fNIRS as a Method to Capture the Emotional User Experience: A Feasibility Study

Kathrin Pollmann[1,2(✉)], Mathias Vukelić[1,2], Niels Birbaumer[3,4], Matthias Peissner[2],
Wilhelm Bauer[2], and Sunjung Kim[3]

[1] Institute of Human Factors and Technology Management IAT, University of Stuttgart,
Stuttgart, Germany
{kathrin.pollmann,mathias.vukelic}@iat.uni-stuttgart.de
[2] Fraunhofer Institute for Industrial Engineering IAO, Stuttgart, Germany
[3] Institute of Medical Psychology and Behavioral Neurobiology, University of Tübingen,
Tübingen, Germany
[4] Istituto di Ricovero e Cura a Carattere Scientifico, Ospedale San Camillo, Venice, Italy

Abstract. User experience (UX) has become a key factor in interface design. Still, so far, no satisfying solution exists for measuring the emotional user experience (UX) during human-technology interaction (HTI) and linking them to design elements of the interface. Non-invasive brain imaging techniques are promising tools to assess the underlying causes and generation of emotional experiences in the brain. Against this background, especially functional near-infrared spectroscopy (fNIRS), a rather new and portable method, appears to have strong potential for measuring UX in real-world HTI settings. However, so far fNIRS has scarcely been used in emotion research. The present research evaluates the feasibility of using fNIRS to detect emotional user responses during HTI by comparing it to the well-established method of fMRI which, due to its set-up, is difficult to use in HTI context. Our feasibility study shows that fNIRS can detect brain activity patterns which are similar to those obtained using fMRI and can be used to distinguished positive and negative emotional reaction in an HTI context and displays brain activities which cannot be examined when fMRI is used. Future research should investigate whether similar results can be found when fNIRS is used in less controlled and more realistic HTI scenarios.

Keywords: User experience · Emotion research · Non-invasive brain-imaging · fNIRS · Human-technology interaction

1 Introduction

1.1 Measuring User Experience

In the past years human-technology interaction (HTI) research has started to direct more attention towards emotional and affective user reactions. Nowadays, usability is no longer the only important factor for successful interface design. There is an increased interest in designing products which do not only support the users in efficiently reaching their goals, but also maximize their positive experience during the

M. Kurosu (Ed.): HCI 2016, Part III, LNCS 9733, pp. 37–47, 2016.
DOI: 10.1007/978-3-319-39513-5_4

interaction. Recent studies in the field of HTI showed that users are more inclined to use or buy a technical product that induces a positive user experience (UX) compared to those products that do not [1, 2]. Still, until now, no satisfying solution has been proposed for identifying moments of positive and negative emotional experience during human-technology interaction (HTI). It is hence difficult to determine whether a technical product really has the intended emotional effect on its users and causes a positive UX. Furthermore, it is hard to find reliable solutions for linking the emotional state of the user to design elements of the interface.

Taking into account the distinct subjective nature of emotions, HTI mainly makes use of subjective methods like surveys, questionnaires or self-reports to assess UX. However, the value of these methods for measuring UX during HTI is limited: Subjective methods can either be employed in retrospective after the interaction or require the interruption of the interaction process. Both approaches are prone to attribution errors and cognitive biases and thus likely to yield a distorted image of the real UX [3]. Subjective measures also fail to identify emotional changes occurring over time [4]. Moreover, given that humans show limited abilities for introspection, subjective reports do not yield information about those emotional reactions that are rather implicit results from unconscious cognitive processing [5].

Subjective data can be complemented by results from behavioral methods such as observing the user or their task performance, which can be employed to continuously monitor the user's emotional reactions during HTI. In addition, video-based face recognition tools can be used to deduce the user's emotional state from their mimics [6]. However, like subjective measures, behavioral methods might fail to detect implicit emotional reactions which are most likely too subtle to be reflected in human behavior or mimics [5].

