2011 IEEE International Conference on Rehabilitation Robotics
Rehab Week Zurich, ETH Zurich Science City, Switzerland, June 29 - July 1, 2011

# Task difficulty adjustment in biocooperative rehabilitation using psychophysiological responses

Domen Novak, Matjaž Mihelj, Jaka Ziherl, Andrej Olenšek, Marko Munih
Faculty of Electrical Engineering
University of Ljubljana
Ljubljana, Slovenia
domen.novak@robo.fe.uni-lj.si

*Abstract*—**This study presents a biocooperative feedback loop where the difficulty of an upper extremity rehabilitation task is adjusted based on four psychophysiological measurements: heart rate, skin conductance, respiration and skin temperature. They are used both by themselves and in combination with task performance and biomechanics. Different variants of linear discriminant analysis are used for data fusion, including a variant that can adjust the fusion rules online and thus gradually adapt to the subject. Both healthy subjects and hemiparetic patients participated in the study. The accuracy rate of the biocooperative controller was defined as the percentage of times it matched the subjects' preferences. Psychophysiological measurements yielded a relatively low accuracy rate by themselves (76.4% for healthy subjects and 68.2% for patients). Task performance, on the other hand, yielded an accuracy rate of approximately 82%. Combining task performance with psychophysiology increased the accuracy rate to 84.7% for healthy subjects and 89.4% for patients. Psychophysiology can thus provide additional information, but factors such as the increased cost and complexity of the system should also be taken into account.**

***Keywords-psychophysiology; rehabilitation robotics; biocooperative robotics; upper extremities***

## I. INTRODUCTION

Over the last decade, robotic interfaces have become increasingly common in motor rehabilitation. Long-term exercise with such devices has been shown to improve motor control, though the effect on functional abilities is still inconsistent [1, 2]. Additionally, they allow objective estimation of the patient's motor performance and abilities [3].

State-of-the art rehabilitation robots are able to not only actively move the patient's limb, but also adapt their robotic assistance to the patient's movement intentions and motor abilities [4, 5]. Recently, this concept of adaptive assistance has been extended to biocooperative robotics, which take into account not only motor abilities, but also psychological factors such as stress and boredom. In biocooperative rehabilitation, exercise parameters are automatically adjusted so that the patient is challenged in a moderate but engaging way without causing undue stress or harm [6]. However, no commonly accepted unobtrusive, real-time method of measuring psychological factors currently exists in rehabilitation. One possibility would be to use psychophysiological measurements: measurements of physiological responses to changes in psychological state. These can be taken without the subject's awareness, providing an objective and unobtrusive method of estimating stress, engagement etc. They have already been used for psychological state assessment in situations such as interaction with mobile robots [7], and have been found to provide useful information about stroke patients' psychological states during robot-aided upper extremity rehabilitation [8]. It should thus be possible to use them in a biocooperative feedback loop in order to obtain information about the patient that cannot be extracted from forces and movements.

Our paper presents a biocooperative controller that adjusts the difficulty of an upper extremity rehabilitation task based on a fusion of four psychophysiological signals (heart rate, skin conductance, respiration and peripheral skin temperature) as well as task performance and biomechanical signals. Discriminant analysis [9] is used for data fusion, and we also propose a method of gradually adapting the system to a particular subject online.

This work was funded by the EU Information and Communication Technologies Collaborative Project MIMICS grant 215756 and additionally supported by the Slovenian Research Agency.

## II. MATERIALS AND METHODS

### A. Hardware

The HapticMaster robot [10] from Moog FCS was used as the haptic interface. Its end-point was equipped with a two-axis gimbal and a passive grasping module. The subject's arm was supported by two cuffs fastened above and below the elbow. A 1.4 m x 1.4 m screen was used to display visual data. Subjects sat approximately 1.25 m in front of the screen.

Physiological signals were sampled at 1.2 kHz using a g.USBamp signal amplifier (g.tec Medical Engineering GmbH). The electrocardiogram (ECG) was recorded using four surface electrodes placed on the torso. Skin conductance was measured using a g.GSR sensor (g.tec). The electrodes were placed on the second and third fingers of the idle hand. Respiratory rate was measured 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 fifth finger of the idle hand.

