Psychophysiological Responses to Robotic Rehabilitation Tasks in Stroke

Domen Novak, Jaka Ziherl, Andrej Olenšek, Maja Milavec, Janez Podobnik, Matjaž Mihelj, and Marko Munih, *Member, IEEE*

***Abstract*—This paper presents the analysis of four psychophysiological responses in post-stroke upper extremity rehabilitation. The goal was to determine which psychophysiological responses would provide the most reliable information about subjects' psychological states during rehabilitation. Heart rate, skin conductance, respiration, and skin temperature were recorded in a stroke group and a control group during two difficulty levels of a pick-and-place task performed in a virtual environment using a haptic robot and during a cognitive task. Psychophysiological measurements were correlated with results of a self-report questionnaire. All four responses showed significant changes in response to the different tasks. Skin conductance differentiated between the two difficulty levels and was correlated with self-reported arousal in both stroke and control groups. Skin temperature differentiated between the two difficulty levels for the control group, but provided poor results for the stroke group. Heart rate and respiration increased during tasks, but their connection to psychological state was unclear. Results suggest that, of the four measured responses, skin conductance offers the most potential as a psychological state indicator, with other measures providing supplementary information. Psychophysiological measurements could thus be used in closed-loop biocooperative systems that would detect the user's psychological state and change the course of therapy accordingly.**

***Index Terms*—Biocooperative robotics, human factors, multimodal interfaces, psychophysiological measurements, rehabilitation robotics, stroke.**

I. Introduction

ROBOTIC interfaces are becoming increasingly common in motor rehabilitation [1], [2]. Such devices enable more intensive therapy and offer an objective estimation of the patient's motor performance and functional improvement [3]. Rehabilitation can be additionally enhanced by combining robots with virtual reality [4]–[6]. Such systems improve patient motivation by allowing exercise in interesting, varied virtual environments. However, while rehabilitation robots can provide objective information about a patient's motor performance, they do not offer insight into his or her psychological state: mood, motivation, engagement, etc. Such factors, however, are known to be very positive to the success of rehabilitation, and encouraging nonmotivated stroke patients improves the likelihood of their eventual recovery [7], [8]. Thus, a quantitative method of measuring patients' psychological states during rehabilitation would be very useful.

One possible way for rehabilitation robots to sense a patient's psychological state would be through the use of psychophysiological measurements. Psychophysiological measurements are defined as the measurements of physiological responses to changes in psychological state. They have been used to assess mood and engagement in a variety of settings and applications, including computer games [9] and virtual reality exposure therapy [10]. Such measurements are noninvasive and can be obtained without the subject's active cooperation, thus allowing the subject to focus on the task at hand. They can be divided into measurements of the central nervous system (electroencephalography, functional magnetic resonance imaging, etc.) and measurements of the autonomic nervous system.

Our paper focuses primarily on four measurements of the autonomic nervous system, for which the required equipment is relatively inexpensive and easy to apply to the subject: heart rate, skin conductance, respiration, and skin temperature. Previous studies have shown that heart rate increases and heart rate variability decreases as a response to cognitive workload [11]–[13]. However, interpretations of heart rate are still controversial, with sometimes conflicting results from different studies (see [14] for examples). Skin conductance increases with general psychological arousal and cognitive workload [12], [15], [16]. Respiratory rate also increases with arousal and cognitive workload [11] while respiratory variability decreases during mentally demanding tasks [17]. Skin temperature decreases as a result of cognitive workload [18] as well as a result of tension or anxiety [19].

However, there are two issues with using psychophysiological measurements in motor rehabilitation. First, physiological responses are influenced not only by a person's psychological state, but also by physical activity. Thus, most psychophysiological studies try to limit physical activity. This is not possible in motor rehabilitation. Second, motor rehabilitation is often performed with patients who have suffered damage to the autonomic nervous system. Stroke patients, for instance, show long-lasting abnormalities in sweating and heart rate variability [20].

We were specifically interested in whether psychophysiological responses could be harnessed for use in closed-loop biocooperative systems that could adapt therapy based on the state of the subject. Prototypes of closed-loop psychophysiological systems have been developed for applications such as flight simulators [15]. However, since stroke patients show physiological

Manuscript received September 25, 2009; revised January 27, 2010; accepted March 17, 2010. First published April 12, 2010; current version published August 11, 2010. This work was supported in part by the EU Information and Communication Technologies Collaborative Project MIMICS under Grant 215756 and in part by the Slovenian Research Agency.

The authors are with the Faculty of Electrical Engineering, University of Ljubljana, SI-1000 Ljubljana, Slovenia (e-mail: domen.novak@robo.fe.uni-lj.si).

Digital Object Identifier 10.1109/TNSRE.2010.2047656

abnormalities, we were uncertain whether psychophysiological responses would provide any useful information in a task with no physical workload, let alone a physically demanding rehabilitation task.

The goal of our study was to answer four questions.

1) How different are stroke patients' physiological responses from responses of healthy subjects?
2) Do stroke patients show distinctive psychophysiological responses to a physically undemanding task despite damage to the autonomic nervous system?
3) Do stroke patients show distinctive psychophysiological responses to a motor rehabilitation task despite damage to the autonomic nervous system and effects of physical workload?
4) Which physiological responses would provide the most reliable information about subjects' psychological states in motor rehabilitation?

To answer our questions, stroke patients and a healthy control group were presented with two tasks. The first was a task that required only cognitive effort with no movement of the afflicted arm. The second was a reaching and grasping exercise task performed in a multimodal environment. Psychophysiological data from both tasks were compared between the stroke and control groups.

II. Materials and Methods

A. Ethical Approval

Before the study began, ethical approval was obtained both from the National Medical Ethics Committee of the Republic of Slovenia and from the Medical Ethics Committee of the Institute for Rehabilitation of the Republic of Slovenia.

B. Subjects

The stroke group consisted of 23 subjects (age 51.0 $\pm$ 13.3 years, age range 23–69 years, 16 males, 7 females). They were diagnosed with subarachnoid hemorrhage (4 subjects), intracerebral hemorrhage (9 subjects) or cerebral infarction (10 subjects). As a result of the stroke, 13 suffered from hemiparesis of the left side of the body and 10 suffered from hemiparesis of the right side of the body. All were right-handed before the stroke. All were undergoing motor rehabilitation at the Institute for Rehabilitation of the Republic of Slovenia. Time between stroke onset and the experiment session was 154 $\pm$ 79 days. A majority of the group had received secondary stroke prevention drugs (including antihypertensives) prior to participation in the study. Three had received insulin due to diabetes (but had no diabetes-related complications), two had received ischemic heart disease treatment drugs, five had received SSRIs, three had received low doses of antiepileptics (preventive doses following aneurysm surgery), three had received short-acting sedatives and one had received a low dose of antipsychotics.

A day before the session, subjects in the stroke group were tested with both the mini-mental state examination (MMSE) [21] and the functional independence measure (FIM) [22]. Score on the MMSE was 27.2 $\pm$ 3.6 (out of a possible 30). All but three subjects scored between 26 and 30. Of the remaining three, one scored 24 but was not excluded from the study since he was able to communicate and comprehend the tasks without problems. The other two subjects had lower scores due to dysphasia. These two were interviewed by a clinical expert and approved for participation in the study. Score on the FIM was 101 $\pm$ 13 (out of a possible 126).

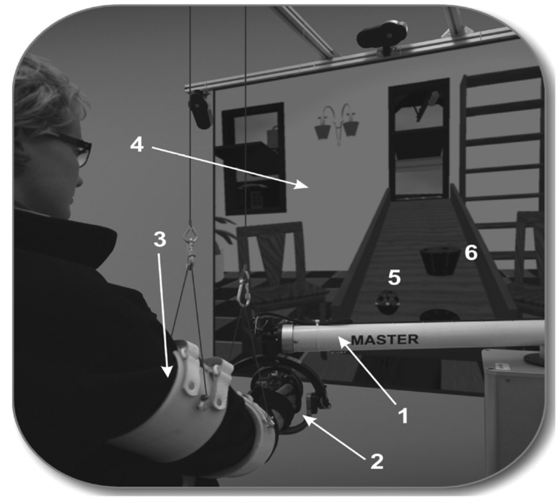

Fig. 1. A subject performing the virtual rehabilitation task. The subject performs the task using the robot (1) and grasping device (2) while his/her arm is supported by cuffs (3). The screen (4) shows a sloped table, a ball (5) and a basket (6).