In addition, psychophysiological measures such as electrocardiography (ECG), electromyography (EMG) electrooculography (EOG) or electrodermal activity (EDA) are emerging as quantitative metrics to identify cognitive and emotional state changes underlying UX [7]. Psychophysiological methods detect electrical activity in the peripheral nervous system right after the user experiences an emotional event. Based on the measured activity the emotional reaction can be characterized on different dimensions such as its valence and the level of arousal, and based on the revealed patterns the emotional state can be inferred. However, psychophysiological methods are currently the least feasible approach for UX measurements as they require the attachment of sensors to the user's body, which restricts their comfort and wellbeing, and might ultimately alter the experience. Although psychophysiological methods are more suitable than subjective or behavioral methods to assess implicit emotional user reactions, activity measured from the peripheral nervous system only yields indirect information about the UX, as the main processing of the emotional event and generation of an emotional state takes place in the brain.

Thus, the more promising approach for measuring UX appears to be to focus on the immediate source of the user's emotional reaction by investigating neural activity in the brain rather than in the peripheral nervous system.

1.2 fNIRS: A New Tool for Emotion Research

Non-invasive brain imaging techniques, primarily developed for clinical settings, are powerful tools for assessing the source of mental and emotional states. Being able to detect and visualize changes in neuronal activity during task performance, brain imaging is becoming a realistic tool for HTI research. The most common method to gain deeper insights into emotional processing is functional Magnetic Resonance Imaging (fMRI) [8]. fMRI measurements provide three-dimensional functional images of the brain showing hemodynamic changes in blood volume and oxygenation, the so called blood oxygen level-dependent (BOLD) hemodynamic response, which occurs in relation to a stimulus. Using strictly controlled experimental set-ups fMRI studies were able to show that different emotional states can be distinguished from each other [9, 10]. However, having been designed for clinical or strictly controlled set-ups, fMRI has limitations regarding the applicability to real-world interaction scenarios and external validity and is hence less suitable to be used in an HTI context. The stationarity, set-up and functionality of fMRI scanners require participants to lie down in unnatural positions restricting their movements and the extent to which the interaction can be performed, thus distorting the UX.

Against this background, functional near-infrared spectroscopy (fNIRS), a rather new and portable method, appears to have strong potential for measuring UX in real-world HTI settings. Unlike other imaging techniques fNIRS does not require a strictly controlled environment and does not have as many restrictions such as stationarity, long set-up time or intolerance to movement.

Similarly to fMRI, fNIRS is a non-invasive, optical brain imaging technique that detects hemodynamic responses based on blood oxygen changes in the brain. To do so, it employs near-infrared light (wavelengths usually in the range from 760 nm to 850 nm) to measure concentration changes of oxygenated-hemoglobin (oxy-Hb) and deoxygenated-hemoglobin (deoxy-Hb) after local neuronal activations. Being based on similar physiological mechanism as fMRI, fNIRS can be regarded as a reliable and valid measurement for detecting activations in cortical regions [11]. fNIRS has successfully been used to asses various mental states during HTI [12] and recent studies show that fNIRS can be employed to detect emotional brain responses [11, 13]. Still, these studies were performed with receptive stimuli and did not involve any interaction of the participant. It is hence still unclear whether fNIRS is a suitable method to detect emotional responses during HTI and link them to the user's experience throughout the interaction process.

1.3 Research Question

The current study was carried out as part of a larger research initiative investigating the neuronal underpinnings of emotional user reactions during realistic HTI environments. It is aimed at investigating the feasibility of using fNIRS as a UX measurement tool by comparing its performance to the results of an fMRI-study which was conducted by Kim and colleagues within the same research project [14]. Using fMRI, we found different cortical activation patterns and hemispheric differences for events that were positively and negatively rated by participants within an HTI context. In our fNIRS-study we expect to find concentration changes in oxygenated- and deoxygenated-hemoglobin that reflect similar brain activity as detected in the fMRI-study.

