Measuring Task Engagement as an Input to Physiological Computing

Stephen H Fairclough
Liverpool John Moores University
Henry Cotton Campus, Liverpool
s.fairclough@ljmu.ac.uk

Katie C Ewing
Liverpool John Moores University
Henry Cotton Campus, Liverpool
k.c.ewing@ljmu.ac.uk

Jenna Roberts
Liverpool John Moores University
Henry Cotton Campus, Liverpool
j.roberts@ljmu.ac.uk

Abstract

Task engagement is a psychological dimension that describes effortful commitment to task goals. This is a multidimensional concept that combines cognition, motivation and emotion. This dimension may be important for the development of physiological computing systems that use real-time psychophysiology to monitor user state, particularly those systems seeking to optimise performance (e.g. adaptive automation, games, automatic tutoring). Two laboratory-based experiments were conducted to investigate measures of task engagement, based on EEG, pupilometry and blood pressure. The first study exposed participants to increased levels of memory load whereas the second used performance feedback to either engage (success feedback) or disengage (failure feedback) participants. EEG variables, such as frontal theta and asymmetry, were sensitive to disengagement due to cognitive load (experiment 1) whilst changes in systolic blood pressure were sensitive to feedback of task success. Implications for the development of physiological computing systems are discussed.

Introduction

Physiological computing (PC) describes systems that capture psychophysiological changes in the user in order to inform real-time software adaptation [1, 2]. PC systems rely on psychophysiology to create a representation of the psychological state of the user in real-time, e.g. changes in cognitive activity, positive and negative emotions, high vs. low task motivation. The system consults this representation to select an appropriate category of adaptive response. For example, if the user is frustrated, changes in user state should prompt the presentation of help information; if a player is bored by a computer game, the representation of user state should trigger an increase of game difficulty [3, 4]. The purpose of this approach is to create real-time software adaptation that is both implicit and intuitive.

The PC paradigm encompasses several existing strands of research/applications, from the control of adaptive automation [5, 6] to the use of psychophysiology to represent user emotion [7]. Unlike BCI applications [8], the PC approach is essentially passive (i.e. requiring no additional activity on the part of the user) and works mainly at the meta-level of the human-computer interaction (HCI) (i.e. ensuring that negative psychological states are minimised), i.e. whereas BCI represent an alternative form of input control [9].

The cycle of data collection and system response wherein psychophysiological change is transformed into adaptive control may be described as a biocybernetic loop [10]. This category of biocybernetic system control creates a symmetrical form of HCI where the availability of system information to the user is balanced by data about user state being at the disposal of the system [11]. Making a computer system privy to psychophysiological states has the potential to enable so-called ‘smart’ technology, i.e. systems that are characterised by increased autonomy and adaptive capability [12]. If technology develops in this direction, there is a subtle shift in the dynamics of HCI, from the master-slave dyad that characterises the way we currently use computers towards a collaborative, symbiotic relationship that requires computer technology to extend awareness of the user in real-time [13, 14].

One fundamental question surrounding the development of PC systems concerns how best to operationalise and represent the user state. There are several aspects to the question that should be considered during the initial stage of system design. In the first instance, what kind or dimension of user state is the most important one for a particular application domain? For example, physiological computing systems designed to control automation in the aircraft or vehicle cockpit have traditionally been concerned with representing the cognitive capability of the operator, specifically the prevention of Hazardous States of Awareness (HSA) [15]. Systems that employ psychophysiological measures for affective computing application emphasise the monitoring of negative affective states, such as anxiety [16] and frustration [17]. Similarly, psychophysiological monitoring has been used to identify quasi-emotional states, such as enjoyment, for those investigating this approach in the context of computer games [18]. At the second stage of system design, the researcher must identify those psychophysiological measures that provide the best operationalisation of the required psychological dimension. This stage may involve perusal of background literature followed by a series of validation experiments in the laboratory or the field, see [2] for full description of these issues.

This paper is concerned with how to measure the psychological dimension of task engagement as the basis for the development of PC systems. Task engagement is defined as “effortful striving towards task

978-1-4244-4799-2/09/$25.00 ©2009 IEEE

goals" [19]. This multidimensional concept incorporates at least three psychological dimensions: (1) the investment of mental effort to optimise cognitive performance, (2) motivation to successfully achieve task goals, and (3) affective changes associated with the likelihood of goal attainment. This dimension is important because engagement has a predictive relationship with human performance (i.e. greater engagement = superior performance) and wellbeing (i.e. disengagement from a task is associated with negative psychological states such as boredom or anxiety).

