

Emotion-sensitive voice-casting care robot in rehabilitation using real-time sensing and analysis of biometric information

Peeraya Sripian^{a,*}, Muhammad Nur Adilin Mohd Anuardi^a, Teppei Ito^a, Yoshito Tobe^b and Midori Sugaya^a

^a *Shibaura Institute of Technology, 3 Chome-7-5 Toyosu, Koto City, Tokyo 135-8548, Japan*

E-mails: peeraya@shibaura-it.ac.jp, nb17106@shibaura-it.ac.jp, doly@shibaura-it.ac.jp

^b *Aoyama Gakuin University, 4 Chome-4-25 Shibuya, Shibuya City, Tokyo 150-8366, Japan*

Received 22 April 2020

Accepted 19 July 2021

Abstract. An important part of nursing care is the physiotherapist's physical exercise recovery training (for instance, walking), which is aimed at restoring athletic ability, known as rehabilitation (rehab). In rehab, the big problem is that it is difficult to maintain motivation. Therapies using robots have been proposed, such as animalistic robots that have positive psychological, physiological, and social effects on the patient. These also have an important effect in reducing the on-site human workload. However, the problem with these robots is that they do not actually understand what emotions the user is currently feeling. Some studies have been successful in estimating a person's emotions. As for non-cognitive approaches, there is an emotional estimation of non-verbal information. In this study, we focus on the characteristics of real-time sensing of emotion through heart rates – unconsciously evaluating what a person experiences – and applying it to select the appropriate turn of phrase by a voice-casting robot. We developed a robot to achieve this purpose. As a result, we were able to confirm the effectiveness of a real-time emotion-sensitive voice-casting robot that performs supportive actions significantly different from non-voice casting robots.

Keywords: Voice casting robot, rehabilitation, emotion analysis, emotion estimation, biometric information

1. Introduction

In recent years, the proportion of older people has increased due to the trend of declining birth rates and people living longer. The super-aged society is progressing in such a way that by 2025 in Japan, the elderly population will have reached approximately 30% of the total population [1]. Along with the increase in seniors aged 65 years and over, the increase in the burden of nursing care is becoming very significant. As part of nursing care, the physiotherapist's physical exercise recovery training (for walking, etc.), which is aimed at restoring the athletic ability to people whose physical function has declined, is called rehabilitation (rehab). In rehab, the burden on the caregiver is enormous. Meanwhile, for the practitioners, there is the problem that it is difficult to maintain motivation as the work is troublesome, sometimes does not have clear best practice, and seems to feel that there is no clear achievement [2].

* Corresponding author. E-mail: peeraya@shibaura-it.ac.jp.

Robot therapy has been introduced in medical and welfare facilities, to tackle problem such as caregiver shortage [3,4]. Animal-type robots were proved that they could provide patients such as elderly people with good psychological and social effect, and reduce the onsite workload of the caregiver. *Paro*, the therapeutic robot developed by AIST, has been used in environment such as hospitals or nursing care facility where live animals cannot be introduced due to possible harm and hygiene concern. *Paro* is equipped five perception sensors; tactile, light, audition, temperature and posture sensors. It can learn names and behaviors of the people by combining these sensors. Gradually, *Paro* can build relationship with the owners through its interaction and the owner is expected to interpret as if it had feelings like living things. However, this therapeutic robot's stimulus-response learning was a rule-based model. It could not really understand what emotions the owner is feeling at that time. In addition, *Paro* can only make the sound like a seal's bark, in situations with a specific purpose such as rehab, no consideration is given to how exactly to promote the owner's motivation.

Verbal interaction is an important mechanism in establishing social interaction. Verbal encouragement is a common technique used by therapists to encourage patients to perform to their maximum potential during rehabilitation sessions. It was found that verbal encouragement significantly increased performance of participants in a motor endurance task [5]. Carrillo et al. developed the humanoid social robot *NAO* as a social robot for pediatric rehabilitation aid [6]. *NAO* was designed that it would perform the exercise in front of the child and provide verbal encouragement during the rehabilitation session. It also provided enticements upon the completion of exercise sets. It was found slight improvement in patient's motor skill after their training with the robot compared to conventional treatment [7].

The motivation of this research is to reduce the workload of the caregiver by proposing a robot equipped with verbal encouragement for the patient's rehabilitation. In this work, the robot is called a voice-casting robot. We propose the method that automatically select encouragement phrases according to the patient's condition. According to Kishiba et al. [8], the caregiver's message should not be instructive nor imperative but should try to be supportive. Though, many literatures suggest that personality of the patients also effect desired caregiver message in social agents [9,10]. An imperative voice-casting would concern mainly on the person calling (casting their voice), a supportive voice-casting would consider the condition and emotion of the listener, or the rehabilitation patient instead.

In order to estimate a person's condition and emotion, many studies have been done. For example, Takeuchi et al. [11] proposed a cognitive approach that uses natural language processing to estimate emotion through the data analysis of dialogue. However, our work focuses on elderly patients. Elderly patients are believed to be more skilled to the social expression of emotion than younger adults, so they are skilled in inhibiting expression [12], and more likely to express positive emotional expression [13]. This makes the estimation of condition and emotion inherently difficult. On the other hand, there is a non-cognitive approach that could estimate emotion from non-verbal information such as people's gaze, head position, facial and mouth expression [14], or even body posture [15]. However, there are many challenges to recognizing emotion from these kinds of expressions. In a laboratory, emotions can be acted or primed, but in the real world, the onset of emotion is often unclear, which could make the data difficult to window, causing noises in the windowing the data [16]. Also, every subject has their own experience of emotion, which could lead to a misleading self-description of emotions that are difficult to classify [17]. Recently, machine learning or deep learning techniques made it possible to learn such subjectively changing expressions and estimate emotion based on the learning data [18]. It is still not sufficient to associate the physiological state with cognition and its expression.

In the fields of affective computing [19] and personality computing [20], a method has been proposed for estimating the state of a person using the mechanically collected characteristics of a person, usually collaborated by subjective questionnaires. As mentioned above, however, it is difficult to correctly identify one's state in a real-time [20]. Therefore, it is necessary to find a suitable method that can objectively estimate one's state (human condition) and emotion.

To identify one's state, biometric information for estimating emotion has been proposed in recent years. Biometric information such as electroencephalogram (EEG), electrocardiogram (ECG) is widely studied and used in the classification of emotion or state of mind [19]. However, it is difficult to practically measure EEG using electroencephalographs in situations such as rehabilitation in rehab sites or clinics. In contrast, the ECG measurement of heart rate is considered to be suitable for human state estimation because data can be easily collected. Hoshishiba et

al. [21], analyzed the heart rate variability using regular music stimulation and pointed out that music has an influence on the parasympathetic nerve activity and that there is a positive correlation between the change in tempo and the respiratory components responsible for heart rate variability [22]. As a result, heart rate variability can unconsciously reveal what a person has experienced through the autonomic nervous system when responding to external stimuli.

Based on the above findings, we will use the heart rate in this work as the biometric information for evaluating the external voice stimulation. In this research, we propose the development of a rehabilitation-aiding voice-casting robot that can estimate human emotion through real-time sensing of biometric information. During the rehabilitation, the robot will automatically select suitable supportive phrases from the database for voice-casting to encourage the patient. The evaluation of the developed system is done through questionnaire and biometric information evaluation. The evaluation result shows that our proposed voice-casting based on emotion estimation performs effectively.