2 Methods and Materials

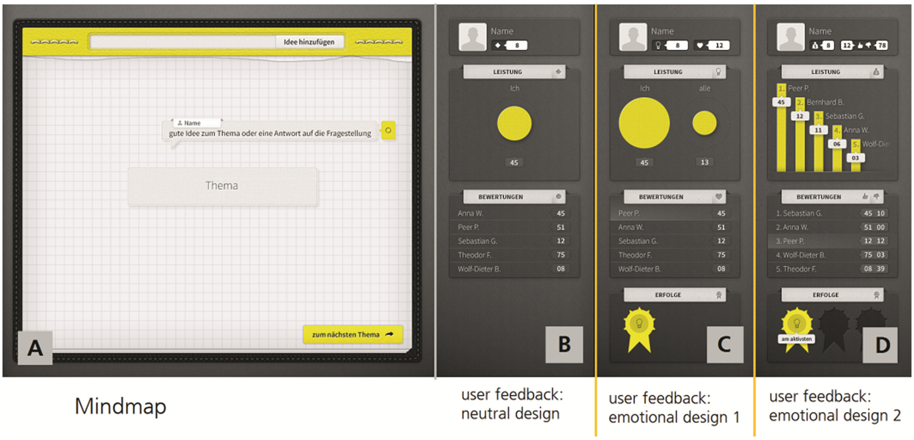
2.1 Subjects

Ten participants from the fMRI-study [14] were re-invited to take part in the present study (mean age: 24.90 ± 2.18 years, 4 females). All participants were right-handed. Participants gave their written informed consent before participation and received monetary compensation. The study protocol was approved by the local ethics committee of the Medical Faculty of the University of Tübingen.

2.2 Stimulus Material

Just like in the fMRI-study, participants were put into an HTI-scenario by confronting them with an interactive ideation tool, which was especially developed in a dedicated study [15].

There are two variations of the ideation tool (in the following referred to as “emotional 1” and “emotional 2”) of which each contains certain graphical design elements that have shown to evoke different emotional states in the user and should ultimately cause a positive UX. The third variation of the tool contains none of these elements, thus causing an emotionally neutral UX (see Fig. 1).



Anna W.	45
Peer P.	51
Sebastian G.	12
Theodor F.	75
Wolf-Dieter B.	08

Peer P.	45
Anna W.	51
Sebastian G.	12
Theodor F.	75
Wolf-Dieter B.	08

1. Sebastian G.	45	10
2. Anna W.	51	00
3. Peer P.	12	12
4. Wolf-Dieter B.	75	03
5. Theodor F.	08	39

Fig. 1. Screenshot of the ideation tool: **1A.** Exemplary shows the part of the screen designed for ideation in the form of a mind map which is placed next to an area that gives user feedback about their ideas. **1B, C and D**. show three different design variations of the user feedback area: the “neutral” version und two “emotional version” (from left to right).

As the physical set-up of the fMRI makes it difficult to realize a scenario were participants can interact with a technical system, in the fMRI-study for each of the three versions of the ideation tool five screenshots of the most prominent design elements were used as stimuli (15 screenshots in total). To guaranty the comparability of the fMRI and fNIRS results, the present study makes use of the exact same stimulus material and follows the same experimental procedure.

2.3 Experimental Procedure

In each session, participants were confronted with screenshots from all three versions of the ideation tool. To put these screenshots in the context of the whole tool, participants got an introduction to the tool and its functionalities at the beginning of each experimental session. Each session comprised two measurement blocks consisting of 15 trials each. Figure 2 shows the overview of the time course of one trial. Every trial consisted of a cued task design with different task epochs. The experiment procedure started with the presentation of the word "attention". This screen was only shown once to indicate the beginning of the experiment and direct the participant's attention to the screen. Each trial was initiated by a preparatory fixation phase followed by a short description of the area of the screen that the screenshot was taken from (e.g. user feedback category: performance). After another fixation phase, a screenshot of the ideation tool was visually presented. Each trial was completed by a rating period during which participants had to rate their emotional experience on a 10-point scale ranging from extremely negative (1–2: "I dislike it") to extremely positive (9–10: "I like it"). The selected screenshots were randomly sorted and each screenshot was presented twice to the participant.