The control group consisted of 23 subjects (age 50.5 $\pm$ 12.6 years, age range 24–68 years, 16 males, 7 females) with no major physical or cognitive defects. All were right-handed. To better match the control group and the stroke group, 13 controls performed the tasks with their left hand while 10 performed the tasks with their right hand.

C. Hardware

The HapticMaster robot [23], developed by Moog FCS, was used as the haptic interface. This robot offers movement with three degrees-of-freedom. The end-point is equipped with force sensors. A two-axis gimbal with a two-degree-of-freedom passive grasping module instrumented with force cells is attached to the end-point of the robot [24]. The subject's arm was additionally supported by two cuffs fastened above and below the elbow. These cuffs were connected to a pulley which applied a constant torque in order to compensate for the gravity acting on the subject's arm. A 1.4 $\times$ 1.4-m screen was used to display visual data. Subjects sat approximately 1.25 m in front of the screen, with the robot situated between the seat and the screen (Fig. 1).

Physiological signals were sampled at 2.4 kHz using a g.USBamp signal amplifier (g.tec Medical Engineering GmbH). The electrocardiogram (ECG) was recorded using four disposable surface electrodes placed in a configuration suggested by the manufacturer of the signal amplifier (one electrode on the left part of the chest, one on the right part of the chest, one on the left part of the abdomen, and a ground electrode on the upper left part of the back). Skin conductance was measured using a g.GSR sensor (g.tec). The electrodes were placed on the medial phalanxes of the second and third fingers of the unaffected hand. The sensor generated a constant voltage between the two electrodes and measured the current between the electrodes in order to estimate skin conductance. Respiratory rate was obtained using a thermistor-based SleepSense Flow sensor placed beneath the nose. Peripheral skin temperature was measured using a g.TEMP sensor (g.tec) attached to the distal phalanx of the fifth finger of the unaffected hand.

D. Virtual Rehabilitation Task

The virtual rehabilitation (VR) task combines reaching and grasping exercise. A photograph of a subject performing the task using the robot is shown in Fig. 1. In the center of the screen, there is a table sloped toward the subject. At the beginning of the task, a ball appears at the top of the slope and starts rolling downward. The subject's goal is to catch the ball before it reaches the lower end of the table. Once the ball is grasped, a basket appears above the table. The subject must then hold the ball and place it in the basket. Once the ball is dropped into the basket or falls off the table, another ball appears at the top of the table, the basket disappears and the task continues. The robot allows the subject to feel each virtual item.

The robot offers various modes of support. If a subject is unable to perform any or all of the following, the robot will actively guide his or her arm in order to move left or right and reach the ball, squeeze the grasping device in order to grasp the ball, and/or lift the ball into the basket. For reaching support (left–right movements), the robot pulls the subject's hand toward the ball with a maximum force of 10 N. The subject can thus reinforce or resist the robot's guiding force with his or her active arm movement. If the subject does not resist the robot, the reaching support system will reach the ball in a majority of cases (but may miss the ball if the starting position of the subject's hand is sufficiently far away from the ball). For grasping support, the robot automatically grasps the ball as long as the subject's hand is in the correct position, regardless of whether the subject is squeezing the grasping module. For lifting support, the subject's hand is pulled along a predefined trajectory [25] toward the basket. The subject can reinforce or resist the guiding force with his or her active arm movement. If the subject does not resist lifting support or release the ball, the ball will always be successfully placed into the basket by the lifting support system. Four subjects in the stroke group required reaching support, seven required grasping support, and eight required lifting support (with some subjects requiring multiple types of support). The control group did not receive any support from the robot.

A second, harder version of the task (henceforth referred to as the harder VR task) was also designed. Meant to be more mentally demanding but equally physically demanding, the harder VR task had inverted left–right controls. If the subject moved his or her arm to the left, the virtual hand on the screen moved right (and vice–versa).

In addition to the VR task, we also wished to evaluate psychophysiological responses to a less complex task. To this end, a physical control task was introduced where subjects moved the robot left and right at a moderate speed while nothing was shown on the display and all force feedback was disabled. While both the physical control task and the VR task require reaching and some degree of coordinated movement, the physical control task is less complex and less cognitively demanding since it does not require lifting movements, does not provide visual stimuli and is not timed. Thus, since psychophysiological responses are strongly influenced by cognitive workload, the physical control task should evoke weaker psychophysiological responses.

E. Cognitive Task

In addition to the VR task, subjects were presented with a task that required only cognitive effort. We used a variant of the Stroop word-color interference task [26] that has been extensively studied by psychologists. Subjects were shown a word on the screen. The word was either "red," "blue," or "green." The color of the letters was also red, blue or green—but the word and the color of the letters did not always match. Subjects were given a keypad with three grey buttons, with the words "red," "blue," and "green" written above the buttons in black. They were told to ignore the meaning of the word and, as quickly as possible, press the button corresponding to the color with which the word was written. Once a button was pressed, a new word was generated. Occasionally, however, the word was generated in black color. In this case, subjects had to push the button corresponding to the meaning of the word rather than its color. The words were randomly generated with the following probabilities: 40% chance of word with matching color, 40% chance of word with different color, 20% chance of word in black.

The two subjects with dysphasia were excluded from the Stroop task.

F. Questionnaires

During the experiment, subjects were periodically presented with nine-point *arousal* and *valence* scales from the Self-Assessment Manikin (SAM) [27]. These scales allow subjects to rate their level of emotional valence and arousal graphically by choosing the picture that best represents their current mood. *Valence* (sometimes also called pleasure) is defined as positive versus negative affective states (e.g., humiliation, disinterest, and anger at one end versus excitement, relaxation, and tranquility at the other end) while *arousal* is defined in terms of mental alertness and physical activity (e.g., sleep, inactivity, boredom, and relaxation at the lower end versus wakefulness, tension, exercise, and concentration at the higher end) [28]. *Valence* and *arousal* were converted to numerical values for purposes of analysis. For *valence*, 1 represented extremely negative *valence* while 9 represented extremely positive *valence*. For *arousal*, 1 represented extremely low *arousal* while 9 represented extremely high *arousal*. The SAM was chosen over other questionnaires for two reasons. First, the physiological effects of arousal and valence are well-documented (see [14] for examples). Second, the SAM is graphical in nature and thus very simple to use; in pretesting, some subjects in the stroke group had difficulty comprehending more complex self-report questionnaires. Despite this simplicity, the SAM has been shown to yield results similar to those of more complex self-report scales such as the semantic differential [27].

G. Experiment Protocol

The experiment was conducted in a dedicated room at the Institute for Rehabilitation of the Republic of Slovenia. Three people were present: the subject, experiment supervisor and occupational therapist. Upon arrival, subjects were informed of the purpose and procedure of the experiment, then signed an informed consent form. Then, they were seated in front of the robot. The affected arm was strapped into the cuffs and grasping device, and the physiological sensors were attached. The normal VR task was demonstrated, and subjects were allowed to practice it briefly. Each subject practiced for at least 2 min, and more time was given to any subject who had not yet attained a basic level of proficiency. During practice, the three modes of support were set manually for each subject. Then, subjects went through the following procedure: rest period, physical control task, rest period, normal VR task, harder VR task.