2. Related works

A therapeutic baby seal robot named “Paro” was developed by Dr. Takanori Shibata of Japan’s National Institute of Advanced Industrial Science and Technology (AIST) and has been widely used around the world [2,3,23]. A therapeutic robot has the advantage that it can be carried out more easily than live-animal therapy because the robot is safer in terms of sanitation and also behavior. The baby seal-like robot Paro [4] made possible with tactile, visual, auditory, and balance senses using its internal sensors, and by combining these data, it can learn people’s names and actions. With Paro, it is possible to gradually build up the relationships between Paro and its owners through interactions, and the owners are expected to interpret it as if Paro had feelings. It is recognized that animal therapy has mainly (1) psychological effects: an increase in the number of smiles and motivation, mitigation of “depression,” etc., (2) physiological effects: a decrease in stress, blood pressure, etc., and (3) social effect: an increase in communication, etc. Shibata et al.’s research aimed to verify empirically whether or not the effects of (1)–(3) exist in robot therapy. Shibata et al., have performed experiments using Paro robots and have shown that they have been effective [2]. However, Paro is mainly healing patients with its animal-like voice and soothing appearance. It does not, for example, improve the specific motivation of patients in rehab by speaking a natural language.

One of the factors that can enhance the effectiveness of rehabilitation therapy is the level of ambition of the rehab patient themselves [11]. Motivation is the driving force of action and the necessity of all kinds of activities. It is an important factor in the success of rehabilitation and frequently it is used as a determinant of the rehabilitation outcome [24]. Thus, if a patient’s motivation is decreased, it could mainly effect the rehabilitation therapy. Kibishi et al. [8] tried to tackle this problem by investigating the degree of patients’ motivation based on many phrases used by physiotherapists, occupational therapists, physicians, etc., during the rehabilitation therapy at a rehab hospital [4]. The investigation was done through a questionnaire, based on eight rehabilitation scenarios: rehab starting time, rehab ending time, before the practice of walking, before the practice of using toilet, before the practice of getting dressed, waiting during rehab, describing self-pain level, when the patient was successfully able to do something that they could not do before. The staff participants were asked to classify spoken phrases that were commonly used in those scenes as appropriate or inappropriate. Then, the rehabilitation patients were asked to rate all the phrases into three levels: increase motivation, neutral, or decrease motivation. As a result, the content of the phrases affects the patient’s rated motivation level. Also, they found that advisory/supportive phrases rather than instructive/imperative phrases were an essential factor that affects the motivation in rehabilitation. Although these results are based on human-human communication, it is possible to apply them to the voice-casting robot in our work.

3. Voice-casting robot

3.1. The objective of this work

The purpose of this research is to reduce the workload of the human caregiver, by using a voice-casting robot in rehabilitation therapy. Our proposed voice-casting robot will estimate the patient’s emotion using real-time sensing of biometric information and select suitable supportive phrases for voice casting in order to increase the patient’s motivation during rehabilitation therapy.

3.2. Proposed methods

As a preliminary system, we focus on walking rehabilitation therapy, because walking is highly necessary for daily life, and home rehabilitation can be performed. Unlike Kibishi et al. [8], who used individual staff's voices for voice casting, we use a robot for voice casting. We selected some spoken phrases that were empirically evaluated as effectively motivating the patients from Kibishi et al.'s finding [8] and store in the robot's phrase database for selective voice-casting during the rehabilitation therapy. The voice casting is performed automatically by selecting suitable natural language phrases from the phrase database according to the patient's state. The patient state (emotion) will be interpreted into two types: positive or negative. If the patient is in a positive state, it can be assumed that motivation can be increased.

There are three parts to this paper:

- (1) Design and implementation of the voice-casting robot
- (2) Real-time measurement of the patient's emotion while doing rehabilitation. The emotion estimation result is classified as a positive or negative feeling.
- (3) The robot will encourage positive emotion by making voice-casting based on the judgment of sensing and analysis of the emotion at that time.

In (1), the design and implementation of the voice-casting robot, we design a robot that speaks while accompanying a walker during rehabilitation. The robot is implemented using commercially available hardware and micro-computers. The robot will be equipped with a built-in speaker so that it can cast supportive phrases to the patient while accompanying the patient in the rehabilitation. To achieve appropriate voice-casting, the robot should keep a certain distance from the patient. The rehabilitation scenario used in the evaluation is a walking training scenario, as it usually requires the human caregiver to guide while walking. Biometric information will be used to objectively measure the positive/negative emotion throughout the rehabilitation to realize (2) and (3).

3.3. Measurement and analysis of emotion using heart rate sensors

Biometric information is the activity data of the body, unconsciously generated by humans, such as heart rate, blood pressure, and brain waves. It is generally used in the investigation of the health condition of a person. In recent years, engineering applications have been developed as a means to evaluate human emotions and psychology [25]. Unconscious activities are controlled by the autonomic nervous system (ANS), and we perceive them as emotion. The ANS transmits stimuli from the outside world to the brain as sound and vision, and when they are interpreted as stimuli, a reaction of the ANS is generated as a response. The index related to the ANS as described measures the mental and physical activity, tension and excitement, arousal and relaxation level, mental burden, and so forth [26].

Heart rate variability (HRV) analysis has become a popular noninvasive tool for assessing the activities of the autonomic nervous system [27]. It was suggested that HRV could be an objective tool to evaluate emotional responses [28]. Also, there is a report that shows a correlation between a subject's emotional state and HRV [29]. HRV shows the variation of the R wave to R wave interval of heartbeat (RRI), that is, the fluctuation of the instantaneous heart rates. Since the fluctuation of the R-R intervals is large at rest, it can help suggest times of stress and tension. By measuring RRI (R-R Intervals), we can evaluate whether the person is relaxed (experiencing comfort), or tense/experiencing stress (discomfort) [25,26,30–32].

3.4. Comfort/discomfort judgment by pulse

pNNx is a method for easily measuring RRI. pNNx indicates the ratio of consecutive RRI differences lasting x milliseconds and more. The value of x is considered to have a strong negative correlation with "stress" or "discomfort" at $20 \leq x \leq 50$, and with "comfort" at larger values [22]. Based on this fact, we used $p = 50$ (pNN50) to judge comfort/discomfort by whether each of the last 30 consecutive RRI (R-R Interval) differences last over 50 ms. pNN50 has better statistical properties and has been used in several studies to measure physiological stress and emotional state classification for robotics application [33–36].

Therefore, in this study, heartbeat and pulses were assumed to be equal, and pulses were used for calculation of pNN50. Pulses can help observe how the heart works and delivers blood to the whole body through blood vessels

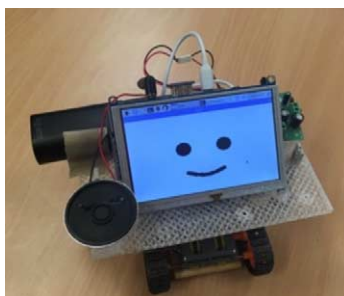


Fig. 1. Appearance of our proposed voice-casting robot used in preliminary experiment.

in all parts of the body. Therefore, despite a slight time delay, we can obtain waveforms substantially synchronized with the heartbeat. Especially in the case of pNN50, since only R waves can be detected, pulses were considered to be sufficient.

For pulse observation, the pulse sensors by Sparkfun Electronics™ SEN-11574 were used. They can target the capillary vessels close to the epidermis and measure the amount of hemoglobin in blood by gauging the strength of the visible light reflected by the green light-emitting photodiode. In addition to the ability to attach to fingertips, this sensor was used because data can be easily obtained with a microcomputer.

3.5. Prototype implementation

A small computer (Raspberry Pi 3 Model B) was built into the body, and a 5-inch liquid crystal display (LCD) was used for the face display. The robot was also equipped with a dynamic speaker for voice casting, one mono amplifier, and a crawler for movement. The crawler used the assembled type for the model. For the power supply, we used two AA batteries to drive the crawler, two 006P types (9 V) batteries for Arduino and speaker operation, and a commercially available lithium-ion battery (5 V) for the Raspberry Pi and LCD. For the screen display, the image used as the facial expressions displayed on the LCD was designed with reference to the expression shown on NAO, developed by Softbank Robotics in 2004 [37]. It is a very simple face with only two eyes and one mouth. For the speech model of the robot, we use `open-jtalk--mecab-naist-jdic` package for dictionary and `hts-voice-nitech-jp-atr503-m001` package from OpenJTalk, which is an open source software that inputs natural language in a string format and consistently performs speech synthesis. For sound synthesis, we used the `aplay` command included in Advanced Linux Sound Architecture (ALSA), which is a kernel component of Linux for audio playback. In order to execute console commands from Python, we used the `subprocess` module, a package for Python.