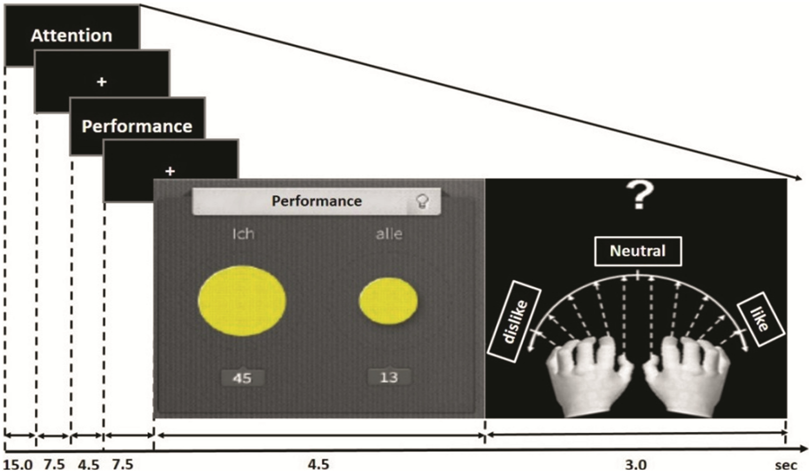


Fig. 2. Experimental paradigm showing the time course of each trial

2.4 Experimental Set-up

Participants were comfortably seated in a chair in front of a computer screen. Instead of a keyboard participants had a self-designed rating device, developed by the Institute of Medical Psychology and Behavioral Neurobiology and the MEG center, at the University of Tübingen, with ten keys lying on the table in front of them to perform the ratings for the different screen shots.

The changes of cerebral blood volume (CBV) in different brain regions during the interaction with the ideation tool were measured in the form of concentration changes of oxygenated- (oxy-Hb) and deoxygenated-hemoglobin (deoxy-Hb) by using a 30-optode fNIRS-system (ETG-400, Hitachi Medical Corporation, Japan). Two probe sets with 5*3 are used as a measurement mode, and the center position between them corresponds to Cz based on the international 10–20 system for electroencephalographic (EEG) electrode placement (Fig. 3). Each probe was composed of eight light sources and seven detectors, whereby a fixed distance of 30 mm between a source and a detector was ensured. Two wavelengths (695 and 830 nm) were used for measurement of the concentration of oxy-Hb and deoxy-Hb with a sampling period 0.1 s. The fNIRS system ETG-400 from Hitachi applies already the modified Beer-Lambert law [16] during data acquisition to obtain concentration changes of oxy- and deoxy-Hb.