After the harder VR task, the keypad used for the Stroop task was strapped to the subject's upper leg so that it would not fall off. The Stroop task was explained and demonstrated, and at least 2 min were once again given to practice. Then, subjects went through a 3-min rest period followed by the Stroop task. They pushed the buttons on the keypad with the thumb of their unaffected hand. The skin conductance and skin temperature sensors were not removed, but remained on the other fingers of the unaffected hand. While this may have affected measurements, it was necessary since many subjects were unable to push buttons with the affected hand. After the Stroop task, the experiment was concluded and a brief informal interview was conducted.

Each task and rest period lasted 3 min, and the SAM was presented on the screen after each period. Subjects verbally made a selection for both arousal and valence scales. Subjects remained quiet during rest, as these periods served as baseline periods for physiological measurements.

While the experiment supervisor and occupational therapist maintained silence in the room during most baseline and task periods, the therapist was permitted to provide verbal guidance and encouragement during the VR task. This was unavoidable, as some subjects in the stroke group required guidance to perform the task without becoming excessively frustrated. To establish similar conditions, the therapist also provided verbal guidance to the control group. If guidance was not necessary, encouraging statements were provided to ensure that all subjects were verbally stimulated.

H. Evaluation of Task Performance and Work

In both versions of the VR task, performance was evaluated using three parameters: *percentage of caught balls, percentage of balls placed into the basket* (calculated as percentage of all balls in the period, not as percentage of caught balls), and *total work* performed during the entire period. Work was calculated as the integral of power (force times velocity), with force data obtained from the robot's force sensors and velocity data obtained from the robot's movement sensors.

The *percentage of caught balls* was calculated only for subjects who did not receive catching support from the robot while the *percentage of balls placed into the basket* was calculated only for subjects who received neither catching nor lifting assistance. In the Stroop task, performance was evaluated using two parameters: *percentage of correct answers* (i.e., correctly chosen colors) and the *mean answer time* (the interval from the moment a color was displayed to the moment the subject pressed any button).

I. Physiological Signal Processing

Each subject's physiological signals were recorded during the experiment. Afterwards, the signals were analyzed offline and several parameters were extracted for each 3-min period. From the ECG, *mean heart rate* as well as two standardized measures of heart rate variability (HRV) were calculated: the standard deviation of NN intervals (*SDNN*) and the square root of the mean squared differences of successive NN intervals (*RMSSD*) [29].

The skin conductance signal can be divided into two components: the skin conductance level (SCL) and skin conductance responses (SCRs). The SCL is the baseline level of skin conductance in the absence of any discrete environmental event. *Mean SCL* was calculated. SCRs are transient increases in skin conductance. An increase in skin conductance was classified as a SCR if its amplitude exceeded 0.05 μS and its peak occurred less than 5 s after the beginning of the increase. *SCR frequency* and *mean SCR amplitude* were calculated. The signal obtained from the respiration sensor was a signal which increased during inspiration and decreased during expiration. *Mean respiratory rate* was calculated by measuring distances between two consecutive peaks. *Respiratory rate variability* was estimated by calculating the standard deviation of the respiratory rate time series.

Final skin temperature was calculated as the mean temperature during the last five seconds of each period.

Although changes in psychological state naturally do occur in the course of each 3-min period, each parameter was calculated over the entire period. This is because some parameters (e.g., HRV, SCR frequency) can only be calculated over a period of several minutes while others (e.g., skin temperature) change relatively slowly in response to stimuli.

J. Data Analysis

The first step of the data analysis was to examine performance and work during the two levels of the VR task and the Stroop task. For the two levels of the VR task, a mixed-design analysis of variance (ANOVA) with one between-subjects factor (group: stroke or control) and one within-subjects factor (task difficulty: normal or harder VR task) was used for each parameter. For the Stroop task, t-tests were used to compare the two performance parameters between groups. The goal of this step was to determine whether the control group performed better than the stroke group and whether performance during the harder VR task was worse than during the normal VR task.

The second step of the data analysis was to compare absolute values of psychophysiological parameters between baseline and task periods. This was done separately for each task in a mixed-design ANOVA with one between-subjects factor (group: stroke or control) and one within-subject factor (period type: baseline or task). The goal of this step was to determine whether baseline values of physiological parameters are different between the groups, whether each task causes significant psychophysiological changes and whether these changes are different between the stroke and control groups.

The third step of the data analysis was to compare relative values of psychophysiological parameters between the different task periods. Relative values of physiological parameters were defined by either subtracting the preceding baseline value from the task value or by subtracting the preceding baseline value from the task value and dividing the result by the baseline value. The second definition was used for those parameters where baseline values varied highly among subjects: both measures of HRV (*SDNN* and *RMSSD*), *SCR frequency*, and *respiratory rate variability*. The comparison of relative values was done in a mixed-design ANOVA with one between-subjects factor (group: stroke or control) and one within-subjects factor (task type: physical control task, normal VR task, harder VR task, and Stroop task). The goal of this step was to determine whether the different task periods cause different psychophysiological responses at an aggregate level. For instance, the harder VR task could evoke larger psychophysiological responses than the normal VR task since most subjects can be expected to find it more cognitively demanding.

The fourth step of the data analysis was to correlate relative values of psychophysiological parameters with results of the SAM and with performance data. Spearman correlations were used in cases involving results of the SAM (where the data is ordinal). Pearson correlations were used in other cases. The goal of this step was to analyze the connections between physiological and nonphysiological responses. For instance, due to the large differences between subjects, it is entirely possible that the ANOVAs performed in the previous step would show no significant difference between the normal and harder VR task. However, a correlation would show that the psychophysiological response to a task does depend on the subject's arousal, valence, or performance during the task.

It should be noted that four subjects from the stroke group (including the two with dysphasia) and two from the control group reported no changes in *valence* or *arousal* during the experiment. Such a lack of changes in self-reported *arousal* is likely to be the result of a misunderstanding of the SAM. In fact, during the final informal interview, three of these six subjects mentioned how active they were during the tasks compared to the baseline period. Thus, these subjects' SAM results were considered unreliable and discarded.

The threshold for significance was always set at $p = 0.05$. Due to space constraints, most results with $p > 0.05$ are not reported. For significant effects in ANOVA, effect size for a factor is also reported as partial η^2 (the proportion of total variability attributable to the factor, excluding other factors from the total nonerror variation) [30]. The Sidak correction for multiple comparisons was used for all post-hoc tests. The Huynh-Feldt correction was used in cases of violations of sphericity in ANOVA. The Kolmogorov-Smirnov test with Lilliefors' modification was used to test for normality.

III. Results

A. Performance and Work

In the normal VR task, the stroke group caught 63% of all balls placed and placed 51% of all balls into the basket while the control group caught 84% of balls and placed 72% of balls into the basket. In the harder VR task, the stroke group caught 48% of balls and placed 36% of balls into the basket while the control group caught 51% of balls and placed 44% of balls into the basket. As previously stated, these percentages were calculated only for subjects who did not receive the relevant haptic support.

In the two levels of the VR task, there was a main effect of task difficulty (normal versus harder task) on *percentage of caught balls* ($p < 0.001$, partial $\eta^2 = 0.66$), *percentage of balls placed into the basket* ($p < 0.001$, partial $\eta^2 = 0.61$), and total work ($p = 0.002$, partial $\eta^2 = 0.20$). There was also a main effect of group (stroke versus control) on *percentage of caught balls* ($p = 0.018$, partial $\eta^2 = 0.14$), *percentage of balls placed into the basket* ($p = 0.005$, partial $\eta^2 = 0.20$), and *total work* ($p < 0.001$, partial $\eta^2 = 0.36$). There was an effect of interaction between task difficulty and group for *percentage of caught balls* ($p = 0.003$, partial $\eta^2 = 0.22$), *percentage of balls placed into the basket* ($p = 0.004$, partial $\eta^2 = 0.22$), and *total work* ($p = 0.001$, partial $\eta^2 = 0.23$). Post-hoc tests showed that both groups caught fewer balls (stroke: $p = 0.002$; control: $p < 0.001$) and placed fewer balls into the basket (stroke: $p = 0.02$; control: $p < 0.001$) in the harder VR task. The control group performed less *total work* in the harder VR task than in the normal VR task (46.9 ± 17.4 J versus 38.2 ± 13.7 J, $p < 0.001$) while the difference was not significant for the stroke group (30.1 ± 9.2 J versus 30.5 ± 8.2 J).