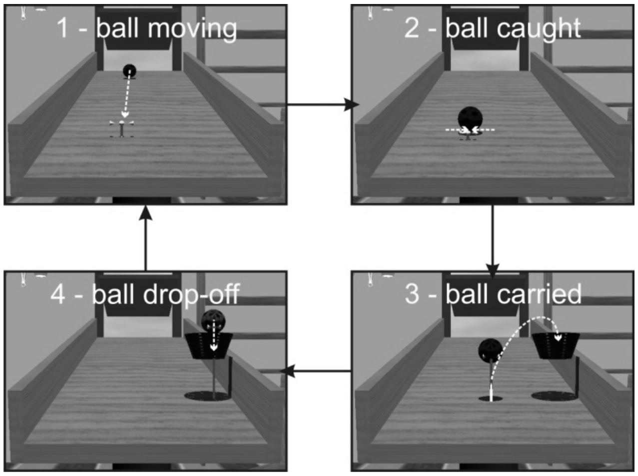


Figure 1. The virtual rehabilitation task. A ball appears on the top of a sloped table (1) and begins to roll down. The subject then catches it (2) and carries it toward a basket that appears above the table (3). Once the ball is above the basket (4), the subject drops it into the basket and a new ball appears.

### B. *Virtual rehabilitation task*

The virtual rehabilitation task was previously used to examine stroke patients' psychophysiological responses to robot-aided upper extremity rehabilitation [8]. The screen shows a table sloped toward the subject. At the beginning, a ball appears at the top of the slope and starts rolling downward. The subject must catch the ball and then place it into a basket that appears above the table. Once the ball is placed into the basket or falls off the table, another ball appears at the top of the table, the basket disappears and the task continues. Screenshots are shown in Fig. 1.

Seven difficulty levels were implemented, with higher difficulties featuring progressively smaller and faster balls. The first level is very easy (the ball is very large and requires approximately fifteen seconds to cross the table), but the seventh is almost impossible (the ball falls off the table in less than three seconds and its radius is one-fifth the radius from the first level). The ultimate goal of the biocooperative feedback loop was to adjust the difficulty level so that the subject is optimally challenged.

### C. *Study protocol*

The study was divided into two phases: the open-loop phase (where difficulty is adjusted according to the subject's preferences) and the closed-loop phase (where difficulty is adjusted automatically by the biocooperative controller). The open-loop phase was conducted first. It was performed first with healthy subjects, then with hemiparetic patients. Since the connections between psychophysiological responses and psychological states are still uncertain, we performed the open-loop phase in order to obtain a large data set for analysis and biocooperative controller training. After training the biocooperative controller on the open-loop data, the controller was tested in the closed-loop phase with both healthy subjects and hemiparetic patients.

The experiment was similar for both phases. It was conducted at the University Rehabilitation Institute of the Republic of Slovenia. Upon arrival, subjects were informed of the purpose and procedure of the study, then gave informed consent. They were seated in front of the robot. The patient's paretic arm or healthy subject's right arm was strapped into the cuffs and grasping device, and the physiological sensors were attached. Subjects were allowed to practice the third level of the task briefly.

After practice, the subject rested for two minutes while baseline psychophysiological measurements were recorded. Then, the task began at level 3, 4 or 5 (chosen randomly). After two minutes at that difficulty level, the task was paused and the subject was asked whether he or she would prefer the difficulty to increase or decrease. The option to stay at the same difficulty level was not given.

In the open-loop phase, once the subject had stated his or her preference, the difficulty changed by one or two levels in the direction preferred by the subject. In the closed-loop phase, the difficulty changed in the direction chosen by the biocooperative controller. After task difficulty was changed, the task continued. The subject went through six two-minute periods in total, with the subject's preference noted and the difficulty changing after each one.

### D. *Subjects*

Twenty-four healthy subjects (20 males, 4 females, age 31.1 ± 10.9 years, age range 21-61) and eleven hemiparetic patients (8 males, 3 females, age 43.2 ± 13.5 years, age range 22-69) participated in the open-loop phase. Ten healthy subjects (9 males, 1 female, age 33.9 ± 12.6 years, age range 22-62) and six hemiparetic patients (4 male, 2 female, age 58.3 ± 6.3 years, age range 54-67) participated in the closed-loop phase. The patients were undergoing rehabilitation at the University Rehabilitation Institute of the Republic of Slovenia.