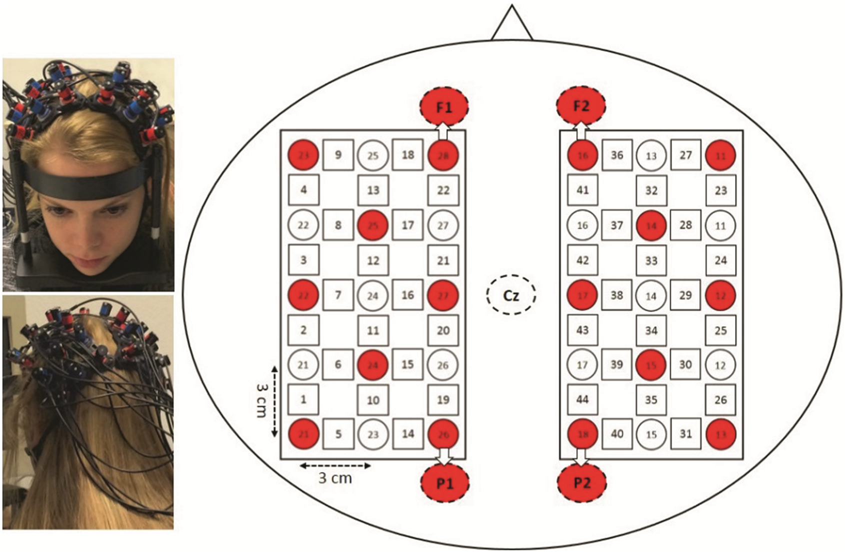


Fig. 3. Schema of the two fNIRS probe locations in each hemisphere: Each probe system included eight light source (red circles) and seven detectors (white circles) resulting in 22 measurement channels in each hemisphere (white squares). The number 27 (red circle) in the left hemisphere and the number 17 (red circle) in the right hemisphere are located so that the center between the both responds to Cz in the international 10–20 system for EEG electrode placement. Optode positions responding to the location of electrodes in the international 10–20 system are presented as red circles with dotted line (for example, the position of the number 28 in the left hemisphere responds to F1) (Color figure online).

2.5 Data Analysis

Subjective and Behavioral Data. First, we examined whether the emotional design of two versions of the ideation tool had an effect on participants' ratings. A repeated measures ANOVA with pairwise comparisons including emotionality of design as a with-in subject factor was used to compare participants' ratings for the screenshots of the three different version of the ideation tool.

In order to explore differences in participants' ratings and the response times between different emotional conditions, the screenshots with ratings 1 and 2 were categorized as dislike (D), the screenshots with ratings 3–8 as neutral (N) and the screenshots with ratings 9 and 10 as like (L). To examine differences between these three categories we performed a Kruskal-Wallis Test, as the data was not normally distributed. The response time was calculated by the time difference between the presentation of the rating task and participants' ratings. A participant's rating is considered correct, if the rating was performed within 3.0 s after the onset of the task. Otherwise, the rating was considered as missed and excluded from the analysis.

fNIRS Data. Offline fNIRS analysis was performed to examine the differences of activation levels (oxy-Hb and deoxy-Hb) between the different emotional graphical design elements of each trial and each channel. We focused one temporal window for the analysis of concentrations changes of oxy-Hb and deoxy-Hb: epoch of the visually presented screenshot of the ideation tool (3 s). Epochs in which the rating was missed have been excluded from further analysis. The epochs were pre-processed by high-pass filtering the raw data with a cut-off frequency of 0.02 Hz to eliminate baseline drifts and pulsation artifacts caused by heartbeat activity [17]. Next, the epochs were ordered according to the participants' rating belonging to each emotional category (D, L and N). Furthermore, we averaged the oxy- and deoxy-Hb changes across epochs on an individual basis for each emotional category. Finally, possible spatial differences of concentration changes during the respective time window were explored by analyzing statistical difference among the emotional categories (D > N, L > N, D > L) with the Mann-Whitney-U-Test, since data did not obey Gaussian distribution.

3 Results

3.1 Subjective and Behavioral Data

The statistical analysis showed an effect of emotionality of design on subjective ratings ($F(2, 1.82) = 24{,}06$, $p < .001$). Post-hoc pairwise comparisons revealed significant differences in participants' ratings between the "neutral" prototype ($M = 3.96$, $SD = 2.09$) and the "emotional 1" prototype ($M = 5.80$, $SD = 2.22$; $p < .001$) as well as between the "neutral" version and the "emotional 2" prototype ($M = 6.23$, $SD = 2.75$; $p < .001$). However, we did not observe any difference in evaluation rating between the two emotional versions of the ideation tool ($p = .675$).

Regarding intensity of the ratings, a Man-Whitney-U-Test revealed no significant difference between L ($Mdn = 2.0$) and D ratings ($Mdn = 2.0$; $U = 918.5$, $p = .065$).

