Towards relating physiological signals to usability metrics: a case study with a web avatar

Pierfrancesco Foglia
Dipartimento di Ingegneria dell'Informazione
Università di Pisa
Via Diotisalvi 2, Pisa
ITALY
foglia@iet.unipi.it

Michele Zanda
R&D
ION Trading
Via San Martino 54, Pisa
ITALY
michele.zanda@gmail.com

Abstract: - Inferring the user's approval of a graphical interface with non-invasive devices can be effective in improving its design and in implementing adaptive pleasant interactions. This paper investigates how 3 common physiological signals, i.e. skin conductance, heart rate and respiration, can be exploited to infer users' approval of an online avatar embedded in a health care Website. A between group experiment is performed with participants who have the avatar support and participants who do not. During the experiment, skin conductance, hearth rate and respiration were monitored, together with traditional usability metrics (visited pages, completion times, errors, etc). At the end of each experiment, a feedback questionnaire is proposed to infer information related to the user experience, ease of use and approval. Results indicate that the respiration overshoot rate is closely related to the users' appreciation of the avatar based interaction. Further steps of our research will consider improvements in the results by investigating and exploiting mutual effects induced by the multiple collected signals.

Keywords: Smart Interfaces, Human-Computer Interaction, Usability Assessment, Physiology, GSR, HRV, respiration.

1 Introduction

Ensuring a pleasant interaction is a key factor to ensure usability, that ISO standard 9241-11 [1], in the *Ergonomic requirements for office work with visual display terminals*, defines as "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use". The satisfaction aspects indicate qualitatively how pleasant is to use the design [2], and they are usually assessed indirectly through feedbacks from the users [3]. Physiological signals can be used to infer emotions, and they can be used to have a more objective measure of the user satisfaction while interacting with a computer [4], [13]. Succeeding in tracking, via physiological signals, the enjoyment or the negative impact of a graphical interface may constitute a mean to deliver better human-computer interactions. This can be achieved at design time, by improving the design of the interface [5], [56] but also in real-time, defining interfaces that adapt themselves to the measured affective state of the user [6].

While different works have focused on the usefulness of physiological acquisitions to state users' emotional response ([4], [6], [7], [8], [9], [10], [11]) as well as on relating emotions to usability measures [12], [13], generally in specific applications contexts [14], the opportunities offered by physiological signals elaboration are now increased, due the evolution of sensors and signal processing technologies [14]. Such evolution permits to have cheaper and less intrusive sensors in systems with more powerful processing and knowledge discovery power. In this way, it is possible to apply such techniques to the framework of developing and using general purpose, every-day life interactive applications [14, 15].

In this paper, we explore the feasibility of inferring the approval level of users when they are navigating on a Web system, through three physiological signals: Galvanic Skin Response (GSR), Heart Rate Variability (HRV) and Respiration (Figure 1). These three physiological signals can be measured little invasively, have rather cheap sensors, and in the years proved effective to detect lies [16] and in emotion states recognition [11]. As a test case, we choose a web site for the online reservation of clinical exams, with an helping system enriched with avatar support. The avatar has been added as

neurophysiologic studies have proven that face-to-face interactions stimulate human attention [17], [18], [19], and anthropomorphic agents like avatars have a positive influence on users [20], [21], also by increasing positive arousal [22].
Two prototypal Websites were developed, with and without the avatar, and a between two groups experiment was performed. During the experiment, the three physiological signals were collected, as well as a post experiment questionnaire was performed to evaluate users' satisfaction and approval. Results indicate that the respiration overshoots rate (stronger inspirations) has a statistical significant correlation with the degree of approval stated by the participants in the post experiment questionnaire. These finding indicates that, in our test case, such physiological measure can be exploited to infer a usability feature, and eventually utilized to develop an adaptive system. Future works will enforce the results by adopting multiple inputs (channels) classifiers [11], with the aim of investigating the mutual effects induced by the collected signals.

2 Related Work

In recent years, tests to infer the user experience have been largely based on subjective feedback collected by means of surveys and questionnaires [3], [40], [41], [57], [61]. With the purpose to have a more direct assessment of the users' experience, psychophysiology (an area of psychology that measures the individual's physiological responses to infer psychological state and emotions) and affective computing, emerged as alternative approaches for assessing the users' experience and achieving better human-computer interactions [4], [5], [6], [13], [42]. A general framework for achieving such goal can be described as follow [6], [9], [38]: psychophysiology related data (such as facial expressions [43], voice [44], body gestures [45], eye gaze [46], [60] as well as physiological signals [47], [48]), are acquired via a set of sensors; for each of these measures, one or more features are analyzed to derive meaningful information (essential, to correlate them to emotions) and refined metrics. These metrics are interpreted and can be exploited for two purposes: to adapt the system interface in real time [38], [39] or to infer usability issues in the design of a system interface [5].