In the Stroop task, the control group had a higher *percentage of correct answers* (stroke: 93.9 ± 5.6%; control: 97.9 ± 4.9%; $p = 0.007$) and lower *mean answer time* (stroke: 2.8 ± 1.6 s; control: 2.0 ± 1.2 s; $p = 0.016$).

B. Temporal Changes of Psychophysiological Signals

The usefulness of psychophysiological signals strongly depends on how quickly and how strongly the signals react to stimuli. The physiological signals of a typical control subject during baseline (rest), the physical control task and the normal VR task are shown in Figs. 2, 3 and 4 in order to illustrate how quickly and how much the signals change. For skin conductance (Fig. 2), an increase can be seen during tasks (with the signal responding to the beginning of the task within seconds), and this general increase is gauged by *mean SCL*. Additionally, a greater number of brief increases in the skin conductance signal appear during the two task periods—there is a higher *SCR frequency*. For skin temperature (Fig. 3), there is a slight decrease during the physical control task followed by a return to baseline as well as a larger decrease during the VR task (although temperature only begins decreasing approximately half a minute after the task begins). Heart rate (Fig. 4) shows high variability during both baseline and task periods, but increases during tasks (though not as quickly as skin conductance). Respiratory rate is not shown since most subjects' differences between baseline and task periods were obscured by the high variability of the signal.

C. Baseline-Task Comparisons

Comparison of absolute values of psychophysiological parameters between baseline and task periods was done separately for each task in a mixed-design ANOVA. Table I shows differences between baseline and task for all four task periods and for

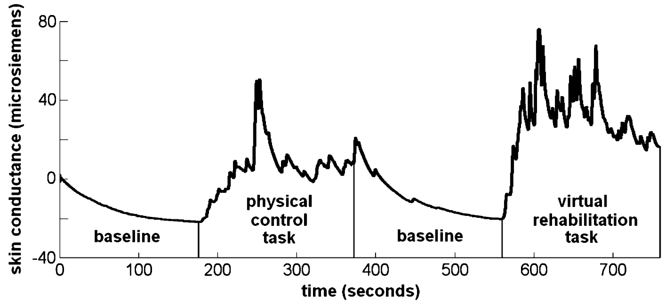

Fig. 2. A typical subject's skin conductance as a function of time during two baseline periods, the physical control task and the virtual rehabilitation task. The initial value was defined as zero.

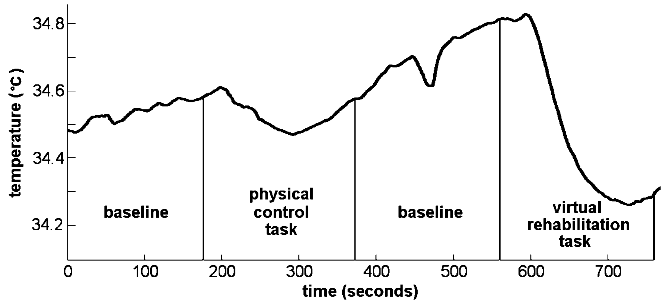

Fig. 3. A typical subject's skin temperature as a function of time during two baseline periods, the physical control task and the virtual rehabilitation task.

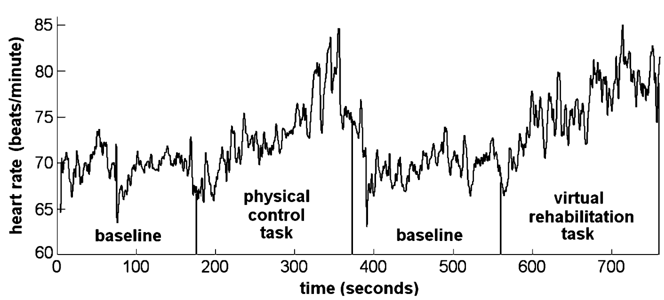

Fig. 4. A typical subject's heart rate as a function of time during two baseline periods, the physical control task and the virtual rehabilitation task.

both groups, as well as the main effect size of time period (baseline versus task).

Significant effects of group (stroke versus control) were found for *arousal* (partial $\eta^2 = 0.12$ in physical control task and 0.11 in normal VR task; higher *arousal* in stroke group in both cases), *mean heart rate* (partial $\eta^2 = 0.23$ in physical control task, 0.18 in normal VR task and 0.15 in harder VR task; higher *mean heart rate* in stroke group in all cases) and *SDNN* (partial $\eta^2 = 0.26$ in physical control task, 0.21 in normal VR task, 0.27 in harder VR task and 0.21 in Stroop task; lower *SDNN* in stroke group in all cases).

Significant effects of group-time period interaction were found for *arousal* (partial $\eta^2 = 0.18$ in harder VR task; the control group showed a larger increase in *arousal*), *mean heart rate* (partial $\eta^2 = 0.16$ in physical control task, 0.15 in harder VR task and 0.14 in Stroop task; the control group showed a larger increase in heart rate in all cases), *RMSSD* (partial $\eta^2 = 0.17$ in physical control task and 0.19 in normal VR task; the control group showed a smaller increase in *RMSSD* in both cases), *mean SCR amplitude* (partial $\eta^2 = 0.12$ in normal VR task and 0.15 in Stroop task; the control group showed a larger increase in SCR amplitude in both cases), and *final skin temperature* (partial $\eta^2 = 0.14$ in harder VR task and 0.11 in Stroop task; the control group showed a larger decrease in skin temperature in both cases).

In order to better illustrate some of the findings, Figs. 5, 6, 7 and 8 show box plots of differences between baseline and task for four physiological parameters: *mean SCL, final skin temperature, mean heart rate and mean SCR amplitude*. A positive value represents an increase from baseline.

D. Comparison of Relative Values Between Tasks

The comparison of relative values was done in a mixed-design ANOVA with one between-subjects factor (group: stroke or control) and one within-subjects factor (task type: physical control task, normal VR task, harder VR task and Stroop task). As the focus was primarily on differences between the physical control task and the two versions of the VR task, differences between the Stroop task and the other three task periods are not reported.

Analysis of self-reported *arousal* found an effect of task type ($p < 0.001$, partial $\eta^2 = 0.24$). Post-hoc tests found higher *arousal* in the harder VR task than in the physical control task ($p < 0.001$) and the normal VR task ($p < 0.001$). There was also an effect of interaction between task type and group ($p < 0.001$, partial $\eta^2 = 0.18$). Post-hoc tests found that the control group showed significant differences between the physical control task, the normal VR task and the harder VR task ($p < 0.05$ for all three pairwise comparisons), but that the stroke group showed no significant differences in *arousal* between these three tasks.

Analysis of *SDNN* found an effect of task type ($p < 0.001$, partial $\eta^2 = 0.15$). However, post-hoc tests found that the only significant differences were between the Stroop task and the other tasks. Similarly, analysis of *RMSSD* found an effect of task type ($p = 0.031$, partial $\eta^2 = 0.08$), but post-hoc tests found no significant differences between the tasks.

Analysis of *respiratory rate variability* found an effect of task type ($p = 0.039$, partial $\eta^2 = 0.07$), but post-hoc tests found no significant differences between the tasks. There was also an effect of interaction between task type and group ($p = 0.024$, partial $\eta^2 = 0.08$). Post-hoc tests showed a difference between the physical control task and the harder rehabilitation task in the control group, but no difference in the stroke group.

Analysis of *mean SCL* (Fig. 5) found an effect of task type ($p < 0.001$, partial $\eta^2 = 0.18$). Post-hoc tests showed significant differences between all three tasks ($p < 0.05$ for all three pairwise comparisons).

Analysis of *mean SCR amplitude* (Fig. 8) found an effect of task type ($p = 0.049$, partial $\eta^2 = 0.08$), but post-hoc tests found no significant differences between the tasks. There was also an effect of group ($p = 0.009$, partial $\eta^2 = 0.17$).