The analysis of the response times indicated a tendency of participants to respond faster to L stimuli (Mdn = 0.78) than to N stimuli (*Mdn* = 1.03 and D stimuli (*Mdn* = 0.95), but no significant difference was found.

3.2 fMRI Results

The findings from the fMRI-study conducted by Kim et al. [14] indicate that the screenshots taken from the three versions of the ideation tool activated a distributed network of interacting cortical and subcortical brain regions which are known to be involved during emotional processing.

We found a strong left-hemispheric activation of the parietal cortex, i.e. the supramarginal gyrus for positively rated design elements, while negatively rated design elements mainly activated right-hemispheric midline (i.e. cingulum) and bi-hemispheric parietal cortical regions (i.e. supramarginal gyrus). Moreover, for negatively rated design elements the activated cortical brain regions showed a stronger involvement with subcortical brain regions (a network of putamen, thalamus and insula) than for positively rated design elements.

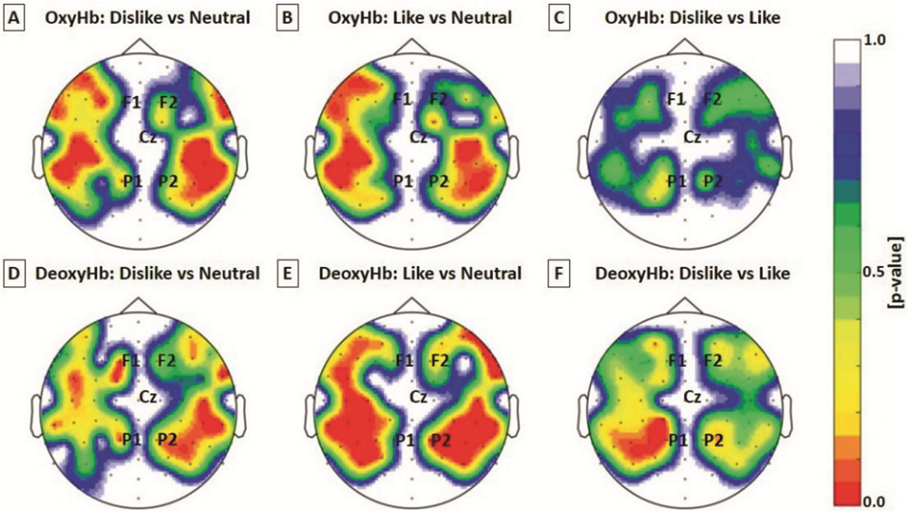


Fig. 4. Differences of the oxy- (**A-C**) and deoxy-Hb (**D-F**) concentration changes between the different emotional categories. **A**: Differences between Dislike (D) versus Neutral (N) showing higher concentration changes for D as compared to N. **B**: Differences between Like (L) versus Neutral (N) showing higher concentration changes for L as compared to N. **C**: Differences between Dislike (D) versus Like (L) showing no significant differences among these two categories. **D**: Differences between Dislike (D) versus Neutral (N) showing higher concentration changes for D as compared to N. **E**: Differences between Like (L) versus Neutral (N) showing higher concentration changes for L as compared to N. **F**: Differences between Dislike (D) versus Like (L) showing higher concentration changes for D as compared to L. Channels, showing significant differences in the Mann-Whitney-U-Test test, are indicated by colors (orange to red). Colors indicate different p-values (uncorrected) taken from the statistical test (Color figure online).