For what concerns previously related works, they can be grouped in i) studies concerning the adoption of physiological signals for investigating emotions of users while interacting with computer systems and ii) studies concerning the identification of whether and which physiological signals can be correlated to conventional or new usability metrics.

As for the first group, the seminal works by Picard [4], [10] present fundamental theories and methods for affective computing, while the study by Kim & André [6] investigates the potential of physiological signals as reliable channel for emotion recognition. It shows how some physiological signals (electromyogram, electrocardiogram, skin conductivity, and respiration changes) and classifiers can be utilized to recognize music appreciation. [8] utilizes video clips and Hearth Rate variability and Galvanic Skin Response to individuate user emotion via neural network, highlighting the difficulties in obtaining stable data. In [9], via a slide show, it was successful investigated if and how GSR, heart rate and temperature were successful in individuating users emotional state. [11] explores the use of GSR, HRV, and respiration for induced negative emotions recognitions, in an experiment constituted by a shock test, where users hear a loud voice and then a main experiment concerning series of questions. [52] finds that cardiorespiratory activity can be utilized to recognize basic emotion (fear, anger, sadness and happiness).

For what concerns studies focused on usability assessment, Phukan [12] investigates, in an exploratory study, whether and how physiological signals can support the assessment of the user experience, with participants viewing a variety of media. [13] try to correlate GSR and heart rate to the game play experience, in first person shooter games. Results indicate a significant correlation between psychophysiological arousal and the self-reported gameplay experience. [14] presents opportunities and feasibility in applying neurophysiologic analysis methods, based on electroencephalography (EEG), for usability evaluation. They focused on reading and readability and succeeded in discrimination the reading activity (silent, attentive and continuous) and the verification of the decreased readability, associated with the user's increased mental workload. [25] finds evidence that with a decrease of user's task performance level in a video game, the GSR increases. [50] highlights that the visual complexity of web sites has effect on physiology and user's experience. They sense GSR, HR, and EMG (i.e. Electromyography activity of the musculus corrugator supercilii), and show that the increased visual complexity of websites is related to an increased level of arousal measured via GSR, a decreased heart rate, and increased facial muscle tension. Besides it determines increased reaction times that were related to increases in heart rate and electrodermal activity. They conclude that the increased visual complexity of web sites has a detrimental cognitive and emotional impact on users. Haapalainen et al. [51] find that features in the ECG and heat flux help in distinguishing between low and high levels of users' cognitive load, useful in the context of ubiquitous computing applications, to determine when or whether interrupt an user. [53], in the field of adaptive game, shows that GSR positively and significantly correlates with negative game events (frustration events) and continuous failures during game playing can trigger distress (negative stress), which may be detected by combining GSR and heart rate signals.

An important class of interactive systems is represented by Web systems, due to their diffusion and their utilization in everyday life, thanks also to the advent of mobile devices [58] and the porting of many conventional applications to web based interfaces [49], [57]. Usability is a major driver in the design of such systems, because of their wide adoption in many non-classical domains (e.g.: e-government, e-health, telemedicine), and people with different skills and of different ages are supposed to use such systems [54], [55], [57], [58], [59], [63]. Understanding and characterizing the behaviors and the reactions of users when utilizing such systems is relevant to derive design guidelines and test Web systems usability, so findings and methodologies of psychophysiology for human-computer interaction can also be applied to the Web domain. In particular, there is the need for specific experiments that verify if and how research findings can be applied to the Web. Following this path, [50] investigates the effects of visual complexity of Web sites on physiology; Ward & Marsden [56] positively investigated a correlation between badly- and well-designed Websites, and physiological signals (skin conductance, heart rate and finger blood volume). Foglia et al. [22] sensed GSR to have a preliminary assessment of users' cognitive reactions to avatar-based interactions with Public Administration Websites. In this paper, we try to extend such results, by exploring the feasibility of inferring the approval level of users when they are navigating on a Web system.

3 Physiological signals

In our study we considered 3 physiological signal: GSR, HRV and respiration. These three physiological signals can be measured little invasively, have rather cheap sensors, and in the years proved effective in emotion states recognition [11].

The GSR is a measure of human skin conductivity, it is affected when the sympathetic nervous system is active [23], [4] and its use in psychiatric evaluations was early documented by Jung [24]. It represents a good measure of arousal, and its relationship to usability aspects was described in [25], while its relationship to the cognitive workload was shown in [37]. Previous studies [7], [25], [4] indicated that the GSR can be an indicator of arousal/emotional response, but also an indicator of stress/anxiety. [22] showed that, in a system similar to the one studied in this paper, discriminating among the two classes of emotions is difficult by only looking at the GSR signal.