Patients in the open-loop group were hemiparetic due to intracerebral hemorrhage (3 subjects), cerebral infarction (4 subjects), or surgery of a brain neoplasm (4 subjects). Six had a paretic left arm and five had a paretic right arm. Patients in the closed-loop group were hemiparetic due to subarachnoid hemorrhage (1 subject), intracerebral hemorrhage (2 subjects), cerebral infarction (2 subjects), or surgery of a neoplasm of the brain (1 subject). Three had a paretic left arm and three had a paretic right arm.

### E. *Feature extraction*

Twenty-six features were calculated from the raw signals for each two-minute task period. They are divided into three groups: task performance (4 features), biomechanics (8 features) and psychophysiology (14 features).

*1) Task performance:* Performance features describe how well a subject did and how long he or she had been performing the task. The four features used were the *difficulty level* (1-7), the *time period* (1 – first, 6 – last), the *percentage of caught balls*, and the *percentage of balls placed into the basket*.

*2) Biomechanics:* Biomechanical features describe the kinematics and dynamics of the subjects' actions. The eight features used were *mean absolute force*, *mean absolute velocity*, *mean absolute acceleration*, *total work*, *mean* *frequency of the position signal*, *mean frequency of the velocity signal*, *mean frequency of the acceleration signal*, and *mean frequency of the force signal*.

*3) Psychophysiology:* Four physiological signals were recorded: the electrocardiogram, skin conductance, respiration and skin temperature. From the ECG, the intervals between two normal heartbeats (NN intervals) were extracted. Then, *mean heart rate* as well as several measures of heart rate variability (HRV) were calculated: *the standard deviation of NN intervals* (SDNN), *the square root of the mean squared differences of successive NN intervals* (RMSSD), *the number of interval differences of successive NN intervals greater than 50 ms divided by the total number of NN intervals* (pNN50), *total power in the high-frequency heart rate band*, and *total power in the low-frequency heart rate band*. Details about these measures are available in [11].

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 discrete environmental events. *Mean SCL* and *mean derivative of SCL* were calculated. A SCR is a transient increase in skin conductance whose amplitude exceeds 0.05 µS and whose peak occurs less than five seconds after the beginning of the increase. *SCR frequency* and *mean SCR amplitude* were calculated.

*Mean respiratory rate* and *standard deviation of respiratory rate* were calculated from the respiration signal.

*Final skin temperature* was calculated as the mean temperature during the last five seconds of the period. Additionally, the *mean derivative of skin temperature* was calculated over the entire period.

Due to large intersubject differences, absolute values of psychophysiological features were not used in data fusion. Instead, relative values were calculated by either subtracting the baseline value from the absolute value or by subtracting the baseline value from the absolute value and dividing the result by the baseline value. The second definition was used for *SCR frequency*, *standard deviation of respiratory rate*, and all measures of HRV except *pNN50*.

### F. *Data fusion*

After feature extraction, a number of features are available for each time period. These features must be fused into an estimate of how task difficulty should be changed.

In the open-loop phase, subjects were regularly asked whether they would prefer difficulty to be easier or harder, and their responses were noted. Assuming that the responses were accurate, this gave us a training data set with known inputs (performance, biomechanics and psychophysiology) and known desired outputs (subject's preference). The rules that use the input features to estimate how difficulty should be changed were defined using linear discriminant analysis.

*1) Linear discriminant analysis:* Originally devised by Fisher [9], linear discriminant analysis (LDA) is used to find a linear combination of features which best separate data points into two or more classes. The LDA equation for classification of data into two classes can be written as:

$$D(\boldsymbol{x}) = b + \boldsymbol{w}^{\mathrm{T}} * \boldsymbol{x} \quad (1)$$

$$b = -\boldsymbol{w}^{\mathrm{T}} * 0.5 * (\mu_1 + \mu_2) \quad (2)$$

$$\boldsymbol{w} = (\boldsymbol{S}_1 + \boldsymbol{S}_2)^{-1} * (\mu_1 - \mu_2) \quad (3)$$

where $\boldsymbol{x}$ is the vector of input features, $D(\boldsymbol{x})$ is the discriminant function, $b$ and $\boldsymbol{w}$ are the weights of $D(\boldsymbol{x})$, $\boldsymbol{S}_k$ is the covariance matrix for class $k$ and $\boldsymbol{\mu}_k$ is the mean value for class $k$. $\boldsymbol{x}$ is then assigned to one class if $D(\boldsymbol{x})$ is positive and to the other class if $D(\boldsymbol{x})$ is negative.