3.3 fNIRS Results

We observed significant changes of oxy-Hb for the emotionally loaded screenshots (dislike and like) as compared to the neutral one (Fig. 4A, B). Here, positively rated screenshots activated to a stronger degree left-hemispheric frontal, central and parietal regions and to some extend right-hemispheric central and parietal regions. On the other side, negatively rated screenshots activated bi-hemispheric central and parietal regions and to a minor extend left-hemispheric frontal regions. Furthermore, we found significant changes of deoxy-Hb for the emotionally loaded screenshots (dislike and like) as compared to the neutral one (Fig. 4D, E). Here, negatively rated screenshots activated bi-hemispheric central and parietal regions. For positively rated screenshots stronger activity was observed in bi-hemispheric central and parietal regions. Moreover, we observed significant changes of deoxy-Hb for the negatively loaded screenshots in comparison to the positively loaded ones (Fig. 4F). Here, the negative rated ones activated more left-hemispheric parietal regions.

4 Discussion

Participants' ratings reflected that the emotionally-loaded design elements induced positive and negative emotional states, while the neutral design elements did not. These findings are in line with the study by Sonnleitner and colleagues [15] who obtained similar results when initially evaluating the UX caused by the different versions of the ideation tool. From the fNIRS measurements we were able to distinguish on the cortical level positively rated design elements from negatively rated ones by comparing changes in oxy-Hb in frontal, central and parietal regions to neutrally rated ones. Here, similar to the results during the fMRI measurements positive design elements resulted in a stronger left-hemispheric activation, while negative design elements resulted in a bi-hemispheric activation pattern. Interestingly, the data revealed a strong involvement of deoxy-Hb concentration changes for the emotionally-loaded design elements i.e. a bi-hemispheric increase of central and parietal regions for positive rated design elements and right-hemispheric increase of central and parietal regions for negative rated design elements – a pattern that cannot be observed via fMRI measurements. The different activation patterns of oxy-Hb and deoxy-Hb in fronto-central and parietal regions between the two hemispheres show that fNIRS can be used to reliably differentiate between positive and negative emotional user reactions. Our results show that fNIRS is a potential, non-invasive and portable alternative method to fMRI measurements for sensing emotional states in realistic HTI settings. As mentioned in the introduction, the main benefit of fNIRS compared to fMRI is its portability, easy and fast set-up and low sensitivity to movement artefacts. However, in the present research these advantages of fNIRS only played a minor role. To assure comparability of the results between the fMRI and fNIRS measurements, our feasibility study was conducted as an exact replication of the fMRI-study, and thus only screenshots were used as stimulus material. These screenshots are merely snapshots from the actual interaction process and therefore have limited ecological validity.

To further evaluate the suitability of fNIRS as a tool for measuring UX, future studies will have to consider its use during ongoing interactions with a technical product, e.g. the ideation tool, rather than merely looking at receptive materials before drawing definite conclusions.

In addition, it would be interesting to compare the performance of fNIRS to other non-invasive brain imaging techniques, e.g. Electroencephalography (EEG) which is another portable and more popular neuroscientific method. To advance the approach of using non-invasive brain imaging techniques for UX measurement, it should be determined which portable method yields the best results and whether the informative value of the desired mental and emotional states could be enhanced by combining EEG and fNIRS.

Ultimately, the goal should be to develop a solution that is able to reliably detect and classify emotional user experience in real-time, thus providing the means to monitor the user throughout the whole interaction process. This would provide a suitable method to disclose moments of positive UX and to link them directly to the design elements of the technical system.

Acknowledgements. This research was supported by a grant from the Ministry of Science, Research and Arts of Baden-Wuerttemberg, Germany (Az: 32-729.63-0/5-14)