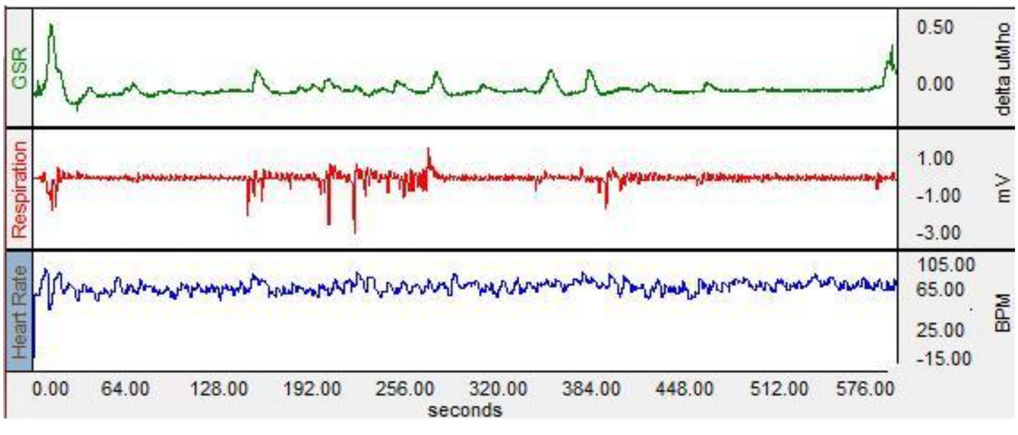

Fig. 1: Physiological signals as gathered by the acquisition system: the Galvanic Skin Response (top), Respiration (middle), and Heart Rate (bottom).

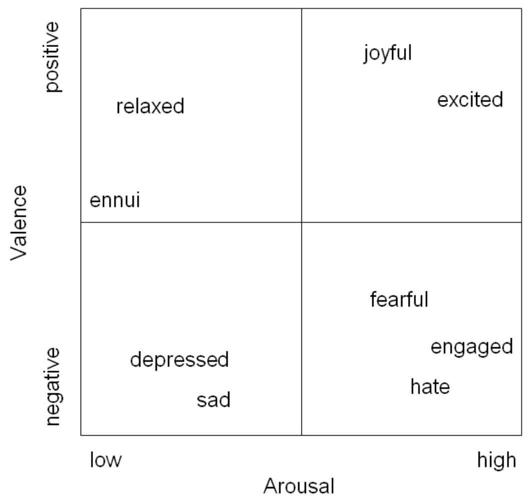

Fig. 2: Emotions classification according to the International Affective Picture System [29], [30]. On the X-axis emotions differ for arousal intensity, on the Y-axis emotions differ for negative or positive valence.

The HRV describes the heart activity when the autonomic nervous system attempts to tackle with the human body demands depending on the stimuli received [11]. In our experiments, the signal was acquired using three pre-gelled disposable electrodes [26], according to the methodology of [27]. Many studies have been performed about HRV and emotions. Phukan [12] reports a study on images perception, with heart rate being "*significantly higher for pleasant and neutral pictures. Unpleasant pictures lead to larger heart rate decelerations*". Other studies, like [8], describe as well how heart rate variability can be considered a good measure of stress and relaxation degree.

We also collected the respiration signal using a velcro stripe and a piezo sensor, with the sensor mounted on the stripe tightened to users' chest. The sympathetic nerve activity has an indirect effect on it [28], and findings by Kim & André [6] have shown that an increasing respiration rate implies a high arousal (joy, anger), while a decreasing respiration rate showed to result in a low arousal (relaxation, bliss).

The aim of our experiments is to investigate whether features in any of the previous described signals can be exploited to infer a positive or negative valence in users' emotions while navigating in the web site described in the next session. A general classification of emotions is reported in Figure 2, [29], [30], with valence levels on the y-axis, and arousal levels on the x-axis. Valence defines whether the emotion is positive and negative, and to what degree. Arousal defines the intensity of the emotion, rating from calm (lowest value) to excited (highest value). In interactive systems, it is important to infer the valence, and not only the arousal, of an induced emotion, as it indicates whether the user is approving the interaction or not.

4 Experiment

4.1 Test case Web site

We set up a Web site for the online reservation of clinical exams. This test case site was developed in two versions, one with an embodied avatar and one without as shown in Figure 3. The avatar was used to support the users during their navigation and to help them accomplishing the assigned tasks. During the execution of the experiment all users' activities were logged into a database. We followed published guidelines to make the avatar pronounce text effectively, and we run the Web site on a multicore server [31], [32], [33]. The face was animated from a 2D photograph. Reallusion Crazytalk was used as morphing software, synchronous with the TTS voice. The TTS engine was Loquendo TTS. Crazytalk provides a client-side plugin, so the only bandwidth overhead for the adopted avatar was the 2D image and the speech to be pronounced.

Fig. 3: The two versions of the Website. On the top the “personal area” without the avatar; on the bottom the “reservation area” with the avatar. The avatar is a simple photo animated with morphing software.