Analysis of *final skin temperature* (Fig. 6) found an effect of task type ($p = 0.010$, partial $\eta^2 = 0.10$). Post-hoc tests found that temperature was lower in the normal VR task than in the physical control task ($p = 0.010$). There was also an effect of group ($p = 0.001$, partial $\eta^2 = 0.24$) and an effect of interaction

TABLE I
DIFFERENCES BETWEEN BASELINE AND TASK (MEAN ± STANDARD DEVIATION) AND MAIN EFFECT SIZE OF BASELINE-TASK

	physical control task			normal VR task			harder VR task			Stroop task		
	stroke	control	p. η^2	stroke	control	p. η^2	stroke	control	p. η^2	stroke	control	p. η^2
valence [scale 1-9]	-0.4±1.0	0.1±0.6	0.00	-0.2±1.2	0.1±0.9	0.00	-0.5±1.6	-0.1±1.4	0.04	-0.2±1.4	0.1±0.7	0.00
arousal [scale 1-9]	1.5±1.8	1.0±1.6	0.35**	1.5±1.9	2.4±1.4	0.59**	1.9±1.7	3.5±1.7	0.72**	1.0±2.1	1.6±1.9	0.31**
mean heart rate [bpm]	2.7±2.4	4.9±2.7	0.70**	4.2±3.8	6.6±5.3	0.59**	4.0±4.2	7.9±5.1	0.63**	2.6±2.2	7.2±7.9	0.43**
SDNN [% of baseline]	12±35	25±49	0.13*	12±34	15±41	0.04*	8±34	27±58	0.08	-19±21	-7±26	0.27**
RMSSD [%]	10±46	-7±31	0.06	26±85	7±35	0.03	12±53	11±38	0.01	-3±24	-8±34	0.09
mean resp. rate [bpm]	1.7±2.8	3.0±2.3	0.47**	2.1±2.9	2.4±2.8	0.39**	1.7±2.7	2.2±3.4	0.30**	2.6±2.0	3.6±2.3	0.69**
resp. rate var. [%]	76±115	20±73	0.05	45±117	64±98	0.11*	52±107	89±99	0.24**	26±92	15±79	0.01
mean SCL (µS)	8±19	12±13	0.31**	25±32	24±21	0.47**	31±39	33±28	0.48**	22±31	22±21	0.43**
SCR frequency [%]	41±68	128±325	0.34**	91±136	82±148	0.57**	95±144	91±165	0.41**	172±456	164±351	0.65**
mean SCR amp. (µS)	-0.6±4.9	1.0±1.3	0.00	0.7±1.5	2.4±2.4	0.38**	0.8±1.4	1.7±1.9	0.37**	0.3±1.4	1.6±2.2	0.22**
final temperature [K]	0.2±0.6	-0.2±0.4	0.01	-0.3±0.6	-0.4±0.6	0.24**	0.2±0.4	-0.8±0.9	0.25**	-0.1±0.8	-0.9±1.0	0.24**

bpm = beats per minute or breaths per minute. %= percentage of baseline value. μS = microsiemens. K = kelvin.
p. η^2 = partial eta − squared value for main effect of time period (baseline versus task). Asterisks indicate significance of effect at $p < 0.05(^*)$ or $p < 0.01(^{**})$.
Main effect size of group (stroke versus control) and interaction effects are listed separately in text.

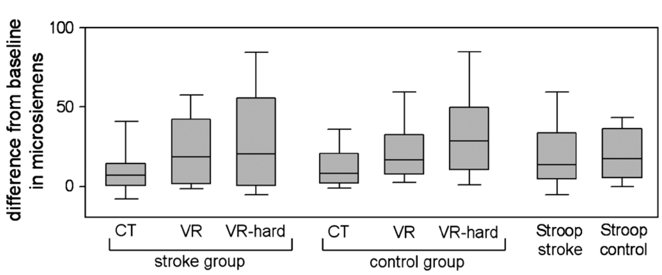

Fig. 5. Changes in *mean SCL* as a response to different tasks. CT = physical control task, VR = normal virtual rehabilitation task, VR-hard = harder virtual rehabilitation task.

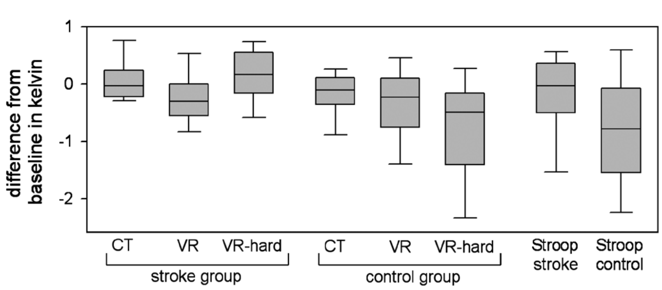

Fig. 6. Changes in *final skin temperature* as a response to different tasks. CT = physical control task, VR = normal virtual rehabilitation task, VR-hard = harder virtual rehabilitation task.

between task type and group ($p = 0.005$, partial $\eta^2 = 0.11$). Post-hoc tests found that, in the stroke group, temperature in the normal VR task was lower than in both the physical control task and the harder VR task. In the control group, temperature in the harder VR task was lower than in the physical control task and the normal VR task.

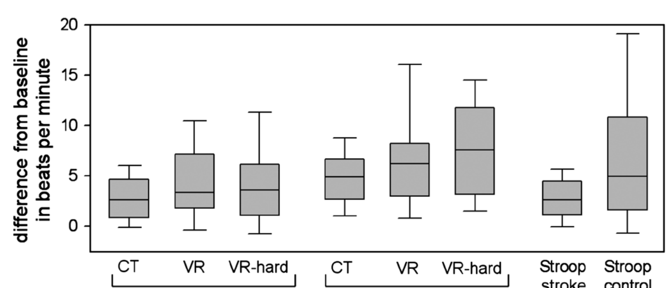

Fig. 7. Changes in *mean heart rate* as a response to different tasks. CT = physical control task, VR = normal virtual rehabilitation task, VR-hard = harder virtual rehabilitation task.

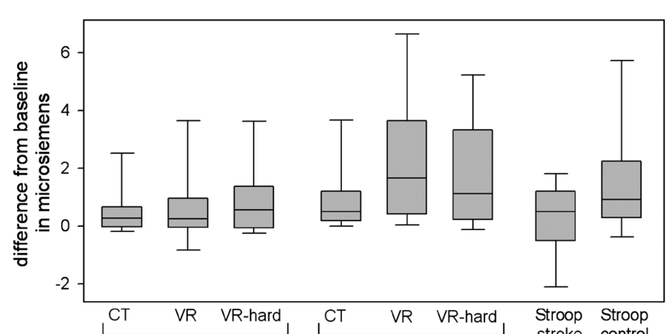

Fig. 8. Changes in *mean SCR amplitude* as a response to different tasks. CT = physical control task, VR = normal virtual rehabilitation task, VR-hard = harder virtual rehabilitation task.

E. Correlations—Arousal, Valence, Performance, and Work

There was no significant correlation between *valence* and *arousal* during any of the tasks in either patients or controls.

For the control group, neither *arousal* nor *valence* was correlated with performance in the VR task or physical control task. In the Stroop task, *valence* was correlated with the *percentage of correct answers* ($\rho = 0.46$, $p = 0.043$).

For the stroke group, *arousal* in the VR task was correlated with the *percentage of balls placed into the basket* ($\rho = 0.66$, $p = 0.029$) while *valence* was correlated with the *percentage of caught balls* ($\rho = 0.56$, $p = 0.045$) and the *percentage of balls placed into the basket* ($\rho = 0.61$, $p = 0.047$). In the Stroop task, *valence* was correlated with the *percentage of correct answers* ($\rho = 0.65$, $p = 0.016$). Neither *arousal* nor *valence* was significantly correlated with work in the physical control task.