In our case, LDA can be used to classify the multiple input features into an estimate of how task difficulty should be changed (easier or harder difficulty). However, there are two problems. First, there are a very large number of available features, some of which may not be relevant. If all are used in discriminant analysis, a very large training set is required to obtain an accurate discriminant function. Two possible solutions are discussed in sections *F.2* and *F.3*.

Second, biomechanical and especially psychophysiological features exhibit high intersubject variability. A discriminant function trained using data from many subjects will be generally accurate, but may fail for some subjects. As a subject exercises, it would be useful for the discriminant function to gradually adapt to that subject and become more accurate. A possible method for this is Kalman adaptive LDA, described in sections *F.4* and *F.5*.

*2) Stepwise linear discriminant analysis:* Stepwise LDA is a variant of LDA where only the most informative features are used. Starting with no features in the function, all the features are evaluated to determine which one best discriminates between classes. That feature is included in the function, and another evaluation is made to determine which feature best discriminates between classes once the information already obtained from the included feature has been taken into account. At each step, a feature already in the function can also be removed if it does not contribute sufficiently. The feature selection process ends when no feature contributes enough additional information to be added. This approach has already been used with psychophysiological data [12]. In our biocooperative feedback loop, we can use stepwise LDA to select the most relevant features for task difficulty adjustment.

*3) Diagonal linear discriminant analysis:* Diagonal LDA is a special case of LDA which ignores correlations between input features. Despite its simplicity, it has proven very effective for classification [13]. In our biocooperative feedback loop, it may be a useful alternative to normal LDA since it usually requires a smaller training set.

*4) Kalman adaptive linear discriminant analysis:* Kalman adaptive LDA (KALDA) [14] is a variant of LDA where the discriminant function is initialized using the training data, then recursively updated online as new information becomes available (after each two-minute task period). In the update process, the weights of the discriminant function are updated using Kalman filtering. For this process, Eqs. (1)-(3) are expanded with [14]:

$$\boldsymbol{H}_k = [1, \boldsymbol{x}_k^T] \quad (4)$$
$$e_k = y_k - \boldsymbol{H}_k * \hat{\boldsymbol{w}}_{k-1} \quad (5)$$
$$Q_k = \boldsymbol{H}_k * \boldsymbol{A}_{k-1} * \boldsymbol{H}_k^T + (1 - UC) \quad (6)$$
$$k_k = \boldsymbol{A}_{k-1} * \boldsymbol{H}_k^T / Q_k \quad (7)$$
$$\hat{\boldsymbol{w}}_k = \hat{\boldsymbol{w}}_{k-1} + k_k * e_k \quad (8)$$
$$\tilde{\boldsymbol{A}}_k = \boldsymbol{A}_{k-1} - \boldsymbol{k}_k * \boldsymbol{H}_k * \boldsymbol{A}_{k-1} \quad (9)$$
$$\boldsymbol{A}_k = \mathrm{trace}(\tilde{\boldsymbol{A}}_k) * UC / p + \tilde{\boldsymbol{A}}_k \quad (10)$$

where $e_k$ is the one-step prediction error, $y_k$ is the current class label, $\boldsymbol{x}_k$ is the current input vector, $\hat{\boldsymbol{w}}_k$ is the state vector ($\hat{\boldsymbol{w}}_k = [b, \boldsymbol{w}^T]$, the estimated weights for the LDA), $Q_k$ is the estimated prediction variance, $\boldsymbol{A}_k$ is the a priori state error correlation matrix, $\tilde{\boldsymbol{A}}_k$ is an intermediate value needed to compute $\boldsymbol{A}_k$, $\boldsymbol{k}_k$ is the Kalman gain, $UC$ is the update coefficient and $p$ is the number of elements of $\hat{\boldsymbol{w}}_k$. The starting values of $\boldsymbol{A}_0$ and $\hat{\boldsymbol{w}}_0$ as well as the optimal value of $UC$ are computed from the training data set.