3.6. Preliminary experiments

We conducted a preliminary experiment ($N = 1$) using the developed voice-casting robot. A commercially available walking assistant robot was used as a walking assistant tool, and our proposed voice-casting robot is placed in front of the walking assistant tool. The voice-casting robot used in this experiment is shown in Fig. 1.

For the procedure, first, a 2-minute rest period is scheduled for baseline pulse measurement. Then, the participant was instructed to walk on a straight path of 8 m using a walking assistant robot while hearing supportive phrases casting by the voice-casting robot that is positioned a certain distance in front of the participant (see Fig. 2). Phrases used in voice casting are as follows:

- (1) “Let’s do our best today,” at the beginning of walking practice
- (2) “A little more”
- (3) “Let’s keep up the good work,” at the end of walking practice

The experiment ended 10 seconds after the casting phase (3). Scenes from the actual experiment are shown in Fig. 2. During the entire rehabilitation, the voice-casting robot is always located about 1 m ahead of the walking assistant robot to accompany the participant.



Fig. 2. Photos from the experiment (left: diagonal rear view, right: side view).

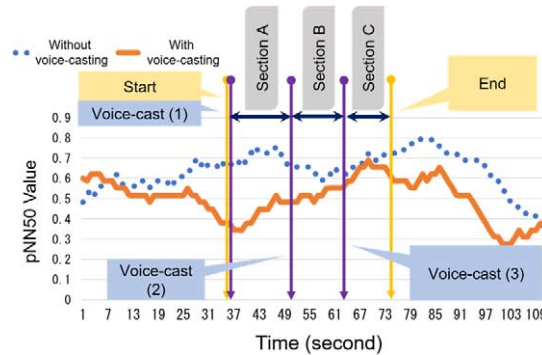


Fig. 3. Changes in the pNN50 value of the participant during the entire experiment.

pNN 50 was used for emotion estimation. We measure the rate of change of pNN50 value throughout the interval. The rate of change (%) of pNN50, Δ_{pNN50} , can be calculated by the following equation;

$$\Delta_{\text{pNN50}} = 100 \left(\frac{E_e - E_s}{t_e - t_s} \right),$$

where E_e is the emotional value of pNN50 at the end of the interval, E_s is the emotional value of pNN50 at the beginning of the interval, t_e is the time at the end of the interval and t_s is the time in the beginning at the interval.

This calculation's results will show the trend of the status of changing emotion. If $\Delta_{\text{pNN50}} \geq 0$, it will be changing to the positive state. On the contrary, if $\Delta_{\text{pNN50}} < 0$, the state of the emotion would be changing to the negative state. The changes in the pNN50 of the participant during the entire experiment is shown in Fig. 3. We stipulate that the time period between voice casting phrase (1) and just before voice casting phrase (2) as Section A, that between voice casting phrase (2) and just before voice casting phrase (3) as Section B, and that between voice casting phrase (3) and the end of the experiment as Section C. From Fig. 3, the average changes rate of pNN50 values were refer from the start time until the end time of each section. The pNN50 value changes for consecutive voice castings was +1.88% in section A, +1.18% in section B, and +1.63% in section C. In the case of no voice casting in the same sections, it was +0.44%, -3.93%, and -1.43%, respectively. As a result of voice casting, we noticed the increase in the value indicating comfort. However, it was also found that the absolute value indicating comfort was high even when there are no voice-casting, therefore, further investigation is required. In fact, this result only preliminary and implied for only one subject which aim to validate the proposed method. Thus, the appropriate approach will be considered for improvement in the next step.

3.7. Issue and improvement

From the participant interview after the preliminary experiment, we found that the robot's motor sound was perceived as an annoying noise, and unstable motion and movement of the walking assistant robot were the causes of the feeling of discomfort. Based on these comments, we believe that it is possible to improve the design and implementation of the robot to better suit the rehabilitation as follows;



Fig. 4. Appearance of our proposed voice-casting robot.

- (1) Improvement of the movement mechanism
- (2) Improvement of the robot's body design
- (3) Improvement of the amplifier
- (4) The addition of a wireless communication function.

Improvement of movement mechanism can be done by replacing the assembly-type setup with the ready-made sealed-type setup for the transmission and the errors in the left and right drivers of the crawler may be reduced and the voice-casting robot may be able to walk in a straighter line.

In the preliminary experiment, the facial expressions implemented a blinking effect because there was feedback saying that “eyes that blink appear to be more life-like.” The facial images were passed to OpenCV functions and displayed on full screen on a 5-inch LCD, and images were switched and displayed every 2 to 5 seconds as determined by randomly generated numbers.

Moreover, a motor driver IC was added so that it could be controlled directly from Raspberry Pi's GPIO pin without going through Arduino. Furthermore, by using Raspberry Pi's GPIO control library pigpio, speed adjustment, and reversal by software PWM control can be performed flexibly. For improvement of the robot design, we consider that the driving unit is exposed to achieve different impressions from the participant. The voice-casting robot must be stable during rehabilitation therapy, so we use a white acrylic plate, which yields the feeling of a smooth surface and also high stability.

For the improvement of the amplifier, the monaural amplifier which yields unstable noise was changed to the commercial stereo amplifier. With the stereo amplifier, the voice is easier to hear with a suitable level of volume.

Finally, we add the wireless communication function for the communication between the pulse sensor and the robot. The system in the preliminary experiment did not use wireless communication so the wire could be accidentally stepped on by the participant during the rehabilitation therapy, leading injury.

The appearance of the voice-casting robot whose design and mounting have been improved is shown in Fig. 4, and the design block diagram is shown in Fig. 5. By using the wireless communication function, a seamless acquisition of pulse data can be achieved. Pulse data are acquired from the pulse sensor and sent wirelessly to the microcomputer, which calculates and analyzes emotional state based on the pNN50 value.

3.8. Phrases used in voice casting

Our voice-casting robot will select phrases differently based on its judgment of comfort/discomfort. Abreu et al.'s threshold was used for the comfort/discomfort emotion judgment based on the pNN50 value [38]. In their work, they found that 0.23 was appropriate as a threshold to indicate comfort. We selected phrases that were found as appropriate voice casting phrases for rehabilitation therapy [8]. The voice casting phrases selected according to the judgment of comfort/discomfort are shown in Table 1.

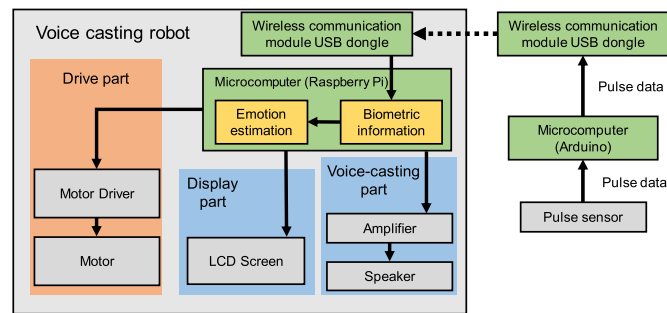


Fig. 5. A block diagram showing the design of our voice-casting robot.

Table 1

Voice casting phases selected according to the judgement of comfort/discomfort by analyzing the level of pNN50

	Discomfort (pNN50 < 0.23)	Comfort (pNN50 ≥ 0.23)
Begin	Thank you for using	Today let's do our best!
Middle of the walk	Do you want to have a break? You do not have to rush	You look happy, I am also happy Only a little bit more
Later	Let's try gradually like this	See you tomorrow

4. Evaluation of voice-casting robots

4.1. Overview

Based on the issues found in the preliminary experiment, we developed an improved version of the voice-casting robot that makes appropriate voice casting based on the comfort/discomfort judgment. The following experiment was conducted to evaluate our updated version of the voice-casting robot.