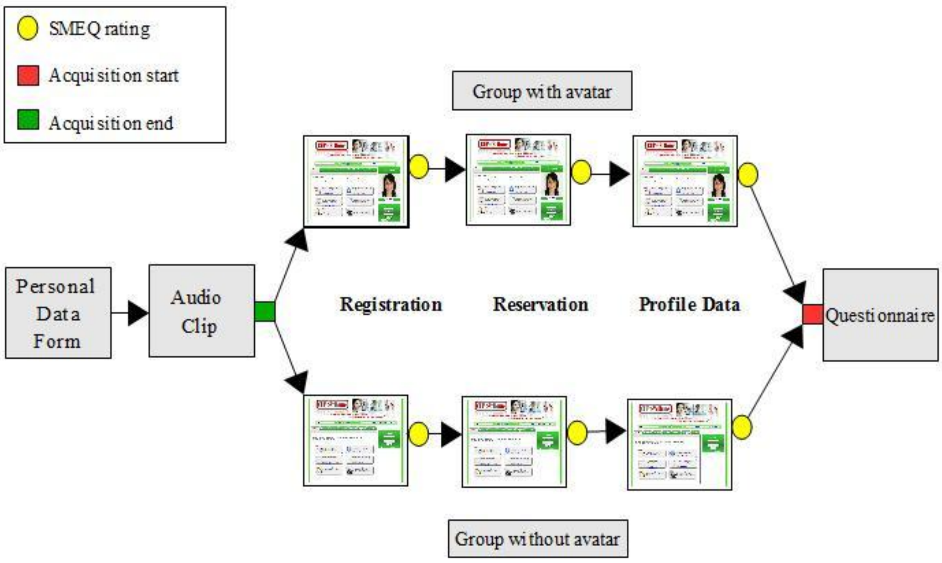

Fig. 4: Usability test procedure with the two groups of users and the two Websites with avatar and without.

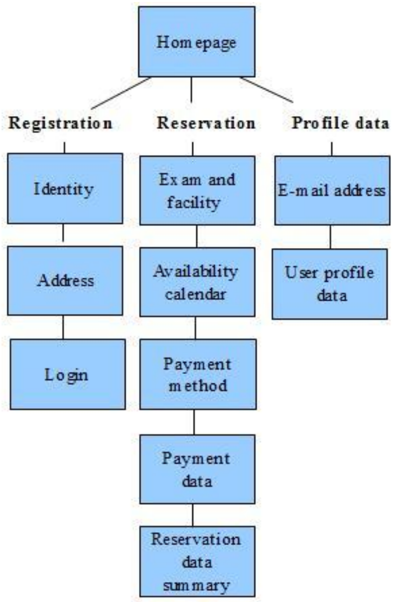

Fig. 5: Sitemap depicting the three tasks that participants had to complete.

4.2 Procedure

For the experiment 35 participants were recruited. Their mean age was 25 and their experience degree with the Internet and computers was intermediate.

All physiological signals were sensed with sensors hardware and software by BioPac Systems Inc.. At the beginning of each test, after all sensors are mounted, users are shown a clip. During such clip the acquisition sensors are calibrated, adjusting the zero-level of the GSR, and testing whether the three mounted sensors (GSR, HRV, respiration) provide proper signal-to-noise ratios. Sensors are self-calibrating, but need few warm-up seconds.

Participants were divided into two groups. For the first group (18 participants) the avatar was embedded into the test case Web site in order to support users, while the second group (17 participants) executed the assigned tasks without the avatar and its extra help: a between two groups experiment was chosen for not having learning effects.

Participants were asked to complete three tasks: registration onto the Website, reservation of a medical examination, adding an e-mail address in user's profile data. In the end, participants were asked to fill out Likert scales in a questionnaire used to assess user satisfaction and approval. The sitemap for these three tasks is depicted in Figure 5.

4.3 Statistical analyses

We tested the statistical significance of experimental results with the Mann-Whitney test, with linear correlation and calculating the Pearson's correlation coefficient [62]. We adopted the non-parametric Mann-Whitney (MW) test to prove the statistical significance of the results. The MW test calculates a *p* value that roughly represents the probability that the observed difference is due to chance. The *p* values standard thresholds are: $p<0.01$ (highly significant), $p<0.05$ (significant), $p>0.05$ (not significant).

The linear correlation fitting [62] is a way to prove that the correspondence between two variables can be approximated by a linear model (good fitting typically with $p<0.01$). The Pearson's correlation coefficient instead indicates the degree of correlation between two variables. So, if they are significantly related, the Pearson coefficient *p* is over the 0.5 value. Instead if they have a low correlation this coefficient is below 0.5. Negative values of *p* indicates inverse correlation.

5 Results

5.1 Non physiological analyses

The avatar presence significantly reduced the number of visited pages (table 1). Our hypothesis is that fewer visited pages are related to a better user experience, in line with [34]. Such reduction of visited pages is caused by the further information given by the avatar. Conversely, the task completion time was not significantly altered (table 1). This result is easy to interpret: participants listened to the avatar, and as a consequence their interaction with the website was not faster. Participants, answering the questionnaire, positively rated the avatar. On average, the effectiveness and the pleasantness were both rated above 3.9 on a 5-point Likert scale. These results confirm those obtained in previous works about Web avatars [21], [22], and confirm the validity of our experimental settings.