In our biocooperative feedback loop, the discriminant function is initialized with the training data. The subject performs the task for two minutes, and the system estimates how difficulty should be changed. The subject is then asked whether he or she would prefer the task to be easier or harder, and the Kalman filter recursively updates the function based on the difference between the system's estimate and the subject's response. This update is performed after each task period, allowing the biocooperative feedback loop to gradually adapt to the current subject. KALDA can be used together with stepwise or diagonal LDA.

*5) Unsupervised Kalman adaptive linear discriminant analysis:* A weakness of KALDA is that it is a supervised learning method: as seen in Eq. (5), the subject's preference ($y_k$) is required to update the weights. We used this supervised KALDA to evaluate whether a biocooperative system can adapt to a subject if given accurate data about him/her. However, a supervised approach is inappropriate in practice - if we know exactly what the subject wants, no automated feedback loop is needed.

KALDA was thus modified so that it updates the weights using its own estimate of the subject's preference rather than the subject's actual preference. While this makes KALDA unsupervised, it must be done carefully since updating the weights with an incorrect estimate would make the function worse. Our solution was to generate a simple measure of whether the estimate is 'reliable'. As previously mentioned, the input vector $\boldsymbol{x}$ is assigned to one class if $D(\boldsymbol{x})$ is positive and to the other class if $D(\boldsymbol{x})$ is negative. If the absolute value of $D(\boldsymbol{x})$ is close to zero, the estimate is probably unreliable. We thus only updated the weights if the absolute value of $D(\boldsymbol{x})$ was larger than a certain threshold. The optimal value of this threshold is calculated from the training data set. While this is not an optimal unsupervised algorithm, it allows us to estimate how accurate an adaptive algorithm will be when working with realistically available data.

The final biocooperative feedback loop designed with these data fusion methods is thus shown in Figure 2.

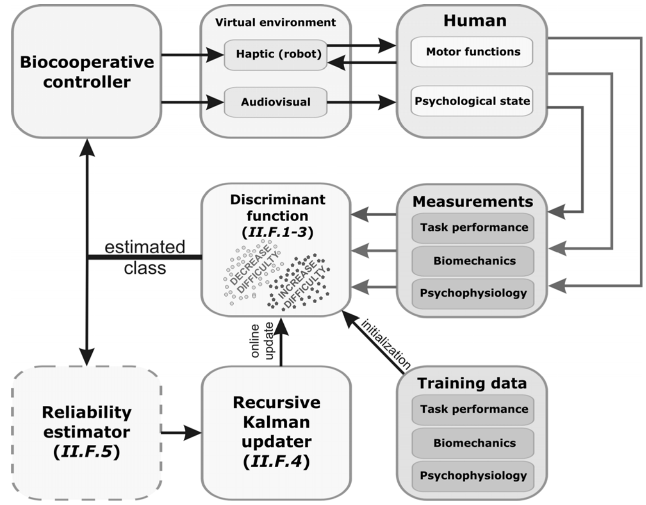


Figure 2. The biocooperative feedback loop. The fusion rules are initialized with prerecorded training data and then recursively updated (with an optional reliability estimator for unsupervised learning) while the subject exercises.

## *G. Open-loop cross-validation*

Data from the open-loop phase (where task difficulty is adjusted according to the subject's preferences) was evaluated using leave-one-out cross-validation. Discriminant functions were created using data from all but one subject, then tested on the remaining subject. This was done as many times as there were subjects. The accuracy rate of a discriminant function was defined as the percentage of times that its estimate matched the subject's preference regarding task difficulty (easier / harder).

Twenty-four discriminant functions were built in total. They varied according to the type of input data (performance, biomechanics, psychophysiology, all) and according to the type of discriminant analysis used (normal, stepwise, diagonal, adaptive, adaptive stepwise, adaptive diagonal LDA). Supervised adaptive LDA (described in section *F.4*) was used in all adaptive cases. The goal was to see how accurate psychophysiological data would be compared to performance and biomechanical data as well as how much the accuracy rate could be improved using stepwise, diagonal or adaptive methods. After calculating accuracy rates for all three variants of supervised adaptive LDA, the most accurate one was used for unsupervised adaptive LDA (described in section *F.5*).