4.2. Subject

Nine participants (two females and seven males, with an age range 20–40 years old) took part in this experiment, having given consent. Before the experiment, the participants were informed of the reason for the development of the voice-casting robot and briefly instructed on the operational method.

4.3. Experiment method

Four patterns of robot movement are used in the experiment. Table 2 shows the combination of robot movements for each pattern. VC stands for voice casting, with or without judgment using biometric information. 'Follow Up' is the supportive behavior added in this updated version of the robot, that is, once the robot goes forward, it would come back to check on the patient. We added this behavior in the updated version as a result of the discussion with physiotherapists that it would be beneficial using follow up behavior.

Follow-up was added to the supportive behavior, as a result of the discussion with the expert researcher in rehabilitation therapy, Prof. Fujimoto from the University of Human Arts and Sciences, Japan. He suggested to consider a patient's social background and the state of the patient in the meantime during the rehabilitation therapy. The voice casting is a supportive behavior but appeals only to the auditory sense. Therefore, we added a follow-up as a supportive visual behavior in our updated system. This follow-up behavior is performed by the robot as it moves back to check on the patient and performs a caring action for the patient, as shown in Fig. 6.

In order to eliminate bias as much as possible, two participants are deliberately assigned to the pattern I, while the remaining patterns II to IV were rearranged to the rest of the participants with random numbers generated by Mersenne Twister.

Table 2
The combination of actions for each robot movement pattern

Pattern of Robot Movement	Movement	
	VC (w/biometric information)	Follow up
I. Move only	X	X
II. Move + Voice Casting	O	X
III. Move + Follow Up	X	O
IV. Move + Voice Casting + Follow Up	O	O

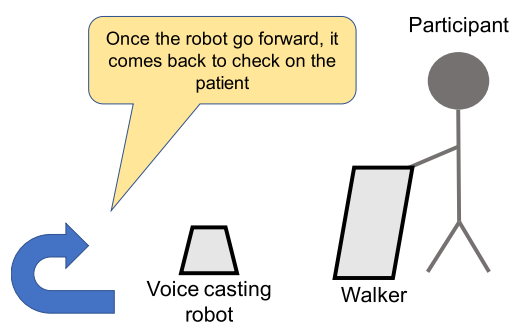


Fig. 6. Follow-up scenario.

4.4. Experimental procedure

According to the comment received from participants in preliminary experiment, the walking path was adjusted to a length of 6 m, and a walking machine (the walker shown in Fig. 6) frequently used by elderly people in need of walking rehabilitation was used. The experimental procedure is as follows:

0. A pulse sensor is attached to the participant for measurement throughout the experiment.
1. 1-minute of resting time was given at the beginning for baseline measurement. During the resting time, the participant was instructed to look at the outside scenery through the window.
2. The participant starts walking using the walker along the path. During this period, the robot made voice castings based on the judgement of participant's state. The voice casting phrases were selected as shown in Table 1.
3. Once the participant walked through the walking path, they were required to answer the subjective evaluation questionnaire.

4.5. Subjective evaluation using a questionnaire

The subjective questionnaire is a self-evaluation of motivation level. There are six levels of motivation: very motivated, motivated, a little motivated, a little unmotivated, unmotivated, and very unmotivated.

4.6. Measurement of biometric information

In order to acquire the direct response at the time of a voice casting, we measured the comfort/discomfort using pNN50, similar to the preliminary experiment. Since the purpose of this experiment is to compare different pattern combinations of the voice-casting robot, we decided not to use the gradient but the average rate of change and measure it for the evaluation. The following formula shows the calculation method of the average rate of change r :

$$r = 100 \frac{\mu_A - \mu_R}{\mu_R},$$

where μ_A is the average pNN50 value while active (moving) and μ_R is the average pNN50 value while resting.

Table 3
Combinations and differences in movement for the *t*-test

Combination		Movement	
Group 1	Group 2	Point of similarity	Point of difference
I	II	None	Voice Casting
	III	None	Follow Up
	IV	None	Voice Casting, Follow Up
II	IV	Voice Casting	Follow Up
III	IV	Follow up	Voice Casting

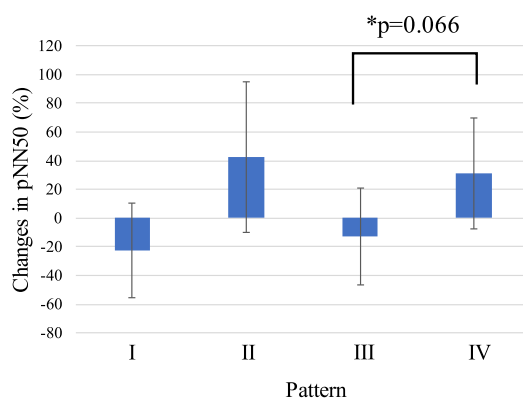


Fig. 7. The comparison of biometric information (pNN50) for all pattern with significant difference tested by pair *t*-test.

4.7. Results

4.7.1. *T*-test comparison of the average rate of change in pNN50 for each pattern

For each of the patterns I to IV, the paired *t*-test ($\alpha = 0.05$) was conducted to compare each group with attention paid to the difference in behavior. The combinations used in the test are shown in Table 3. A paired *t*-test showed a significant effect between the biometric information recorded from pattern III and pattern IV ($t(8) = -2.129$, $p < 0.066$). The comparison of all pNN50 results with the significant effect is shown in Fig. 7.

4.7.2. *T*-test comparison of subjective evaluation scores for each pattern

A paired *t*-test ($\alpha = 0.05$) was performed to find if there was any significant difference between subjective evaluation score. The comparison of pairs is shown in Table 3. From the paired *t*-test results, we found the significant effect of pattern I vs. II ($t(8) = -3.277$, $p < 0.05$), pattern I vs. IV ($t(8) = -6.008$, $p < 0.001$), and pattern III vs. IV ($t(8) = -7.0$, $p < 0.001$) accordingly. The comparison of subjective evaluation scores with *t*-test results is shown in Fig. 8.

4.8. Discussion

The comparison of biometric information shows a significant difference only between patterns III and IV, where there is similarity in follow-ups but voice-casting is different. This may suggest that voice-casting could influence the comfort/discomfort state in addition to follow-up. However, no significant difference is found for patterns I and II, where voice casting is also different. We assume that the number of participants was small, hence a large standard deviation could be an important factor.

For the comparison of subjective evaluation scores, a significant difference was observed between patterns I and IV, and between patterns III and IV at the 1% level. First, between patterns I and IV, the mean values were 2.1 and 4.2, respectively. Since this is an interval scale variable, it indicates motivation improvement as the score of 2.1 is categorized as “a little unmotivated,” while 4.2 is categorized as “motivated.” It can be implied that the subjective

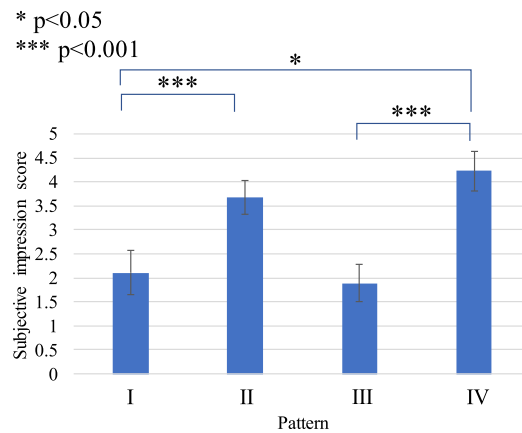


Fig. 8. The comparison of subjective impression score with significant difference tested by pair *t*-test.

evaluation of motivation increases drastically when the combination of follow-up and voice casting is used rather than the robot move only. Next, between patterns III and IV, the mean values were 1.9 and 4.2 respectively. Similarly, it can be implied that the motivation is increased when the combination of follow-up and voice casting is used. For this comparison (III and IV), the significant effect in the difference of biometric information is also observed. Therefore, it is possible to conclude that a combination of voice casting and follow-up features of the robot effects a participant's motivation.