5.2 Analysis of physiological signals with and without avatar

Physiological signals at large are very subjective metrics and as such are very hard to be analyzed to derive robust findings. For this reason, the avatar effect on the users' physiological signals was evaluated by extracting many features out of the three signals.

In particular, for the Galvanic Skin Response we considered: mean, standard deviation, power, overshoot rate, (max-min)/|min|, power spectral density, same features applying either a high-selective pass band filter (> 1.5 Hz) or low-selective filter (<1.5 Hz); in particular, the overshoot rate represents the number of times the signal exceeds its standard deviation in one second.

For the respiration, we considered mean, standard deviation, power, overshoot rate, (max-min)/|min|, power spectral density, same features applying either a high-selective pass band filter (> 1.5 Hz) or low-selective filter (<1.5 Hz).

For the Heart Rate Variability, we considered mean, standard deviation, energy, power, overshoot (signal beyond the standard deviation) rate, power spectral density, same features applying either a high-selective pass band filter (> 1.5 Hz) or a low-selective filter (<1.5 Hz).

To improve the SNR a filtering operation with a low-selective pass filter (<12.5 Hz) was always performed for the three signals, as also suggested in the BioPac manual. The extracted features are common to other similar studies [35],[36], [5].

We investigated whether the avatar presence caused significant changes in the users' signals, comparing users experiencing the avatar with respect to users not experiencing the avatar in similar tasks. As reported in Table 2, the only physiological signal features out of all the tested ones that presented significant differences are the "GSR mean", "GSR overshoots number", the "respiration overshoots rate", the "heart rate mean". All these signal features presented significant differences in users experiencing the avatar with respect to users not experiencing the avatar (MW: $p<0.05$). In line with previous study on GSR [22] and the other considered signals [6], [8], such differences are due to a higher arousal caused by the avatar presence. However, the arousal level (Figure 2) tells nothing about the positive or negative valence of the emotion experienced by the participants.

.

Table 1: Hypotheses related to traditional Web usability metrics.

Hypothesis	Means (or data)	p-value	Significance
Avatar enriched website provides higher task completion rates	18 vs 17	p>0.05	Not Significant
Avatar enriched website provides fewer visited pages	23.8 vs 32.9	p=0.03	Significant
Avatar enriched website provides shorter task completion times	481s vs 487s	p>0.05	Not Significant

Table 2: Mann-Whitney significant tests for the features extracted from the measured signals.

Data	Hypothesis	Mean values	p-value	Significance
GSR mean	Higher in group with AF	0.04 vs 0.01	p<0.05	Significant
GSR number of overshoots	Higher in group with avatar	10.83 vs 8.71	p<0.05	Significant
Respiration overshoot rate	Higher in group with avatar	85.11 vs 47.06	p<0.05	Significant
Heart rate (HRV) mean	Higher in group with avatar	82.56 vs 71.76	p<0.05	Significant

Table 3: Mann-Whitney test for *overshoot rate*.

Data	Feature	Hypothesis	Mean	p-value	Significance
Respiration	Overshoot rate	Higher for approving participants	0.27 vs 0.05	p<0.01	Significant

Table 4: Linear regression for the *respiration overshoot rate* and the *approval score* in the final questionnaire.

Data	Feature	Hypothesis	p-value	Significance
Respiration	Overshoot rate - approval score	Linear Regression model fitting	p<0.01	Significant

Table 5: The Pearson's correlation coefficient for the tested hypothesis.

Data	Feature	Hypothesis	p-value	Significance
Respiration	Overshoot rate	Overshoots number correlated with the approval score	p = 0.88	Significant

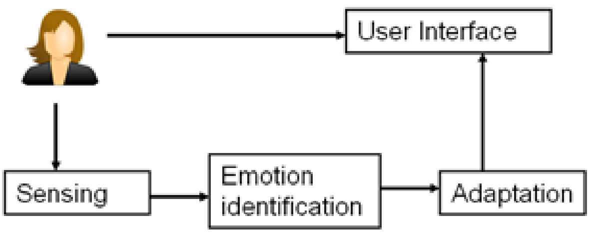

Fig. 6: General scheme to develop emotion-adaptive user interfaces.

5.3 Assessment of the avatar approval via physiological signals

The further step in our analyses is evaluating whether any of the signal features can be considered a metric for identifying the users' approval of the avatar. In our study, following users' questionnaire answers concerning the avatar approval, the group of users that experienced the avatar was divided in two groups: one group appreciating the avatar (11 participants), one group not appreciating the avatar (7 participants). The two groups were identified with the common data mining *k*-means algorithm with inputs the answers from the questionnaire. Once the two groups of users were identified, we investigated whether significant differences could be observed in any of the extracted features while comparing the approving users with respect to the disapproving ones. According to the statistical analyses, significant differences can be observed only in the respiration changes signal, for the feature overshoot rates (MW: p<0.01), wherein an overshoot is signal exceeding the standard deviation. This two feature proved effective in differentiating between users who appreciate the avatar and those who do not (Table 3).