Discriminant functions were first built and cross-validated with data from healthy subjects, then separately built and cross-validated with data from hemiparetic subjects. We also wanted to know which features would be most useful in a discriminant function. The five most relevant features were determined for both healthy subjects and hemiparetic patients as the first five features selected by stepwise LDA.

## *H. Closed-loop validation*

The discriminant function that yielded the highest accuracy rate in open-loop cross-validation was implemented in the closed-loop biocooperative controller. Discriminant functions for closed-loop validation were trained separately for healthy subjects and for patients.

As previously mentioned, the closed-loop measurement protocol was similar to the open-loop protocol. At the end of each period, the biocooperative controller output whether the task difficulty should be increased or decreased. The subject was asked about his or her preference, but task difficulty was adjusted according to the output of the controller. The goal was to show that discriminant analysis can be used online for task difficulty adjustment in a biocooperative feedback loop.

## III. Results

### A. *Open-loop cross-validation*

Accuracy rates (in percentages) for open-loop cross-validation of different types of LDA and different types of data are shown in Table I for both healthy subjects and hemiparetic patients. The highest accuracy rates for nonadaptive and supervised adaptive methods are bolded in each column.

TABLE I. ACCURACY RATES IN OPEN-LOOP CROSS-VALIDATION OF DIFFERENT TYPES OF DISCRIMINANT ANALYSIS ON DIFFERENT INPUT DATA TYPES FOR HEALTHY SUBJECTS (H) AND HEMIPARETIC PATIENTS (P)

| | | psychophysiology | | biomechanics | | performance | | all | |
|---|---|---|---|---|---|---|---|---|---|
| | | **H** | **P** | **H** | **P** | **H** | **P** | **H** | **P** |
| **nonadaptive** | **LDA** | 56.9 | 54.5 | **75.0** | **75.8** | **81.9** | **81.8** | 75.7 | 75.8 |
| | **stepwise** | 56.9 | 60.6 | 73.6 | 75.8 | 81.9 | 81.8 | **84.7** | **89.4** |
| | **diagonal** | **60.4** | **60.6** | 74.3 | 71.2 | 80.6 | 81.8 | 77.8 | 75.8 |
| **adaptive** | **LDA** | 71.5 | 68.2 | 75.7 | **75.8** | **82.6** | **81.8** | 75.7 | 75.8 |
| | **stepwise** | 56.9 | 68.2 | 73.6 | 75.8 | 81.9 | 81.8 | **84.7** | **89.4** |
| | **diagonal** | **76.4** | **68.2** | **80.6** | 71.2 | 82.6 | 81.8 | 83.3 | 76.5 |

Unsupervised adaptive LDA was tested only on psychophysiological data since supervised adaptive LDA offered little increase in accuracy over nonadaptive LDA with other data types (as seen in Table I). For healthy subjects, unsupervised adaptive diagonal LDA achieved an accuracy rate of 70.8% (compared to 76.4% in the supervised approach and 60.4% in the nonadaptive approach). For patients, unsupervised adaptive LDA achieved an accuracy rate of 65.2% (compared to 68.2% in the supervised approach and 60.6% in the nonadaptive approach).

In stepwise LDA for healthy subjects, the first five selected features were the *percentage of caught balls* ($F$ = 129.9), *mean SCR amplitude* ($F$ = 3.9), *pNN50* ($F$ = 4.2), *total power in the low-frequency heart rate band* ($F$ = 2.3) and *mean derivative of skin temperature* ($F$ = 2.1).

In stepwise LDA for hemiparetic patients, the first five selected features were the *percentage of balls placed into the basket* ($F$ = 100.6), *standard deviation of respiratory rate* ($F$ = 3.7*), total power in the high-frequency heart rate band* ($F$ = 7.7), *RMSSD* ($F$ = 5.0) and *final skin temperature* ($F$ = 2.8).

### B. *Closed-loop testing*

As seen in the open-loop phase, the most accurate type of discriminant function was stepwise LDA with all data types. Thus, stepwise LDA was trained on all types of open-loop data and used in closed-loop testing. It yielded an accuracy rate of 88.3% for healthy subjects and 88.9% for patients.