From the participant's comments after the experiment, the participants tend to have a negative impression, such as frustration, for the pattern I, as it is presumed that participants are being forced semi-compulsively, instead of working ambitiously with a sense of work. For pattern II, the participants gave positive comments regarding voice-casting as they could feel that the robot was working together with the participant on the rehabilitation. However, a few participants said that they were surprised by the robot's voice, so the way a robot speaks could also be improved in the future. Meanwhile, comments regarding pattern III revealed many negative emotions, such as intimidation, embarrassment, and indignation. This could be caused by the follow-up feature of the robot, in which the participants could not understand the robot's intention. Hence the feeling of being puzzled or irritation occurred. Finally, the comments for pattern IV reflect many positive feelings such as feeling cared for and feeling secure. In addition, we also received comments that the robot should be made simpler like a pet to show affection and have a soothing effect.

5. Voice casting robot with an awareness of the arousal status

5.1. Motivation

The result of the previous experiment showed that the rehabilitation robot with voice-casting is more effective than no voice-casting. However, it is necessary to further investigate which condition is suitable for voice-casting. The robot in the previous experiment uses different voice phrases according to the subjects' valence level. However, the combination of valence and arousal values may indicates a totally specific emotion based on the valence/arousal model as in Russel's circumflex model of affect [39]. The combination of low valence (unpleasant) and high arousal (activated), could be considered as angry, and also irritation (red area in Fig. 9). On the other hand, low valence (unpleasant) with low arousal (deactivated) could be interpreted as sad or bored (blue area in Fig. 9). Casting a voice on the patients with the specific emotions such as irritated patients or bored patients would yield a different result. The previous experiment did not consider the combination of valence and arousal values. Therefore, in this experiment, we consider these features to compare the effect of voice casting on a different emotion.

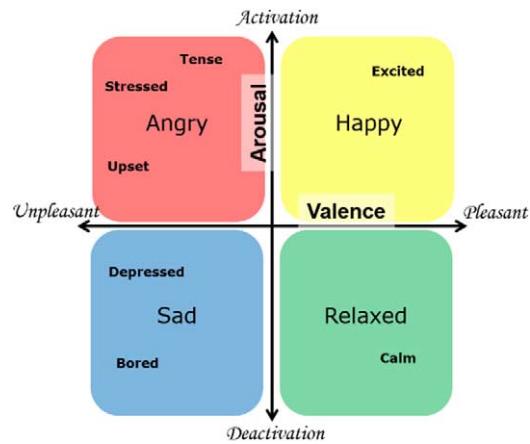


Fig. 9. Russell's circumflex model of affect [26] The horizontal-vertical axes are valence-arousal, accordingly.

5.2. Improvement of the system

To calculate arousal value, the attention and meditation values obtained from EEG of Neurosky's Mindwave mobile are used. The experiment shows that attention value correlates with the β wave, and meditation value correlates with the α wave of the EEG frequency band. Therefore, arousal can be calculated from attention and meditation values. For pNN50, we calculated from the RRI obtained from the pulse sensor.

The improved system contains a module attached to a walker that would transmit data obtained from sensors attached to the patient wirelessly. A microcomputer (Arduino) and an electroencephalograph that performs pulse data acquisition and data transmission using serial communication are used in the module.

The wireless communication between Arduino and the Raspberry Pi for transmission of RRI data obtained from the pulse sensor attached to a patient is done using XBee. For brainwaves transmission, the EEG from Neurosky's Mindwave mobile provided Bluetooth wireless transmission to measure the attention and meditation value. Arousal value is calculated from the difference in attention and meditation value obtained from EEG of Neurosky's Mindwave Mobile [40]. The estimated emotion is sleepy when the difference of attention and meditation is zero or less. Meanwhile, valence value is calculated from the RR wave obtained from pulse sensor similar to previous experiment.

In this experiment, in order to prevent the effect of each voice call from being unclear, the previous voice call was performed so that 4 to 5 voice calls were performed during a 10 m walk. For 20 seconds from the beginning, no call was made even below the threshold.

5.3. Experiment

The following experiment was conducted to compare the effectiveness of voice-casting based on arousal and based on valence. In this experiment, we set the timing for all voice-casting to be at least 20 seconds apart from the previous voice-casting in order to prevent effect confusion from successive voice calls.

5.3.1. Experiment method

Five students (ages around 20–25 years old) participated in the experiment after having given consent. Before the experiment starts, the experimenter attaches a weight to the ankle of the participant to add load to walking. The loads are added to give the burden effects as the subjects are young people instead of the elderly people or those who under rehabilitee treatment. The participant will be guided to walk for 10 m by a voice casting robot and will walk using a walker. The walking distance 10 m was validated in elderly's 6-minutes' walk study by Kawamoto et al. [41]. According to the advice by rehabilitation experts, the voice casting robot is placed in front of the participant to always keep 2 m distance not to lower the line of sight.

In this experiment, we compared the effect of voice-casting with and without considering the arousal level. There are three patterns as follows;

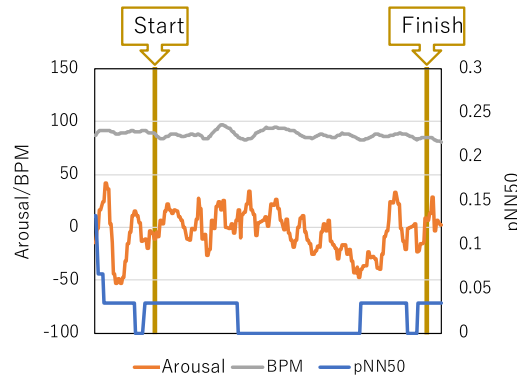


Fig. 10. Changes in BPM, Arousal, and pNN50 for participant A during pattern I (no voice calling).

- (1) No voice-casting at all
- (2) Voice-casting when the participant is unpleasant (low valence, any level of arousal)
- (3) Voice-casting when the participant is unpleasant and deactivated (low valence, low arousal)

All voice-casting is preceded by follow-up, which was described in Section 4.4 and Fig. 6.

5.3.2. Analysis method

For all three experimental patterns, we calculate changes in BPM, arousal value, and pNN50 using the following equations:

$$\Delta_B = \sum_{i=0}^{60} B_B - \sum_{i=60}^M B_W,$$

where B_B and B_W are BPM measured during baseline and walking, and i is the time (secs) from the beginning ($i = 0$) to the end of baseline measurement (60 secs), and M indicates the time at the end of the measurement.

$$\Delta_A = \sum_{i=0}^{60} A_B - \sum_{i=60}^M A_W,$$

where A_B and A_W are arousal value measured during baseline ($i = 0$) to the end of baseline measurement ($i = 60$) and during walking (from $i = 60$ to M), respectively.

$$\Delta_{pNN50} = \sum_{i=0}^N pNN50_B - \sum_{i=N}^M pNN50_W,$$

where $pNN50_B$ and $pNN50_W$ are pNN50 values measured during baseline ($i = 0$) to the end of baseline measurement ($i = 60$) and during walking (from $i = 60$ to M), respectively.

5.4. Result and discussion

5.4.1. Time series analysis result

All the subjects showed the same tendency for the results. Therefore, Fig. 10, Fig. 11, and Fig. 12 show the participant A's results for the time series analysis by visualizing the changes in arousal, BPM, and pNN50 during pattern I, II, and III, respectively.

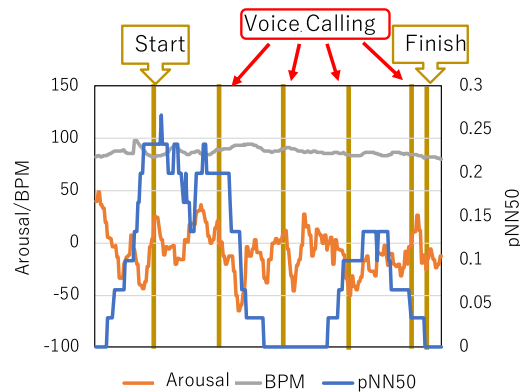


Fig. 11. Changes in BPM, Arousal, and pNN50 for participant A during pattern II (voice calling only when unpleasant).