The further step is trying to quantify the users' approval with respect to the avatar presence. For this reason, the questionnaire answers considered are used to calculate an approval score for each participant.
Following the Mann-Whitney results for the overshoot rates in respiration, we observed a linear relationship between respiration overshoot rates and approval scores by applying with success a linear regression (p<0.01, Table 4).

We confirm these results by observing a Pearson's correlation coefficient equal to 0.88 (significant), Table 5.

Such findings indicate that there is a significant correlation between the respiration overshoot rate and the appreciation of the online interaction with the avatar, and this can be exploited to infer the avatar approval, as explained in the following section.

5.4 Applications and limits

From our results, we obtain that the respiration overshoots rate proved to be a statistical significant feature to infer a positive valence, the approval level, when users interact with a Web avatar. In particular, as there is a linear regression model that fits with user's approval, once obtained the respiration overshoot rate, the regression line can be applied to such rate to derive user's approval level.

An indication of positive or negative valence in users' emotions can be exploited to improve the design of the system (in a controlled environment), or to properly control the avatar, providing greater or smaller support in line with user's approval (in an uncontrolled, every-day environment) [38][39]. In such a scheme (Figure 6), the user interacts with the system; in the meanwhile physiological signals (respiration in our case) are sensed and used to identify emotions via a classifier. The outputs of such classifier (a regression line in our findings) represent a feedback for the computer system, that can adapt the interface to the user's affective state (for instance, by switching off the avatar if the user is not approving it).

In this study, following the methodology presented in [13], we have investigated correlations between a single feature (out of the many extracted) and the approval level. As we have available more than one feature, in future works, we plan to investigate correlations among multiple features and the approval level. This can be done by adopting multiple inputs classifiers [11], for their ability to capture the mutual effects induced by multiple inputs. Our aim is to achieve a more robust classifications, that is needed especially when working in an every-day scenario.

6 Conclusions

Physiological signals can be used to infer emotions, and they can permit objective evaluation of the user satisfaction while she is interacting with a computer. As a consequence, they can be used to improve the design of interfaces or to build adaptive pleasant interactions. In this paper, we have explored the feasibility of inferring the approval level of users navigating on a Web system, through three physiological signals: Galvanic Skin Response, Heart Rate Variability and Respiration.

Two prototypal web sites for the online reservation of clinical exams were developed, with and without an avatar based helping system. The avatar was chosen as it well known that it improves users' arousal and satisfaction. The three physiological signals were collected during the experiments, as well as a post experiment questionnaire was performed to assess users' satisfaction and approval. Results confirm that avatar based interaction is effective in increasing users' arousal, and show that the respiration overshoots rate has a statistical significant correlation with the degree of approval of the avatar stated by the participants in the post experiment questionnaire. These finding indicates that, in our test case, such physiological feature can be exploited to infer a usability measure, the approval level, and eventually, it can be utilized to develop an adaptive interactive system. To achieve a more robust identification of the measure, in future works we plan to adopt multiple inputs (channels) classifiers, for their capability to capture the mutual effects induced by the multiple collected signals.

References:

[1] ISO 9241-11, Guidance on Usability. Available online: http://www.iso.org/iso/catalogue_detail.htm?csnumber=16883.

[2] Nielsen J. Usability 101: introduction to usability, AlertBox, Aug. 23, 2003.

[3] Lewis J.R.. IBM computer usability questionnaires: psychometric evaluation and instructions for use; Int. Journal of Human-Computer Interaction, v.7(1):57-78, 1995.

[4] Picard R.W.. Affective Computing, MIT Press, USA, 2000.

[5] Prendinger H., Ishizuka M.. The empathetic companion: a character-based interface that addresses users' affective states. Applied Artificial Intelligence, 19:267-285, 2005.

[6] Kim J., André E.. Emotion recognition based on physiological changes in music listening, IEEE Tran. on Pattern Analysis and Machine Intelligence, v.30(12):2067-2083, 2008.

[7] Healey, J.A.. Wearable and automotive system for affect recognition from physiology. PhD Thesis, MIT, 2000.

[8] Lee C., Yoo SK., et al. Using Neural Network to Recognize Human Emotions from Heart Rate Variability and Skin Resistance, Engin. in Medicine and Biology, 27th Conf., 2005.

[9] Nasoz F., Alvarez K., Lisetti C., Finkelstein N.. Emotion recognition from physiological signals for presence technologies; Intern. journal of cognition, technology and work; 6(1):32, 2003.