Previous research

Research into biocybernetic control of adaptive automation at NASA focused on the measurement of spontaneous electroencelographic (EEG) activity in order to capture task engagement, i.e. an EEG index ratio measure where the ratio of mean power in the high-frequency beta bandwidth (13-40Hz) is divided by total power in lower-frequency alpha (8-12Hz) and theta (3-7Hz) components ($\beta/(\alpha+\Theta)$) [10]. This prototypical system enabled automation of a laboratory-based task (the Multi-Attribute Task Battery - MATB) provided that the operator was deemed to exhibit high task engagement; if EEG measures of task engagement went into a decline whilst automation was activated, the system switched the user into a manual control mode, i.e. to re-engage with the task and prevent automation-induced complacency. This programme of research is summarised in [20].

The measurement of task engagement using psychophysiology takes on a different complexion in the context of desktop-based systems. For example, detection of negative user states is particularly relevant for computing applications designed to aid learning [21]. Recent work on the detection of user frustration [17] demonstrated the utility of the multimodal approach that combined multiple measures to predict subjective feelings of frustration. These authors measured skin conductance in combination with posture analysis, detection of head gestures (head shakes and nods), facial expression (smiling) and haptic monitoring. These measures were used to predict self-reported episodes of frustration, which was accurately detected in 79% of all cases (chance level = 58%). This experiment demonstrated how covert psychophysiology may be combined with overt behavioural signals in order to define the psychological dimension of interest. Related work on affective computing has also combined different psychological dimensions to yield a suitable representation of user state. For example, Burleson and Picard [22] described a state of "stuck" that may occur during the learning process to the detriment of user motivation. The definition of this state combines negative affect (e.g. anxiety) with cognitive characteristics (e.g. inability to focus, mental fatigue).

Measuring task engagement via psychophysiology

Task engagement can be defined with respect to cognitive activity (mental effort), motivational orientation (approach vs. avoidance) and affective changes (positive vs. negative valence).

Mental effort is conceptualised as energy mobilisation in the service of cognitive tasks or goals. At the cerebral level, the electrical activity of the brain may be quantified via the EEG to study how different states of brain activation represent the level of mental effort investment. The topography of EEG activation may provide important information about the specificity and distribution of activation over the cortex. A series of experiments demonstrated that augmentation of theta activity (4-7Hz) from central frontal sites and suppression of alpha activity from occipital areas were both associated with increased mental effort in response to working memory load (i.e. number of items to be retained in memory) [23, 24].

The pupillary response has a long association with the measurement of mental effort in response to cognitive variables [25, 26]. There is evidence that pupil dilation is greater during the processing of a complex cognitive operation relative to a simple one. The main problem with pupilometry is interference from light adaptation, i.e. for those environments where the level of lighting is not carefully controlled. The Index of Cognitive Activity [27]represents an attempt to quantify small discontinuities in pupil size that are related to cognitive activity. The ICA is derived in a selective manner that minimises the influence of lighting levels.

There is an obvious link between task engagement and the motivation to successfully achieve a given outcome. Motivational intensity theory [28, 29] proposes that goal commitment (i.e. the willingness to invest effort into the task) is a function of perceived: (i) task difficulty, (ii) ability, and (iii) likelihood that successful performance on the task will achieve a desired motive (e.g. monetary incentives, prowess, 'feeling good'). Therefore, if the individual assesses themselves to have the requisite level of skill to achieve success, then effort is invested into performance. Research into motivational intensity theory has used indicators of sympathetic nervous system (usually systolic blood pressure) to describe the "tipping point" where increased difficulty/reduced perception of ability/reduced perception that the task is worthwhile forces participants to switch from effortful striving for goal success to disengagement and a significant reduction of mental effort [30, 31].

Related research has linked changes in frontal EEG asymmetry to the self-regulation of affect and motivational orientation. In broad terms, the experience of positive emotions is associated with high levels of relative left frontal activity, whereas negative emotions is related to increased relative right frontal activity [32, 33]. There is also evidence that increased left frontal activation is correlated with motivational approach whilst right frontal activation is linked with a motivation disposition in the direction of avoidance. Research into the influence of reward on frontal asymmetry supports this connection [34-36], and higher levels of left frontal activation have been associated with trait measures of behavioural activation [37-39]. The relationship between motivational direction and affective valence encapsulated by frontal EEG asymmetry is implicit within a performance setting.

Task engagement is a multidimensional description of user state [40] that incorporates psychophysiological measures of cognition, motivation and affect. The relationship between physiology and psychology may be described as many-to-one [41] as multiple indicators from EEG, pupilometry and cardiovascular activity are deployed in concert to represent this dimension of task engagement. The purpose of this paper is to describe two laboratory experiments, both dedicated to the measurement of task engagement using different types of manipulation. In experiment one, participants are exposed to five levels of task demand using a working memory task. The aim of this experiment is to mentally overload the participants so he or she decides to withdraw effort from the task because it is deemed to be too difficult to achieve. The second experiment manipulated task engagement by providing participants with false feedback about the quality of their performance. One group was informed that performance was successively improving over time whereas the second group of participants received feedback of progressive performance decline. In the case of the second experiment, task engagement is influenced by manipulating participants' perception of their own ability.