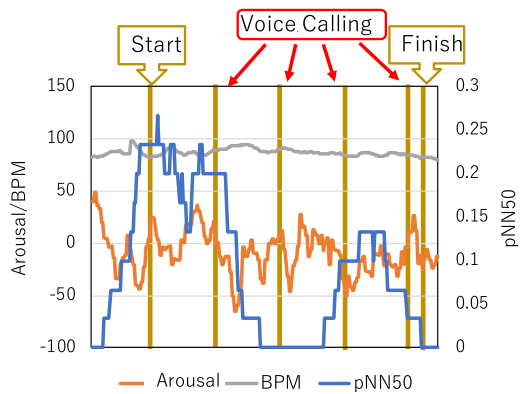


Fig. 12. Changes in BPM, Arousal, and pNN50 for participant A during pattern III (voice calling when unpleasant and deactivated).

From the result shown in Fig. 10, we found that participant's arousal changes are around -20 to 40 in the pattern I (no voice-casting). Valence, which is calculated from pNN50, was quite low, which can be interpreted that the participant felt annoyed. It can be assumed that the robot merely moves in front of the participant without any interaction with the participant; hence, the participant could feel irritated that the robot blocks their way.

On the other hand, results observed from pattern II, shown in Fig. 11, show that the participant's valence (observed by pNN50) was high from the starting point of the experiment, which is interpreted as pleasure. Since the robot would perform voice-casting when it moves, it can be implied that the participant felt pleasure when the robot moves because of the voice-casting, unlike pattern I where the participant felt irritated. This pattern is programmed so that the robot would perform voice-casting when the participant's valence is low. The arousal value of the participant seems to have decreased after the voice-casting. When arousal value is negative, the participant's state enters meditation, which means that the brain is relaxed. Also, it can be noticed from the transition of pNN50 value that the sympathetic nerve became dominant when the robot performs voice-casting. Especially for the first time voice-casting, arousal value increased to as high as 60 . Arousal value is also increased after voice-casting. Therefore, it can be implied that the participant became nervous when the robot performs voice-casting. However, the participant felt more pleasant (the parasympathetic nervous system was more relaxed) when the robot performs voice-casting for the third time. The phrase used in voice-casting was, "You don't need to hurry," which could have made the participant feel more relaxed.

Figure 12 shows arousal, BPM, and pNN50 when the robot performs voice-casting when the participant's arousal value is low, which means that it enters the meditation state. Arousal value tends to increase from the start of voice-casting.

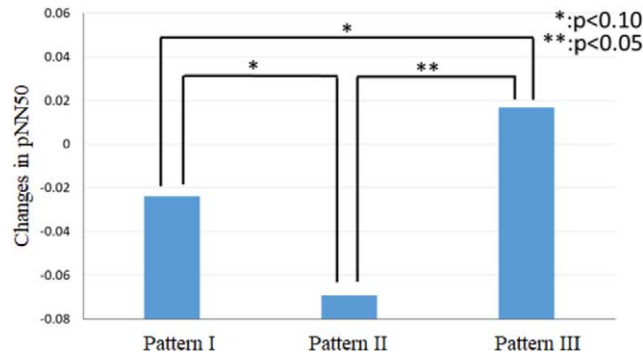


Fig. 13. The comparison of average pNN50 value observed from each pattern, with significant effect showed by ANOVA.

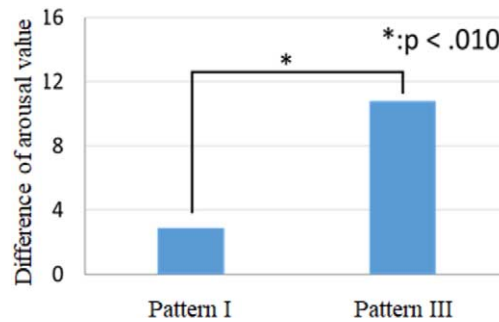


Fig. 14. The comparison of arousal value between pattern I and pattern III, with significant effect showed by *t*-test.

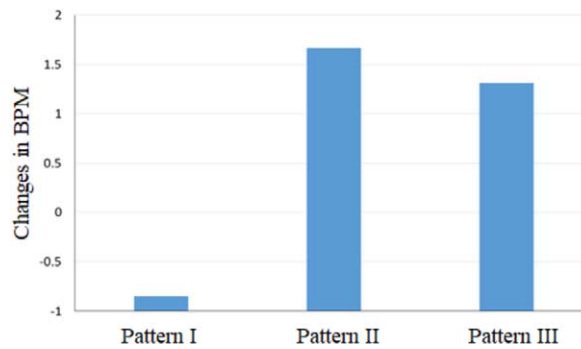


Fig. 15. The comparison of average BPM value observed from each pattern.

5.4.2. Statistical analysis result

We performed a one-way analysis of variance (ANOVA) ($\alpha = 0.05$) on the differences between the average values of BPM, arousal level, and pNN50 during resting (baseline measurement) and during the experiment (all three patterns). The comparisons cover all subjects' results and are shown in Figs 13 to 16.

ANOVA was performed on the difference between the mean of BPM values (see Fig. 13) and the mean of the arousal values (see Fig. 14), but no significant difference was observed ($p > 0.05$). On the other hand, significant differences were observed when ANOVA was performed on the difference between the mean values of pNN50 ($p < 0.05$). Post-hoc comparison by Tukey HSD shows a significant difference between patterns II and III ($p < 0.05$). Also, a significant trend was observed between pattern I and III and pattern I and II ($p < 0.1$), as shown in Fig. 15.

From the above results, it can be implied that voice-casting performed in pattern III when the participant is a little aroused can make the participant feel more relaxed (activated parasympathetic nerve) than the voice-casting

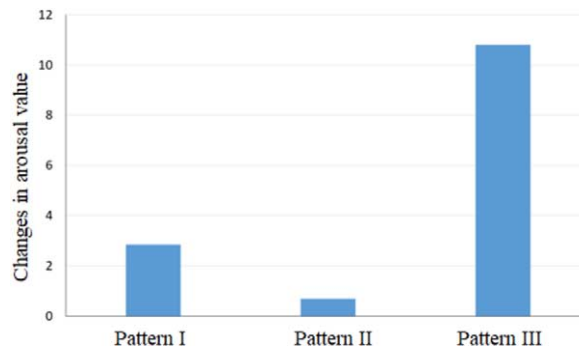


Fig. 16. The comparison of average arousal value observed from each pattern.

performed in pattern II, which occurs when the participant felt unpleasant. Also, pattern II could make the participant feel more relaxed than the pattern I when no voice-casting is made.

Therefore, it can be summarized that the pNN50 of the participant experiencing pattern II where valence (voice-casting when the participant is unpleasant) is low, where the sympathetic nerve is dominant. A sympathetic dominant state is when the blood pressure rises, and body tension is increased, or when the participant was trying hard for rehabilitation. So, performing voice-casting when the valence is low could boost the sympathetic dominant state.

Figure 16 shows the comparison with the *t*-test result of the arousal level of the pattern I and pattern III. This result shows that voice-casting could activate the biometric information of the participant during rehabilitation.

From the above pNN50 (valence) and arousal (attention-meditation) values, it can be implied that the rehabilitation patient could be encouraged to work harder and be more active during rehabilitation by using voice-casting when low valence and low arousal.

5.4.3. Discussion

From the results described above, there were no significant differences between the mean of BPM and the mean of arousal value. This could be because the experiment was performed with a predefined distance and speed, which were similar for all patterns. Therefore, having voice-casting or not, did not particularly change the load of the participant. On the other hand, a significant difference was observed between the means of pNN50. It is considered that when the participant is in a deactivation state, voice-casting could make them feel more pleasant, as seen from the rise in pNN50. Also, performing voice-casting while the participant is in an unpleasant state could stimulate tension more than no voice-casting at all.