F. Correlations—Arousal and Physiology

For the control group, *arousal* was correlated with *SCR frequency* in both the VR task ($\rho = 0.60$, $p = 0.004$) and the Stroop task ($\rho = 0.81$, $p < 0.001$).

For the stroke group, *arousal* was correlated with *SCR frequency* in the VR task ($\rho = 0.59$, $p = 0.019$), with *final skin temperature* in the Stroop task ($\rho = -0.54$, $p = 0.042$) and with *mean SCL* in the physical control task ($\rho = 0.51$, $p = 0.045$).

G. Correlations—Valence and Physiology

For the control group, *valence* was correlated with *SCR frequency* ($\rho = 0.44$, $p = 0.046$) and borderline significantly correlated with *mean respiratory rate* ($\rho = 0.43$, $p = 0.051$) in the VR task. It was not significantly correlated with any physiological parameter in either the Stroop task or the physical control task.

For the stroke group, *valence* was correlated with *final skin temperature* in the VR task ($\rho = 0.62$, $p = 0.017$). In the Stroop task, it was correlated with *mean respiratory rate* ($\rho = 0.59$, $p = 0.026$) and *respiratory rate variability* ($\rho = -0.58$, $p = 0.045$). It was not significantly correlated with any physiological parameter in the physical control task.

H. Correlations—Performance, Work, and Physiology

For the control group, the *percentage of caught balls* in the VR task was correlated with the *RMSSD* measure of heart rate variability ($r = -0.49$, $p = 0.023$). There was no significant correlation between *total work* and physiology in the VR task. In the Stroop task, *mean answer time* was correlated with *respiratory rate variability* ($r = 0.47$, $p = 0.031$). In the physical control task, *total work* was correlated with *SCR frequency* ($r = 0.66$, $p < 0.001$).

For the stroke group, *mean SCL* in the VR task was correlated with *the percentage of balls placed into the basket* ($r = 0.49$, $p = 0.04$) and *total work* ($r = 0.50$, $p = 0.021$). In the Stroop task, *mean answer time* was correlated with *RMSSD* ($r = 0.51$, $p = 0.025$) and *respiratory rate variability* ($r = 0.47$, $p = 0.044$).

IV. Discussion

A. The Stroop Task

In the Stroop task, where no physical activity was required, the expected responses to a cognitive task were noted in both the stroke and control groups: increased SCL (Fig. 5) [12], increased *SCR frequency* [15], [16], decreased *final skin temperature* (Fig. 6) [18], increased *mean heart rate* (Fig. 7) [11], [12], decreased heart rate variability [11], [13] and increased *mean respiratory rate* [11] relative to baseline.

There was a significant main effect of group on physiological responses. The stroke group showed higher *mean heart rate* and lower *SDNN* than the control group during both baseline and task periods, confirming the results of previous studies that have found increased heart rate [31] and decreased heart rate variability [20] after stroke.

In the comparison of physiological parameters between baseline and task, there was also a significant interaction effect. No decrease in *final skin temperature* from baseline to task was noted in the stroke group (Fig. 6), and the control group showed a larger increase in *mean heart rate* than the stroke group (Fig. 7). This shows that the stroke group exhibits different physiological responses even to a cognitive task with no physical activity.

B. Comparison of Relative Values Between Tasks

Mean SCL was the only parameter that showed a significant difference between the physical control task, the normal VR task, and the harder VR task (Fig. 5). Additionally, it showed no significant effect of group or group-task interaction. Since skin conductance is a well-documented indicator of arousal [12], [15], [16] and also showed large, rapid changes in response to our tasks (Figs. 2 and 5), it seems to be the most effective indicator of arousal in virtual rehabilitation. However, another question needs to be answered to ensure reliability. Since the harder VR task was always performed after the normal VR task, was there an influence of task order? In other words, would SCL have kept increasing even if the normal VR task had been followed by an easier task? At the end of a task and beginning of a baseline period, SCL decreases again (Fig. 2). Additionally, a qualitative examination of the recorded signals showed that most subjects' skin conductance reached a plateau within approximately a minute and then stayed at that plateau or even decreased slowly. Such a plateau can be seen for both task periods in Fig. 2, and a slow drift can be observed for the VR task in the same figure. However, some subjects did show a constant rise in skin conductance throughout the VR task, so we cannot rule out the influence of task order.

Another interesting parameter is *final skin temperature*, which showed significant task-group interaction effects. In the stroke group, it decreased from baseline during the normal VR task but was actually higher than baseline in the harder VR task (Fig. 6). In the control group, however, the greatest decrease was in the harder VR task. This could be explained by a transient nature of skin temperature changes [32]. In several subjects, skin temperature began to return to baseline levels toward the end of a task period, thus suggesting that skin temperature will return to baseline after several minutes' exposure to a constant stimulus. This can be seen in the physical control task in Fig. 3. The control group likely does not find the normal VR task to be difficult (in the final informal interview, many stated that they found it too easy) and thus shows a larger decrease in temperature during the harder VR task where subjects need to focus more. The stroke group, however, is already challenged by the normal VR task. If the harder VR task is not much more challenging, temperature may return to baseline. On the other hand, it is possible that the harder VR task is so difficult that some subjects in the stroke group simply give up and no longer try hard. This was also noted in the informal interviews.

One factor that may have blurred differences between the two VR tasks was the verbal assistance of the occupational therapist. In the harder VR task, the stroke group may have relied on the therapist more than the control group. Nonetheless, it was impossible to carry out the study without the therapist's verbal advice since several subjects in the stroke group needed guidance to perform the task properly.

C. Correlations—Arousal, Valence, Performance, and Work

In both groups, *valence* was correlated with task success (*percentage of correct answers*) in the Stroop task. However, only the stroke group showed a correlation between *valence* and task success in the VR task (*percentage of balls placed into the basket*). A possible explanation is that the stroke group finds the task challenging and is thus pleased by success while the control group does not find the task difficult and is thus less concerned about performance.

In the VR task, the stroke group's *arousal* was correlated with the *percentage of balls placed into the basket*. The lack of a correlation between *arousal* and performance in the control group could once again be explained by the fact that control subjects likely did not find the task to be difficult.

D. Correlations—Arousal and Physiology

Looking first at the Stroop task, which requires no physical effort, there was a significant correlation between self-reported *arousal* and *SCR frequency* in the control group. This is in agreement with previous studies that have found SCR frequency to be a good indicator of arousal [15], [16]. However, there was no significant correlation between *arousal* and *SCR frequency* in the stroke group. There was, however, a correlation between *arousal* and *final skin temperature*. While this is also in agreement with studies that found connections between skin temperature and cognitive workload [18], it is interesting that neither group shows both correlations.

In the VR task, both groups showed a correlation between *arousal* and *SCR frequency*. This raises the question of why the stroke group showed a correlation between *arousal* and *SCR frequency* in the VR task, but not the Stroop task. Additionally, in the physical control task, neither group showed a significant correlation between *SCR frequency* and *arousal*. It is possible that, during a task that requires only physical load, *SCR frequency* is not a good measure of *arousal*. Still, of all tested parameters, *SCR frequency* appears to be the most reliable indicator of self-reported arousal. It should thus be possible to use it to control arousal in closed-loop biocooperative systems using methods previously developed for other applications [15].

Mean SCL, the only physiological parameter that showed a significant difference between the different tasks, was only correlated with *arousal* in the physical control task (and only for the stroke group).

E. Correlations—Valence and Physiology

The stroke group's *valence* was correlated with *mean respiratory rate* and *respiratory rate variability* in the Stroop task and with *final skin temperature* in the VR task. Evidence does exist for connections between respiratory variability and anxiety [33] as well as between skin temperature and anxiety [19]. However, respiratory rate has mainly been associated with arousal and cognitive workload. Similarly, the control group's *valence* was correlated with *SCR frequency*, which is a documented indicator of arousal rather than valence.