[10] Picard R.W., Vyzas E., Healey J.. Toward machine emotional intelligence: analysis of affective physiological state; IEEE

Transactions on pattern analysis and machine intelligence, v.23(10):1175-1191, 2001.
[11] Monajati, Mahdis, et al. "Emotions States Recognition Based on Physiological Parameters by Employing of Fuzzy-Adaptive Resonance Theory. International Journal of Intelligence Science2 (2012): 166-175.
[12] Phukan A.. Measuring usability via biometrics. Online Communities, LNCS 5621, pp.:101-107, Springer 2009.
[13] Drachen, Anders, et al. Correlation between heart rate, electrodermal activity and player experience in first-person shooter games. Proc. of the 5th ACM SIGGRAPH Symp. on Video Games. 2010.
[14] Oliveira, I., Isabel P., and Nuno M. Guimarães. "Practical Neurophysiological Analysis of Readability as a Usability Dimension." Human Factors in Computing and Informatics. Springer Berlin Heidelberg, 2013. 194-211.
[15] Liapis, A. and Michalis, X. The physiological measurements as a critical indicator in users' experience evaluation. Proc. 17th Panhellenic Conference on Informatics. ACM, 2013.
[16] Biopac Systems. Physiology lessons for use with the Biopac Student Lab; Lesson 9, Galvanic skin response and the Polygraph.
[17] Clark V.P., Maisog J.M. and Haxby J.V.. fMRI study of face perception and memory using random stimulus sequences, Journal of Neurophysiology, v.79(6):3257-3265, 1998.
[18] Haxby J.V., Hoffman E.A., Gobbini M.I.. Human neural systems for face recognition and social communication, Biological Psychiatry, v. 51(1):59-67, Elsevier, 2002.
[19] Haxby J.V., Gobbini M.I., Furey M.L., Ishai A., Schouten, J.L., Pietrini P.. Distributed and overlapping representations of faces and objects in ventral temporal cortex. Science, Sept. 28; 293(5539):2405-7, 2001.
[20] Qiu L., Benbasat I.. An Investigation into the effects of text-to-speech voice and 3D avatars on the perception of presence and flow of live help in electronic commerce, Transactions on Computer Human Interaction (TOCHI), 12(4):329-355; ACM Press, USA, 2005.
[21] Foglia P., Giuntoli F, Prete C.A., Zanda M.. Assisting E-Government Users with Animated Talking Faces, ACM Interactions, 14(1):24-26, ACM Press, 2007.
[22] Foglia, P., Prete, C.A., Zanda, M.. Relating GSR Signals to traditional Usability Metrics: A Case Study with an anthropomorphic Web Assistant. IEEE Int. Instrumentation & Measurement Technology Conf.. Canada, pp.:1814-1819, 2008.
[23] Abrams, S.. The polygraph in a psychiatric setting, American Journal of Psychiatry, 130(1):94-98, 1973.
[24] Jung, C.G.. On the psychophysical relations of the associative experiment, Journal of abnormal psychology, 1, pp.:247-255, 1907.
[25] Lin T., Omata et al. Do physiological data relate to traditional usability indexes? Proceedings of OZCHI 2005, pp: 1-10, ACM Press, 2005.
[26] Biopac System. pre-gelled disposable electrodes. http://www.biopac.com/disposable-electrode-100-secondary-education. 2014.
[27] Biopac System. BSL PRO Lesson H32: Heart Rate Variability Analysis. http://www.biopac.com/curriculum/pdf/h32.pdf , 2012.
[28] Aysin, B., Colombo J., and Aysin, E. Comparison of HRV analysis methods during Orthostatic Challenge: HRV with respiration or without?. Engineering in Medicine and Biology Society, 2007. 29th Annual Intern. Conf.
[29] Lang P.J.. The emotion probe: studies of motivation and attention, American Psychologist, v.50(5):372-385, 1995.
[30] Lang P.J., Bradley M.M., Cuthbert B.N.. International affective picture system (IAPS): instruction manual and affective ratings. TR:A-6, The Center for Research in Psychophysiology, University of Florida, 2005.
[31] Nielsen J.. Talking-Head video is boring online, AlertBox, Dec. 5, 2005.
[32] Stewart O., Huerta J.M.. Transition relevance place: a proposal for adaptive user interface in natural language dialog management systems. ACM SIGCHI Conference Work in progress, 2006.
[33] Foglia, P., & Comparetti, M. (2014). A workload independent energy reduction strategy for D-NUCA caches. The Journal of Supercomputing, 68(1), 157-182.
[34] Zhu J., Hong J., Hughes. PageCluster: Mining Conceptual Link Hierarchies from Web Log Files for Adaptive Web Site Navigation. Trans. on Internet Technology, 4(2):185-208; 2004.
[35] Lanatà, A., Scilingo, E.P., et al. Comparative evaluation of susceptibility to motion artifact in different wearable systems for monitoring respiratory rate. IEEE Trans. on inform. technology in biomedicine, 14(2):378-386, 2010.
[36] Moore M.M., Dua U.. A galvanic skin response interface for people with severe motor

