irfan is a Research and Development Associate at Evinoks Service Equipment Industry and Commerce Inc. (Turkey), working on the development of customisable software for industrial robots and smart buffets. Prior to that, she was a Research and Development Lab Associate at Disney Research Los Angeles (USA) in 2019, where she worked on dynamic emotional language adaptation in multiparty interactions. She received her PhD from the University of Plymouth (UK) and SoftBank Robotics Europe (France) as an Early-Stage Researcher in the Marie Skłodowska-Curie ITN project APRIL, focusing on multi-modal personalisation in long-term human-robot interaction within rehabilitation and customer service. She holds an MSc in computer engineering and a BSc in mechanical engineering from Boğaziçi University (Turkey). Her research focuses on creating personal robots that can continually learn and adapt to help people in their daily lives.

Nathalia Céspedes is a PhD student at the Queen Mary University of London (UK) in the Human Augmentation and Interactive Robotics team. She graduated as a Biomedical Engineer from the Colombian School of Engineering Julio Garavito and received a master degree in Electronics Engineering from

the same university. Her work is focused on building human-robot interfaces for rehabilitation scenarios and developing techniques to improve the interaction with users. Currently, she is researching social robots as tools for dementia treatment.

Jonathan Casas is a PhD student at the Department of Mechanical and Aerospace Engineering at Syracuse University, where he joined the Bionics System and Control lab as a research assistant in 2019. He received his BS and MS in 2016 and 2019, respectively, from the department of electrical engineering at the Colombian School of Engineering Julio Garavito, Bogotá, Colombia. In 2016, he joined the Department of Biomedical Engineering at the same institution as a research assistant, working in the areas of robotics and human-robot interaction for physical rehabilitation. His primary research interests are the areas of rehabilitation robotics and nonlinear and adaptive control for human-robot interaction in rehabilitation environments.

Emmanuel Senft is a Research Associate at the University of Wisconsin-Madison (USA) in the People and Robots Laboratory. He received his PhD from the University of Plymouth (UK) for his work in human-robot interaction. He researches systems to improve end-user programming of robots in a variety of applications, from therapy to manufacturing. He currently leads the technical development on a NASA University Leadership Initiative project focused on improving human-robot teaming in aviation manufacturing.

Luisa F. Gutiérrez is a Specialist in Physical Medicine and Rehabilitation. She has experience in managing cardiac rehabilitation and cardiovascular prevention. She is the leader of the program of neurodegenerative diseases of childhood in the rehabilitation of the Fundación Cardioinfantil, and she is also director of the Cardiac rehabilitation program of the Colsubsidio network. She has training in non-oncological and neuropathic chronic pain, private consultation. She has competencies in clinical research and the academic training of health professionals at the Universidad de La Sabana and the University of Rosario.

Mónica Rincon-Roncancio is a Medical Specialist in Physical Medicine and Rehabilitation and Bioethics. She holds a PhD in Bioethics, a Master in Bioethics, Philosophy and Physiology. She is a professor at Universidad del Rosario, Universidad de La Sabana, Colombian School of Engineering Julio Garavito, and National of Colombia. She is the Head of the Rehabilitation Service and President of the Hospital Ethics Committee at Fundación Cardioinfantil-Instituto de Cardiología. She is also the President of the Cardiac Prevention and Rehabilitation Chapter of the Colombian Association of Physical Medicine and Rehabilitation.

Carlos A. Cifuentes is a Professor with the School of Engineering, Science and Technology at Universidad del Rosario (Colombia) and Visiting Professor at EPF Graduate School of Engineering (France). In 2017, his work was lauded as one of five history-changing ideas in Latin America by History Channel. In 2020, he was elevated to the grade of IEEE Senior Member, and in 2022, he was appointed as Associate Editor for IEEE Robotics and Automation Magazine. His research on Human-Robot Interaction is world-leading, and his team's research has conclusively shown the benefits of using robots in a healthcare context. Funding agencies such as the Royal Academy of Engineering (RAEng, UK), Ibero-American Programme on Science and Technology for Development (CYTED), and Ministry of Science, Technology, and Innovation (Minciencias, Colombia) present his research as funding success.

Tony Belpaeme is a Professor at Ghent University and Visiting Professor in Robotics and Cognitive Systems at the University of Plymouth, UK. At Ghent, he is a member of IDLab - imec. He received his PhD in Computer Science from the Vrije Universiteit Brussel (VUB) and currently leads a team studying cognitive robotics and human-robot interaction. Starting from the premise that intelligence is rooted in social interaction, Belpaeme and his research team try to further the science and technology behind artificial intelligence and social human-robot interaction. This results in a spectrum of results, from theoretical insights to practical applications.

Marcela Múnera is an Assistant Professor of Biomedical Engineering at the Escuela Colombiana de Ingenieria Julio Garavito. Currently, she works at the Center for Biomechatronic in the development of projects in rehabilitation, the objective assessment of robotic devices and the assessment of the human response in different scenarios. She received her PhD in Mechanics and Biomechanics from Université de Reims Champagne Ardenne, thanks to a FEDER, Region Champagne Ardenne- Doctoral Grant. She

graduated as a Bioengineer from Universidad de Antioquia and from the Ecole Nationale de Metz with a Masters in Mechanics and Materials. During her PhD and-after, as a Lecturer at Université de Reims Champagne Ardenne (France), she worked on industrial research projects in biomechanical assessment related to sports performance and injury prevention. Her research interests are focused on biomechanics, movement analysis, and assessment in rehabilitation and sports.