Responses of the autonomic nervous system appear to be better at indicating arousal than valence. This is to be expected. Skin conductance is regulated exclusively by the sympathetic branch of the autonomic nervous system and is thus poor at distinguishing different levels of valence [14]. Connections between heart rate and valence are, at the moment, controversial (see [14] for examples). Similarly, while some studies have reported a connection between skin temperature and tension/anxiety, others have found that skin temperature is also primarily regulated by the sympathetic branch of the autonomic nervous system [32] and a better indicator of arousal. Respiratory variability may be an indicator of valence, but this is a complex issue since studies have found different respiratory variability responses to different negative emotions (see [33] for examples).

F. Correlations—Performance, Work, and Physiology

For the control group, there was a correlation between heart rate variability and *percentage of caught balls* in the VR task. The lower the *RMSSD*, the more balls the subject caught. Since decreases in heart rate variability have been linked to increased cognitive workload [11], [13], it is probable that subjects who concentrated harder also performed better.

Similarly, a correlation was found between the control group's *mean answer time* and *respiratory rate variability* in the Stroop task. Since decreases in respiratory variability have been linked to increased cognitive load [17], it is again likely that subjects who focused harder were able to answer faster.

In the physical control task, the control group showed a correlation between *SCR frequency* and *total work*. This may indicate an influence of physical activity on skin conductance. Such an influence is expected (exertion would cause sweating), but would likely make it more difficult to separate the physiological effects of cognitive and physical load.

For the stroke group, *mean answer* time in the Stroop task was also correlated with heart rate variability (*RMSSD*) and *respiratory rate variability*. The reasoning for this is identical to the reasoning above for the control group. In the VR task, *mean SCL* was correlated with the *percentage of balls placed into the basket*. Although we found no significant correlation between *arousal* and *mean SCL*, skin conductance is a known indicator of *arousal*, and arousal is also significantly correlated with the *percentage of balls placed into the basket*. *Mean SCL* was also correlated with *total work*. This may also indicate an influence of physical activity on skin conductance.

G. Study Limitations

Three limitations of our study should be highlighted: limitations of self-report measures, lack of task order randomization, and potential physiological effects of drugs.

First, our study found some connections between self-report measures and physiological responses that were at odds with previous research. For instance, correlations were found between valence and skin conductance even though skin conductance has been shown to be almost exclusively affected by arousal. Additionally, there were some discrepancies between physiological and self-report measures. However, these issues are not as problematic as they may appear. Several other studies have found only weak associations between physiological responses and self-reported emotions [34]. Some studies even found that subjects are sometimes not aware of their own emotions or are simply unwilling to report them [35]. This was noted in our study as well. Several subjects commented that they found certain tasks to be highly engaging, but did not report any increase in *arousal* on the SAM. Others reported the same *valence* for the entire session despite obvious frustration and annoyance (expressed by, for example, cursing and annoyed facial expressions). Thus, imperfect connections between self-report measures and psychophysiological responses should be taken not only as a limitation of our study, but also as an unavoidable facet of research in human–robot interaction.

Second, the order in which tasks were performed was not randomized. Thus, there may have been an influence of task order on parameters such as *mean SCL* and *final skin temperature*. While we acknowledge that a fully randomized task order would have been preferable from a methodological viewpoint, we were aware in advance that the number of subjects would be limited. As stroke patients are already likely to exhibit large intersubject variability due to impairments of cognitive and motor ability, it was decided in the planning phase to avoid task order randomization since it would further increase variability and potentially obscure important results. The decision was thus made for purely practical reasons.

Finally, several subjects had received drugs that may have affected psychophysiological responses. For instance, sedatives have been shown to affect skin conductance [36]. Antiepileptics and antipsychotics may have also had an effect. There may have been an influence of secondary stroke prevention drugs, but since these are commonly used in stroke rehabilitation, their effects cannot be avoided and could be taken as inherent in that subset of the population.

V. Conclusion

Of the examined psychophysiological measures, skin conductance appears to be the most useful for patient state assessment. Skin conductance level differentiated between the physical control task, the normal VR task, and the harder VR task while skin conductance response frequency was correlated with self-reported arousal for both the stroke group and the control group. Additionally, skin conductance sensors are very easy to use. Though some issues remain, skin conductance seems to offer great potential for biocooperative rehabilitation robotics and could be used to control a person's level of arousal in closed-loop systems.

Skin temperature, which is also simple to use, unfortunately shows different responses in the stroke and control groups. It also responds to stimuli much more slowly than skin conductance. Heart rate offers uncertain results with regard to psychophysiology, but could at least be used as a measure of physical effort (in which case, a simpler measuring method would suffice). As evidenced by the results from the Stroop task, the stroke group's physiological responses are indeed influenced by psychological activity and thus cannot be only a result of the physical activity required for motor rehabilitation.

While correlations with self-reported arousal and differences between difficulty levels were found, it is not certain whether these results are repeatable. For instance, if physiological responses are affected by novelty and unfamiliarity with the task, they may be weaker if measured a second time. This could be a problem in therapy, which generally consists of many sessions. Of course, since many different virtual scenarios are available for rehabilitation, it would be possible to vary scenarios between sessions in order to keep therapy interesting. Additionally, researchers wishing to use autonomic nervous system responses for patient state assessment should bear in mind that these responses are reliable indicators of arousal, but that connections with valence are uncertain. While it is possible to gauge whether a subject is mentally active, it is harder to tell whether he or she is feeling well. The issues of repeatability and valence measurement certainly warrant further study before psychophysiological responses are used in clinical settings.

Acknowledgment

The authors would like to thank all the volunteers who participated in this study. Special thanks go to N. Goljar, MD, M. Javh, OT, J. Poje, OT, and J. Ocepek, OT, of the Institute for Rehabilitation of the Republic of Slovenia, for their valuable assistance. Moog FCS graciously loaned one of the two HapticMasters used in our research.

References

[1] H. I. Krebs, B. T. Volpe, D. Williams, J. Celestino, S. K. Charles, D. Lynch, and N. Hogan, "Robot-aided neurorehabilitation: A robot for wrist rehabilitation," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 15, no. 3, pp. 327–335, Sep. 2007.

[2] T. Nef, M. Mihelj, G. Colombo, and R. Riener, "ARMIN—A robot for patient-cooperative arm therapy," *Med. Biol. Eng. Comput.*, vol. 45, pp. 887–900, Sep. 2007.

[3] W. S. Harwin, J. L. Patton, and V. R. Edgerton, "Challenges and opportunities for robot-mediated neurorehabilitation," *Proc. IEEE*, vol. 94, pp. 1717–1726, Sep. 2006.

[4] G. C. Burdea and P. Coiffet, *Virtual Reality Technology*, 2nd ed. Hoboken, NJ: Wiley-Interscience, 2003, pp. 304–314.

[5] M. K. Holden, "Virtual environments for motor rehabilitation: Review," *CyberPsychol. Behav.*, vol. 8, pp. 187–211, Jun. 2005.

[6] D. J. Myall, M. R. MacAskill, P. R. Davidson, T. J. Anderson, and R. D. Jones, "Design of a modular and low-latency virtual-environment platform for applications in motor adaptation research, neurological disorders and neurorehabilitation," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 16, no. 3, pp. 298–309, Jun. 2008.

[7] I. Robertson and J. Murre, "Rehabilitation of brain damage: Brain plasticity and principles of guided recovery," *Psychol. Bull.*, vol. 125, pp. 544–575, Sep. 1999.

[8] N. Maclean, P. Pound, C. Wulfe, and A. Rudd, "The concept of patient motivation: A qualitative analysis of stroke professionals' attitudes," *Stroke*, vol. 33, pp. 444–448, Feb. 2002.

[9] R. L. Mandryk and S. M. Atkins, "A fuzzy physiological approach for continuously modeling emotion during interaction with play technologies," *Int. J. Hum. Comput. Stud.*, vol. 65, pp. 329–347, Apr. 2007.