disabilities, ACM conference on computers and accessibility assets, pp:48-54, 2004.
[37] Shi Y., Choi E.H.C., Ruiz N., Chen F., Taib R.. Galvanic Skin Response (GSR) as an index of cognitive workload, ACM CHI Conference Work-in-progress, 2007.
[38] Pantic M. and Rothkrantz L.J.M., 2003. Toward an affect sensitive multimodal human-computer interaction; Proceedings of the IEEE, v.91(9):1370-1390.
[39] Allanson J. and Fairclough S.H., 2004. A research agenda for physiological computing; Interacting with Computers, 16(5):857-878.
[40] Dhillon, B. S. Engineering product usability: a review and analysis techniques. WSEAS Trans. on Circuits and Systems 2 (2005): 86-94.
[41] Rubin, J., Chisnell, D. Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests. John Wiley & Sons, 2008.
[42] Park B., 2009. Psychophysiology as a Tool for HCI Research: Promises and Pitfalls, Proc. of 13th Intern. Conf. on Human-Computer Interaction; v.5610; pp.:141-148; Springer.
[43] Partala T., Surakka V. and Vanhala T., 2006. Real-time estimation of emotional experiences from facial expressions; Interacting with computers; v.18(2):208-226; Elsevier.
[44] Evans R.E. and Kortum P., 2010. The impact of voice characteristics on user response in an interactive voice response system; Interacting with computers; v.22(6):606-614; Elsevier.
[45] Kleinsmith A., Ravindra de Silva P. and Bianchi-Berthouze N., 2006. Cross-cultural differences in recognizing affect from body gesture; Interacting with computers; v.18(6):1371-1389; Elsevier.
[46] Tuisku O., Surakka V., et al., 2012. Wireless face interface: using voluntary gaze direction and facial muscle activations for human-computer interaction; Interacting with computers; v.24(1):1-9; Elsevier.
[47] Zeng Z., Pantic M., Roisman G.I. and Huang T.S., 2009. A survey of affect recognition methods: audio, visual, and spontaneous expressions; IEEE Tran. on Pattern Analysis and Machine Intelligence, v.31(1):39-58.
[48] Calvo R.A. and D'Mello S., 2010. Affect detection: an interdisciplinary review of models, methods, and their applications; IEEE Trans. on Affective Computing, v.1(1):18-37.
[49] Dini, G., Foglia, P., Prete, C. A., & Zanda, M. (2014). Social and Q&A interfaces for app download. Information Processing & Management, 50(4), 584-598.
[50] Tuch A.N., Bargas-Avila J.A., et al., 2009. Visual complexity of websites: effects on users' experience, physiology, performance, and memory; Int. J. of Human-Computer Studies, v.67(9):703-715.
[51] Haapalainen E., Kim S.J., et al, 2010. Psycho-physiological measures for assessing cognitive load. 12th ACM Int. Conf. on Ubiquitous Computing, pp. 301-310.
[52] Rainville P., Bechara A., et al, 2006. Basic emotions are associated with distinct patterns of cardiorespiratory activity; Int. J. of Psychophysiology, v.61(1):5-18; Elsevier.
[53] Wu S. and Lin. T, 2011. Exploring the use of physiology in adaptive game design; IEEE Int. Conf. on Consumer electronics, pp:1280-1284.
[54] Becker S.A., 2005. E-government usability for older adults; Communications of the ACM, v.48(2):102-104.
[55] Powell, J.A., Darvell M. and Gray J.A., 2003. The doctor, the patient and the world-wide web: how the Internet is changing healthcare. J. of the Royal Soc. of Medicine; v.96(2):74-76.
[56] Ward R.D. and Marsden P.H., 2003. Physiological responses to different Web page designs; Int. J. of Human-Computer Studies, v.59(1-2):199-212; Elsevier.
[57] Jitka, K., Jedlicka, M. and Hub, M. . "Usability user testing of selected web-based GIS applications." WSEAS Transactions on Computers 9.1 (2010): 21-30.
[58] Pastore, S. 2012. Website development and web standards in the ubiquitous world: where are we going. WSEAS Trans. on Comp. 11.9.
[59] Dini, G., Foglia, P., Prete, C. A., & Zanda, M. (2012). A social-feedback enriched interface for software download. Journal of Organizational and End User Computing, 25(1), 24-42.
[60] T. Miyake, T. Iwami, et al. Single Video Image Based Eye-Gaze Estimation for Human Interface. WSEAS Trans. on Computers, 4(11), 2005, pp. 1576-1582.
[61] Arh, T., & Blažič, B. (2008). A Case Study of Usability Testing–the SUMI Evaluation Approach of the EducaNext Portal. WSEAS Trans. on Information Science and Applications, 5(2), 175-181.
[62] Glantz, S., Primer of Biostatistics, McGraw-Hill, 2001.
[63] Dini, G., Foglia, P., Prete, C. A., & Zanda, M. (2006). An analysis of user interface factors influencing the acceptance of code download. 28th IEEE Information Technology Interfaces Conference (pp. 233-238).