References

1. Kahneman, D.: Objective happiness. In: Kahneman, D., Diener, E., Schwarz, N. (eds.) Well-being. The Foundations of Hedonic Psychology, pp. 3–25. Russell Sage Foundation Press, New York (1999)
2. Spath, D., Peissner, M., Sproll, S.: Methods from neuroscience for measuring user experience in work environments. In: Proceedings of the International Conference on Applied Human Factors and Ergonomics (AHFE 2010) (2010)
3. Amelang, M., Schmidt-Atzert, L.: Psychologische Diagnostik und Intervention. Springer Medizin, Heidelberg (2006)
4. Hirshfield, L.M., Solovey, E.T., Girouard, A., Kebinger, J., Jacob, R.J.K., Sassaroli, A., Fantini, S.: Brain measurement for usability testing and adaptive interfaces: an example of uncovering syntactic workload with functional near infrared spectroscopy. In: Proceedings of the 27th International Conference on Human factors in Computing Systems, pp. 2185–2194. ACM, Boston, MA, USA (2009)
5. Minnery, B.S., Fine, M.S.: Neuroscience and the future of human-computer interaction. Interactions **16**(2), 70–75 (2009)
6. Gerhäuser, H., Elst, G.: Detection of Faces and Facial Expressions Used in Market Research–ShoreTM. Fraunhofer Institute for Integrated Circuits IIS, Erlangen (2011)
7. Fairclough, S.H.: Fundamentals of physiological computing. Interact. Comput. **21**(1), 133–145 (2009)
8. Braeutigam, S.: Neural systems supporting and affecting economically relevant behavior. Neurosci. Neuroecon. **1**, 11–23 (2012)
9. Phan, K.L., Wager, T., Taylor, S.F., Liberzon, I.: Functional neuroanatomy of emotion: a meta-analysis of emotion activation studies in PET and fMRI. Neuroimage **16**(2), 331–348 (2002). doi:10.1006/nimg.2002.1087

10. Sitaram, R., Lee, S., Ruiz, S., Rana, M., Veit, R., Birbaumer, N.: Real-time support vector classification and feedback of multiple emotional brain states. Neuroimage **56**(2), 753–765 (2011). doi:10.1016/j.neuroimage.2010.08.007
11. Ernst, L.H., Plichta, M.M., Lutz, E., Zesewitz, A.K., Tupak, S.V., Dresler, T., Ehlis, A.-C., Fallgatter, A.J.: Prefrontal activation patterns of automatic and regulated approach–avoidance reactions–a functional near-infrared spectroscopy (fNIRS) study. Cortex **49**(1), 131–142 (2013)
12. Girouard, A., Solovey, E.T., Hirshfield, L.M., Peck, E.M., Chauncey, K., Sassaroli, A., Fantini, S., Jacob, R.J.K.: From brain signals to adaptive interfaces: using fNIRS in HCI. In: Tan, D.S., Nijholt, A. (eds.) Brain-Computer Interfaces, Human-Computer Interaction Series, pp. 221–237. Springer, Heidelberg (2010)
13. Plichta, M.M., Gerdes, A.B.M., Alpers, G.W., Harnisch, W., Brill, S., Wieser, M.J., Fallgatter, A.J.: Auditory cortex activation is modulated by emotion: a functional near-infrared spectroscopy (fNIRS) study. Neuroimage **55**(3), 1200–1207 (2011)
14. Kim, S., Peissner, M., Bauer, W., Scheffler, K., Birbaumer, N.: Neural correlates of emotional preference in work environment. Front. Behav. Neurosci. (under review)
15. Sonnleitner, A., Pawlowski, M., Kässer, T., Peissner, M.: Experimentally manipulating positive user experience based on the fulfilment of user needs. In: Kotzé, P., Marsden, G., Lindgaard, G., Wesson, J., Winckler, M. (eds.) INTERACT 2013, Part IV. LNCS, vol. 8120, pp. 555–562. Springer, Heidelberg (2013)
16. Cope, M., Delpy, D.T.: System for long-term measurement of cerebral blood and tissue oxygenation on newborn infants by near infra-red transillumination. Med. Biol. Eng. Comput. **26**(3), 289–294 (1988)
17. Plichta, M.M., Herrmann, M.J., Baehne, C.G., Ehlis, A.-C., Richter, M.M., Pauli, P., Fallgatter, A.J.: Event-related functional near-infrared spectroscopy (fNIRS): are the measurements reliable? Neuroimage **31**(1), 116–124 (2006)