On the other hand, although the arousal level was increased in both voice-casting and no voice-casting while walking, *t*-test results showed that the arousal level increased more when performing voice-casting during the deactivation condition. From these results, it can be suggested that walking could increase arousal, as well as voice-casting.

From the rise of arousal and tension seen in the results, it can be implied that participants are more active and willing to try hard in rehabilitation. In addition, it is possible to enhance relaxation (increase in valence) by performing voice-casting when the participant's valence is low. Therefore, we can grasp the state of a person more closely by taking into account the two values of valence and arousal, hence, perform the more appropriate action for encouraging the rehabilitation process.

In the social robot domain, it has been found that novelty effects exist in the use of technology. While the use new technology is new to the user, it can promote positive response, however, positive response or interest could drop over time. This is found in many works involving social robot [42,43]. In our work, novelty may affect the positive response of the user. However, human emotion is unpredictable, therefore, the interaction between user and robot can always be something new. For example, it is possible to create the robot that respond to the user according to the emotion, or even mimic the user's emotion at the time being. Although it is not the scope of this work, it could be considered as future work.

6. Conclusion and future work

In this research, we proposed a voice-casting robot to reduce the workload of the human caregiver during rehabilitation therapy. The proposed voice-casting robot would estimate the patient's emotion using real-time sensing of biometric information and select suitable supportive phrases for voice casting in order to increase the patient's motivation during rehabilitation therapy. From the experiment, we found a significant improvement of the participant's motivation when the robot uses voice-casting rather than moving only. As a result, it can be implied that our proposed voice-casting robot is effective for rehabilitation, and hence could reduce the workload of the human caregiver.

One issue with our evaluation experiment is that the participants were not actually older people or rehabilitation patients. Instead, the participants in our experiment are young people in their 20 s to 40 s. So, it may be too quick to assume that our proposed voice-casting robot will be effective in actual rehabilitation therapy. Therefore, it is important to evaluate our experiment on elderly people or rehabilitation patients to further prove our proposed method. Next, the participants' comments also suggested that the robot could be better designed for multiple voice styles, quieter motor noise, and improved appearance. Also, although the evaluation of Paro [2,4,23] used the EEG signal for biometric evaluation, we did not adopt such a signal in this study due to many restrictions such as the requirement of stationary state for measurement.

We also performed another experiment on an improved version of voice-casting robot by also considering arousal level, measurable from EEG signal. As a result, it was found that participants felt more comfortable when voice-casting was performed by referring to both arousal and valence.

For the next step, the sample size should be increased in order to acquire more robust data for statistical analysis. Even though, the sample size is small, it is reasonably enough to do the analysis for the physiological measurements [19,44–47]. Also, in the future, more consideration regarding timing could be investigated. Regarding voice casting, a study has shown that messages conveyed through words during rehabilitation therapy, may only contain 35% of the total information. The remaining 65% could be conveyed through expression or tone [48]. Therefore, it is necessary to examine in more detail behavior such as facial expressions and follow-up used in a fixed way in this research. Others physiological data such as breathing frequency or electrodermal activity (EDA) also may be considered in order to gain better result. Finally, if we can tailor voice casting and follow-ups according to the state and timing of individuals using machine learning and AI in the future, it could be possible to develop a voice-casting robot with better supportive effects considering individual differences.

Conflict of interest

The authors have no conflict of interest to report.

References

- [1] The estimation of supply and demand for human resources for caregiver toward year 2025 (Definite value) (in Japanese), 7, Ministry of Health, Labour and Welfare, 2015.
- [2] T. Shibata and K. Wada, Introduction of field test on robot therapy by seal robot, PARO, *Journal of the Robotics Society of Japan* **29**(3) (2011), 246–249. doi:10.7210/jrsj.29.246.
- [3] T. Shibata, An overview of human interactive robots for psychological enrichment, *Proceedings of the IEEE* **92**(11) (2004), 1749–1758. doi:10.1109/JPROC.2004.835383.
- [4] T. Shibata and K. Wada, Robot therapy: A new approach for mental healthcare of the elderly – a mini-review, *Gerontology* **57**(4) (2011), 378–386. doi:10.1159/000319015.
- [5] M.J. Bickers, Does verbal encouragement work? The effect of verbal encouragement on a muscular endurance task, *Clinical Rehabilitation* **7**(3) (1993), 196–200. doi:10.1177/026921559300700303.
- [6] F. Martí Carrillo, J. Butchart, S. Knight, A. Scheinberg, L. Wise, L. Sterling and C. McCarthy, Adapting a general-purpose social robot for paediatric rehabilitation through in situ design, *ACM Transactions on Human-Robot Interaction (THRI)* **7**(1) (2018), 1–30. doi:10.1145/3203304.