In a follow-up offline analysis, the closed-loop data was also passed through LDA based only on performance data from the open-loop phase. Performance data yielded an accuracy rate of 86.7% for healthy subjects and 83.3% for patients. For an example of psychophysiology increasing accuracy, see Fig. 3. In a second follow-up, the data was passed through supervised adaptive stepwise LDA. However, the adaptive version yielded the same accuracy rates as the nonadaptive one for both healthy subjects and patients.

## IV. Discussion

### A. *Comparison of different data types*

When used by themselves, psychophysiological measurements yielded relatively poor results. While task performance had an accuracy rate of over 80% and biomechanical data had an accuracy rate of over 75% for both healthy subjects and patients even without adaptation, nonadaptive LDA using psychophysiological data yields an accuracy rate of approximately 60%. While adaptive LDA can increase the accuracy rate for psychophysiological data, performance and biomechanical data still yield higher accuracy rates. This suggests that psychophysiological measurements by themselves are not reliable in a biocooperative feedback loop.

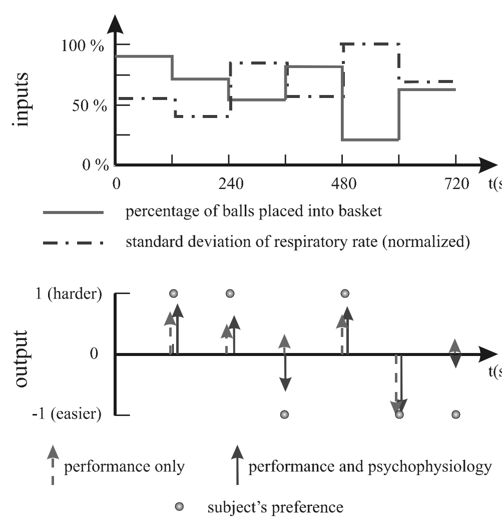


Figure 3. A patient in the closed-loop phase: two input features (one performance, one psychophysiological), the output ($D(\boldsymbol{x})$), and the subject's preferences. High performance and a low *standard deviation of respiratory rate* (even, regular breathing) indicate an easy task. For the first, second, fourth and fifth task periods, difficulty could have been adjusted using task performance alone. In the third period, task performance is moderately high, but breathing becomes uneven, indicating stress. If only performance had been taken into account, the wrong decision would have been made (the patient was performing well, but was stressed and wanted difficulty to decrease). During the last period, performance and psychophysiology are both unreliable, and the patient stated that he would most prefer difficulty to stay the same.

Combining multiple data types using normal LDA or diagonal LDA actually results in a worse accuracy rate than when using only one data type. This probably occurs since the number of features is large and the training set is relatively small, making it difficult to find an accurate discriminant function. Stepwise LDA is more robust since it uses only the most relevant features. While it identifies a task performance feature as the most important, several psychophysiological features are also included in the discriminant function, indicating that they can provide supplementary information.

It is, however, necessary to ask whether the increased accuracy rate offered by psychophysiology is enough to offset the increased cost and complexity of the system. Additionally, user-friendliness should be considered: while the skin conductance and temperature sensors are not problematic, the respiration sensor was considered unpleasant by several patients. A simpler, more attractive option might be to design rehabilitation scenarios with readily available and relevant measures of task performance so that psychophysiological measurements are unnecessary. However, psychophysiology may still be useful when performance measures are unavailable or not connected to the subject's psychological state. It could also be used to change elements other than task difficulty, such as the appearance of a virtual scenario.

### *B. Adaptive linear discriminant analysis*

Supervised adaptive LDA offers little improvement over nonadaptive LDA in the case of performance and biomechanical features. In the case of psychophysiological features, however, supervised adaptive LDA increases the accuracy rate in open-loop cross-validation from 60.4% to 76.4% for healthy subjects and from 60.6% to 68.8% for patients. Unsupervised adaptive LDA also increases the accuracy rate, though to a lesser degree.

In the supervised adaptive LDA, we provided the system with the subject's preference as a 'best-case' scenario. Since this information is not available in practice, we also demonstrated an unsupervised version. Though our modification is not optimal, it is a potential practical implementation of adaptive LDA. Although adaptive LDA offers no advantage when performance and biomechanics are included, it could still be useful for psychophysiological analysis, both in rehabilitation and in other settings.