[10] F. H. Wilhelm, M. C. Pfaltz, J. J. Gross, I. B. Mauss, S. I. Kim, and B. K. Wiederhold, "Mechanisms of virtual reality exposure therapy: The role of the behavioral activation and behavioral inhibition systems," *Appl. Psychophys. Biof.*, vol. 30, pp. 271–284, Sep. 2005.

[11] J. A. Veltman and A. W. Gaillard, "Physiological workload reactions to increasing levels of task difficulty," *Ergonomics*, vol. 41, pp. 656–669, May 1998.

[12] C. Collet, P. Averty, and A. Dittmar, "Autonomic nervous system and subjective ratings of strain in air-traffic control," *Appl. Ergon.*, vol. 40, pp. 23–32, Jan. 2009.

[13] R. W. Backs, J. K. Lenneman, J. M. Wetzel, and P. Green, "Cardiac measures of driver workload during simulated driving with and without visual occlusion," *Hum. Factors*, vol. 45, pp. 525–538, Jan. 2003.

[14] C. Peter and A. Herbon, "Emotion representation and physiology assignments in digital systems," *Interact. Comput.*, vol. 18, pp. 139–170, Mar. 2006.

[15] A. Haarmann, W. Boucsein, and F. Schaefer, "Combining electrodermal responses and cardiovascular measures for probing adaptive automation during simulated flight," *Appl. Ergon.*, vol. 40, pp. 1026–1040, Nov. 2009.

[16] R. Nikula, "Psychological correlates of nonspecific skin conductance responses," *Psychophysiology*, vol. 28, pp. 86–90, Jan. 1991.

[17] F. Boiten, "Component analysis of task-related respiratory patterns," *Int. J. Psychophysiol.*, vol. 15, pp. 91–104, Sep. 1993.

[18] M. Ohsuga, F. Shimono, and H. Genno, "Assessment of phasic work stress using autonomic indices," *Int. J. Psychophysiol.*, vol. 40, pp. 211–220, Apr. 2001.

[19] B. C. Min, S. C. Chung, S. J. Park, C. J. Kim, M.-K. Sim, and K. Sakamoto, "Autonomic responses of young passengers contingent to the speed and driving mode of a vehicle," *Int. J. Ind. Ergonom.*, vol. 29, pp. 187–198, Apr. 2002.

[20] J. T. Korpelainen, K. A. Sotaniemi, and V. V. Myllylä, "Autonomic nervous system disorders in stroke," *Clin. Auton. Res.*, vol. 9, pp. 325–333, Dec. 1999.

[21] M. F. Folstein, S. E. Folstein, and P. R. McHugh, "'Mini-mental state': A practical method for grading the cognitive state of patients for the clinician," *J. Psychiatr. Res.*, vol. 12, pp. 189–198, Nov. 1975.

[22] D. Kidd, G. Stewart, J. Baldry, J. Johnson, D. Rossiter, A. Petruckevitch, and A. J. Thompson, "The functional independence measure: A comparative validity and reliability study," *Disabil. Rehabil.*, vol. 17, pp. 10–14, Jan. 1995.

[23] R. Q. van der Linde and P. Lammertse, "HapticMaster—A generic force controlled robot for human interaction," *Ind. Robot*, vol. 30, pp. 515–524, Oct. 2003.

[24] J. Podobnik, M. Mihelj, and M. Munih, "Upper limb and grasp rehabilitation and evaluation of stroke patients using HenRiE device," in *Virtual Rehabil.*, New York, 2009, pp. 173–178.

[25] J. Ziherl, J. Podobnik, M. Sikic, and M. Munih, "Pick to place trajectories in human arm training movement," *Technol. Health Care*, vol. 17, pp. 323–335, Dec. 2009.

[26] C. M. MacLeod, "Half a century of research on the Stroop effect: An integrative review," *Psychol. Bull.*, vol. 109, pp. 163–203, Mar. 1991.

[27] M. M. Bradley and P. J. Lang, "Measuring emotion: The self-assessment manikin and the semantic differential," *J. Behav. Ther. Exp. Psychiatr.*, vol. 25, pp. 49–59, Mar. 1994.

[28] A. Mehrabian, "Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament," *Curr. Psychol.*, vol. 14, pp. 261–292, Dec. 1996.

[29] Task Force of the European Society of Cardiology, North American Society of Pacing and Electrophysiology, "Heart rate variability: Standards of measurement, physiological interpretation, and clinical use," *Eur. Heart J.*, vol. 17, pp. 354–381, Mar. 1996.

[30] J. Cohen, "Eta-squared and partial eta-squared in fixed factor ANOVA designs," *Educ. Psychol. Meas.*, vol. 33, pp. 107–112, Apr. 1973.

[31] N. Lakusic, D. Mahovic, and T. Babic, "Gradual recovery of impaired cardiac autonomic balance within first six months after ischemic cerebral stroke," *Acta Neurol. Belg.*, vol. 105, pp. 39–42, Mar. 2005.

[32] A. Kistler, C. Mariauzouls, and K. von Berlepsch, "Fingertip temperature as an indicator for sympathetic responses," *Int. J. Psychophysiol.*, vol. 29, pp. 35–41, Jun. 1998.

[33] I. van Diest, J. F. Thayer, B. Vandeputte, K. P. van de Woestijne, and O. van den Bergh, "Anxiety and respiratory variability," *Physiol. Behav.*, vol. 89, pp. 189–195, Sep. 2006.

[34] P. J. Feldman, S. Cohen, S. J. Lepore, K. A. Matthews, T. W. Kamarck, and A. L. Marsland, "Negative emotions and acute physiological responses to stress," *Ann. Behav. Med.*, vol. 21, pp. 216–222, Sep. 1999.

[35] A. Schwerdtfeger, "Predicting autonomic reactivity to public speaking: Don't get fixed on self-report data!," *Int. J. Psychophysiol.*, vol. 52, pp. 217–224, May 2004.

[36] M. Siepmann, B. Heine, A. Kluge, T. Ziemssen, M. Mück-Weymann, and W. Kirch, "The effects of lorazepam on skin conductance responses to aversive stimuli in healthy subjects," *Clin. Auton. Res.*, vol. 17, pp. 160–164, Jun. 2007.

Domen Novak received the degree in electrical engineering, in 2008, from the University of Ljubljana where he is now enrolled as a Ph.D. student.

His research interests include applied psychophysiology, pattern recognition, and biomedical signal processing.

Jaka Ziherl received the degree in electrical engineering, in 2008, from the University of Ljubljana where he is now a Ph.D. student.

His research interests include haptic interfaces, rehabilitation robotics, and virtual reality.

Andrej Olenšek graduated in 2003 from the University of Ljubljana and defended his Ph.D. thesis in electrical engineering at the same university, in 2008.

His primary research interest is multimodal human-robot interaction for rehabilitation purposes.

Maja Milavec received the M.Sc. degree in psychology, in 2008, from the University of Ljubljana, where she is currently pursuing a second degree in computer science at the same university.

Janez Podobnik received the university degree in electrical engineering and the Ph.D. degree from the University of Ljubljana, in 2004 and 2009, respectively.

He is currently a researcher and teaching assistant with the University of Ljubljana. His main research interests include haptic interfaces, real-time control of robots for virtual-reality-supported rehabilitation and sensory fusion techniques.

Matjaž Mihelj received the D.Sc. degreefrom the Faculty of Electrical Engineering, University of Ljubljana, in 2002.

He was a visiting researcher at Tohoku University, Japan, and at ETH Zurich, Switzerland. Since 2009, he has been affiliated as an Associate Professor at the Faculty of Electrical Engineering, University of Ljubljana. His research interests include man–machine interaction, haptic interfaces, and medical robotics.

Marko Munih received the Ph.D. degree in electrical engineering from the University of Ljubljana.

His research interests include functional electrical stimulation, measurement and control, biomechanics, and robotics. In 1995/1996, he was a visiting scientist at University College London. He is now full professor and Head of the Laboratory for Robotics and Biomedical Engineering at the University of Ljubljana.