- [7] J.C. Pulido, C. Suarez-Mejias, J.C. Gonzalez, A.D. Ruiz, P.F. Ferri, M.E.M. Sahuquillo, C.E.R. De Vargas, P. Infante-Cossio, C.L.P. Calderon and F. Fernandez, A socially assistive robotic platform for upper-limb rehabilitation: A longitudinal study with pediatric patients, *IEEE Robotics & Automation Magazine* **26**(2) (2019), 24–39. doi:10.1109/MRA.2019.2905231.
- [8] Y. Kibishi, Y. Takahashi and K. Sasaki, Effective voice casting to patients in rehabilitation, *Japanese Journal of Allied Health and Rehabilitation* **3** (2004), 25–29, (in Japanese).
- [9] A. Tapus, C. Țăpuș and M.J. Matorić, User–robot personality matching and assistive robot behavior adaptation for post-stroke rehabilitation therapy, *Intelligent Service Robotics* **1**(2) (2008), 169. doi:10.1007/s11370-008-0017-4.
- [10] A. Tapus and M.J. Matorić, User personality matching with a hands-off robot for post-stroke rehabilitation therapy, in: *Proceedings of Experimental Robotics*, 2008, pp. 165–175. doi:10.1007/978-3-540-77457-0_16.
- [11] S. Takeuchi, A. Sakai, S. Kato and H. Itoh, An emotion generation model based on the dialogist likability for sensitivity communication robot, *Journal of the Robotics Society of Japan* **25**(7) (2007), 1125–1133. doi:10.7210/jrsj.25.1125.
- [12] P. Andrés and M. Van der Linden, Age-related differences in supervisory attentional system functions, *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences* **55**(6) (2000), 373–380. doi:10.1093/geronb/55.6.P373.
- [13] A. Di Domenico, R. Palumbo, N. Mammarella and B. Fairfield, Aging and emotional expressions: Is there a positivity bias during dynamic emotion recognition?, *Frontiers in Psychology* **6** (2015), 1130. doi:10.3389/fpsyg.2015.01130.
- [14] P. Ekman and W.V. Friesen, *Facial Action Coding System: Investigator's Guide*, Consulting Psychologists Press, 1978.
- [15] M. Coulson, Attributing emotion to static body postures: Recognition accuracy, confusions, and viewpoint dependence, *Journal of Nonverbal Behavior* **28**(2) (2004), 117–139. doi:10.1023/B:JONB.0000023655.25550.be.
- [16] R.A. Calvo, S. D'Mello, J.M. Gratch and A. Kappas, *The Oxford Handbook of Affective Computing*, Oxford University Press, USA, 2015.
- [17] J. Healey, L. Nachman, S. Subramanian, J. Shahabdeen and M. Morris, Out of the lab and into the fray: Towards modeling emotion in everyday life, in: *Proceedings of International Conference on Pervasive Computing*, 2010, pp. 156–173. doi:10.1007/978-3-642-12654-3_10.
- [18] S. Poria, E. Cambria, R. Bajpai and A. Hussain, A review of affective computing: From unimodal analysis to multimodal fusion, *Information Fusion* **37** (2017), 98–125. doi:10.1016/j.inffus.2017.02.003.
- [19] S.M. Alarcao and M.J. Fonseca, Emotions recognition using EEG signals: A survey, *IEEE Transactions on Affective Computing* **10**(3) (2017), 374–393. doi:10.1109/TAFFC.2017.2714671.
- [20] A. Vinciarelli and G. Mohammadi, A survey of personality computing, *IEEE Transactions on Affective Computing* **5**(3) (2014), 273–291. doi:10.1109/TAFFC.2014.2330816.
- [21] T. Hoshishiba, H. Uemura, T. Hojo and T. Ichiro, Heartbeat variability analysis for music stimulation, *The Journal of Acoustical Society of Japan*, *ASJ* **51**(3) (1995), 163–173, (in Japanese). doi:10.20697/jasj.51.3_163.
- [22] M. Trimmel, Relationship of heart rate variability (HRV) parameters including pNNxx with the subjective experience of stress, depression, well-being, and every-day trait moods (TRIM-T): A pilot study, *The Ergonomics Open Journal* **8**(1) (2015). doi:10.2174/1875934301508010032.
- [23] T. Shibata and K. Wada, Therapeutic seal robot “PARO” and its effective operation, *Journal of the Society of Instrument and Control Engineers* **51**(7) (2012), 640–643.
- [24] N. Maclean, P. Pound, C. Wolfe and A. Rudd, Qualitative analysis of stroke patients' motivation for rehabilitation, *BMJ* **321**(7268) (2000), 1051–1054. doi:10.1136/bmj.321.7268.1051.
- [25] K. Rattanyu and M. Mizukawa, Emotion recognition based on ECG signals for service robots in the intelligent space during daily life, *Journal of Advanced Computational Intelligence and Intelligent Informatics* **15**(5) (2011), 582–591. doi:10.20965/jaciii.2011.p0582.
- [26] J. Taelman, S. Vandeput, A. Spaepen and S. Van Huffel, Influence of mental stress on heart rate and heart rate variability, in: *Proceedings of 4th European Conference of the International Federation for Medical and Biological Engineering*, 2009, pp. 1366–1369. doi:10.1007/978-3-540-89208-3_324.
- [27] U.R. Acharya, K.P. Joseph, N. Kannathal, C.M. Lim and J.S. Suri, Heart rate variability: A review, *Medical and Biological Engineering and Computing* **44**(12) (2006), 1031–1051. doi:10.1007/s11517-006-0119-0.
- [28] G. Valenza, A. Lanata and E.P. Scilingo, The role of nonlinear dynamics in affective valence and arousal recognition, *IEEE Transactions on Affective Computing* **3**(2) (2011), 237–249. doi:10.1109/T-AFFC.2011.30.
- [29] R.D. Lane, K. McRae, E.M. Reiman, K. Chen, G.L. Ahern and J.F. Thayer, Neural correlates of heart rate variability during emotion, *Neuroimage* **44**(1) (2009), 213–222. doi:10.1016/j.neuroimage.2008.07.056.
- [30] S. Futomi, M. Ohsuga and H. Terashita, Method for assessment of mental stress during high-tension and monotonous tasks using heart rate, respiration and blood pressure, *The Japanese Journal of Ergonomics* **34**(3) (1998), 107–115, (in Japanese). doi:10.5100/jje.34.107.
- [31] H. Zenju, A. Nozawa, H. Tanaka and H. Ide, Estimation of unpleasant and pleasant states by nasal thermogram, *IEEJ Transactions on Electronics, Information and Systems* **124**(1) (2004), 213–214. doi:10.1541/ieejieiss.124.213.
- [32] F.-T. Sun, C. Kuo, H.-T. Cheng, S. Buthpitiya, P. Collins and M. Griss, Activity-aware mental stress detection using physiological sensors, in: *Proceedings of Mobile Computing, Applications, and Services*, Berlin, Heidelberg, 2010, pp. 211–230.
- [33] D. Li and W.C. Sullivan, Impact of views to school landscapes on recovery from stress and mental fatigue, *Landscape and Urban Planning* **148** (2016), 149–158. doi:10.1016/j.landurbplan.2015.12.015.
- [34] L. Fiorini, G. Mancoppi, F. Semeraro, H. Fujita and F. Cavallo, Unsupervised emotional state classification through physiological parameters for social robotics applications, *Knowledge-Based Systems* **190** (2020), 105217. doi:10.1016/j.knsys.2019.105217.
- [35] S.S. Panicker and P. Gayathri, A survey of machine learning techniques in physiological based mental stress detection systems, *Biocybernetics and Biomedical Engineering* **39**(2) (2019), 444–469. doi:10.1016/j.bbe.2019.01.004.

- [36] A.J. Camm, M. Malik, J.T. Bigger, G. Breithardt, S. Cerutti, R.J. Cohen, P. Coumel, E.L. Fallen, H.L. Kennedy and R. Kleiger, Heart rate variability. Standards of measurement, physiological interpretation, and clinical use, 1996. doi:10.1161/01.CIR.93.5.1043.
- [37] Softbank Robotics, NAO, Available from: <https://www.softbankrobotics.com/emea/en/nao> (2017). Access date: 19 July 2021.
- [38] E.M. de Carvalho Abreu, R. de Souza Alves, A.C.L. Borges, F.P.S. Lima, A.R. de Paula Júnior and M.O. Lima, Autonomic cardiovascular control recovery in quadriplegics after handcycle training, *Journal of Physical Therapy Science* **28**(7) (2016), 2063–2068. doi:10.1589/jpts.28.2063.
- [39] J.A. Russell, A circumplex model of affect, *Journal of Personality and Social Psychology* **39**(6) (1980), 1161–1178. doi:10.1037/h0077714.
- [40] NeuroSky, MindWave Mobile: User Guide, Available from: http://download.neurosky.com/support_page_files/MindWaveMobile/docs/mindwave_mobile_user_guide.pdf (2015). Access date: 19 July 2021.
- [41] K. Kawamoto, H. Kanazawa, H. Iwamoto, K. Fujii, N. Deguchi, T. Shima, S. Kamei and T. Shirakawa, Is the 6-minute walk test for the 10-meter course useful? – Study in the elderly with orthopedic diseases, *Supplementary of Physical Therapy* **36**(2) (2009), C3P2480–C3P2480, (in Japanese). doi:10.14900/cjpt.2008.0.C3P2480.0.
- [42] T. Kanda, T. Hirano, D. Eaton and H. Ishiguro, Interactive robots as social partners and peer tutors for children: A field trial, *Human–Computer Interaction* **19**(1–2) (2004), 61–84. doi:10.1207/s15327051hci1901%262_4.
- [43] K. Dautenhahn, Socially intelligent robots: Dimensions of human–robot interaction, *Philosophical Transactions of the Royal Society B: Biological Sciences* **362**(1480) (2007), 679–704. doi:10.1098/rstb.2006.2004.
- [44] S. Gündoğdu and E. Gülbetkin, P82 assessment of correlation between EEG frequency subbands and HRV in playing puzzle video game, *Clinical Neurophysiology* **131**(4) (2020), e217–e218. doi:10.1016/j.clinph.2019.12.080.
- [45] M. Kumar, D. Singh and K. Deepak, Identifying heart-brain interactions during internally and externally operative attention using conditional entropy, *Biomedical Signal Processing and Control* **57** (2020), 101826. doi:10.1016/j.bspc.2019.101826.
- [46] J.P. Fuentes-García, T. Pereira, M.A. Castro, A.C. Santos and S. Villafaina, Psychophysiological stress response of adolescent chess players during problem-solving tasks, *Physiology & Behavior* **209** (2019), 112609. doi:10.1016/j.physbeh.2019.112609.
- [47] T. Page and F.J. Rugg-Gunn, Bitemporal seizure spread and its effect on autonomic dysfunction, *Epilepsy & Behavior* **84** (2018), 166–172. doi:10.1016/j.yebeh.2018.03.016.
- [48] E. Saito, *Social Psychology (Mainly Role Theory) and Other Theories*, 1995, (in Japanese).