It should be noted that LDA assumes that the underlying model is linear. This may not be the case in psychophysiology, and alternative classifiers such as support vector machines or neural networks should also be considered. Furthermore, our classification approach takes the subject's own opinion as the reference class. However, the subject's opinion may not be the optimal choice with regard to rehabilitation, and alternative ways of defining the reference class may give different results.

## V. CONCLUSIONS

The four psychophysiological responses evaluated in our study should not be used by themselves in a biocooperative feedback loop for upper extremity rehabilitation since they are noticeably less accurate than other measures. They can be used with measurements of task performance and biomechanics to provide supplementary information. However, factors such as cost, complexity and user-friendliness should also be taken into account in order to determine whether the inclusion of psychophysiological measurements is rational.

## ACKNOWLEDGMENT

The authors would also like to thank Nika Goljar, MD and Metka Javh, OT, of the University Rehabilitation Institute of the Republic of Slovenia for their assistance. Moog FCS loaned one of the HapticMasters used in our research.

## REFERENCES

[1] G. B. Prange, M. J. Jannink, C. G. Groothuis-Oudshoorn and H. J. Hermens, "Systematic review of the effect of robot-aided therapy on recovery of the hemiparetic arm after stroke," J. Rehabil. Res. Dev., vol. 43, pp. 171-184, Mar. 2006.

[2] A. C. Lo, P. D. Guarino, L. G. Richards, J. K. Haselkorn, G. F. Wittenberg, D. G. Federman et al., "Robot-assisted therapy for long-term upper-limb impairment after stroke," New Engl. J. Med., vol. 362, pp. 1772-1783, May 2010.

[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] R. Riener, L. Lünenburger, S. Jezernik, M. Anderschitz, G. Colombo and V. Dietz, "Patient-cooperative strategies for robot-aided treadmill training: First experimental results," IEEE Trans. Neural Syst. Rehabil. Eng., vol. 13, pp. 380-394, Sep. 2005.

[5] E. T. Wolbrecht, V. Chan, D. J. Reinkensmeyer and J. E. Bobrow, "Optimizing compliant, model-based robotic assistance to promote neurorehabilitation," IEEE Trans. Neural Syst. Rehabil. Eng., vol. 16, pp. 286-297, Jun. 2008.

[6] M. Munih, R. Riener, G. Colombo, L. Lünenburger, F. Müller, M. Slater et al., "MIMICS: Multimodal immersive motion rehabilitation of upper and lower extremities by exploiting biocooperation principles", in IEEE Conference on Rehabilitation Robotics 2009, Kyoto, 2009, pp. 127-132.

[7] P. Rani, N. Sarkar, C. A. Smith and L. D. Kirby, "Anxiety detecting robotic system – towards implicit human-robot collaboration," Robotica, vol. 22, pp. 85-95, Jan. 2004.

[8] D. Novak, J. Ziherl, A. Olenšek, M. Milavec, J. Podobnik, M. Mihelj et al., "Psychophysiological responses to robotic rehabilitation tasks in stroke," IEEE Trans. Neural Syst. Rehabil. Eng., vol. 18, pp. 351-361, Aug. 2010.

[9] R. Fisher, "The use of multiple measurements in taxonomic problems", Ann. Eugen., vol. 7, pp. 179-188, 1936.

[10] 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.

[11] Task Force of the European Society of Cardiology and the 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.

[12] G. F. Wilson and C. A. Russell, "Operator functional state classification using multiple psychophysiological features in an air traffic control task", Hum. Factors, vol. 45, pp. 381-389, Jan. 2003.

[13] S. Dudoit, J. Fridlyand and T. P. Speed, "Comparison of discrimination methods for the classification of tumors using gene expression data", J. Amer. Statistical Assoc., vol. 97, pp. 77-87, Mar. 2002.

[14] C. Vidaurre, A. A. Schlögl, R. Cabeza, R. Scherer and G. Pfurtscheller, "Study of on-line adaptive discriminant analysis for EEG-based brain computer interfaces", IEEE Trans. Biomed. Eng., vol. 54, pp. 550-556, Mar. 2007.