Experiment 1: Mental Overload

Description of Study

21 participants (11 male) took part in the research, however data from 3 participants was excluded due to EEG artefacts and incorrect task completion. Participants were aged between 19 and 39 years of age. Cognitive effort was elicited with a verbal working memory task known as the n-back task. The task requires participants to indicate if the currently presented stimulus matches one shown on an earlier occasion. Solid black letters (against a white background) were presented to participants on colour monitor at a distance of 80cm. The task consisted of 6 levels of difficulty, with level 1 being the easiest and level 6 the most difficult. For each stimulus presentation participants needed to indicate if the letter matched the previous letter (level 1), the letter 2-previous (level 2), the letter 3-previous (level 3), the letter 4-previous (level 4), the letter 5-previous (level 5) and the letter 6-previous (level 6). Responses were given with a keyboard press of 1 for match and 2 for non-match, using the right index and middle fingers. Participants attended a training session of approximately 4.5 hours on the day before the experiment.

EEG activity was recorded monopolarly from 32 Ag-AgCl pin-type active electrodes mounted in a BioSemi stretch-lycra headcap. Electrodes were positioned according to the international 10-20 system and EEG activity recorded from the following sites: frontal pole (FP1, FP2), Anterior-frontal (AF3, AF4), frontal (F3, Fz, F4), fronto-central (FC5, FC1, FC2, FC6), central (C3, Cz, C4), temporal (T7, T8), parieto-central (CP5, CP1, CP2, CP6), parietal (P7, P3, Pz, P4, P8), occipito-parietal (PO3, PO4) and occipital (O1, Oz, O2). Electrodes were also placed at earlobe sites (A1, A2) allowing electrodes to be referenced off line to a linked ears reference. EEG was recorded continuously throughout a 4 minute baseline prior to the task and continuously throughout the task.

Systolic blood pressure measurements were taken using a Dinamap Vital Signs monitor (PRO100) using a cuff that was worn on the upper arm. Readings of systolic, and diastolic blood pressure along with heart rate and mean arterial pressure were obtained. A baseline reading was taken during a 4 minute period prior to task completion at 180s after the start of this period. Readings were then taken for each experimental trial 60s after onset giving 2 readings for each task level.

Pupil diameter measurements were recorded continuously at a sample rate of 60Hz with two remote infrared video cameras (Seeing Machines Ltd, Canberra, Australia). The cameras used binocular tracking and were mounted on a metal frame 80-90cm in front of the participant, placed beneath the stimulus display monitor. Pupil size resolution was possible at 0.00001mm. Data was recorded using FaceLAB 4.6 software. Illumination from the stimulus display and room lighting (8 x 36W ceiling mounted fluorescent tubes) was maintained within the range of 355-380Lux at the seated position of the participant to avoid a confound with the pupillary light reflex. Pupil diameter was measured throughout a 2min baseline prior to task completion during which participants were required to maintain their gaze at a fixation point (green dot) at the screen centre. Measurements were then made continuously throughout each trial. Participants were asked to keep still and maintain fixation at the centre of the screen minimizing possibility of head movement artifacts in the signal.

All EEG analysis was performed using BESA software (MEGIS software GmbH, Gräfelfing, Germany). First a 50Hz notch filter was applied to the raw data along with a 0.05Hz high pass and 60Hz low pass filter. A linked ears montage was applied. Data was visually inspected for artefacts from external electromagnetic sources which were excluded. Data underwent automatic correction for blink artefacts, horizontal and vertical saccades based on detection through predefined topographies. Average power spectra were then computed for each experimental condition. Power spectra in μV^2 were Log transformed (natural log) to normalise the distribution. Frontal asymmetry values were obtained for all 7 experimental conditions using EEG power values from the following electrode sites: FP2, AF4, F8, F4, FC2, FC6, C4, T8,(right hemisphere sites) FP1, AF3, F7, F3, FC1, FC5, C3, T7. (left hemisphere sites).

Power estimates for frequencies lying within Individual Alpha Bands were then used in the following formula: Ln [right total alpha power] – Ln [left total alpha power] to generate an asymmetry index [42]. Positive values indicated greater relative right alpha power and greater relative left frontal activity, greater relative right frontal activity was indicated by negative values. Asymmetries were also calculated for homologous pairs of electrodes.

Data from the left and right eye of 14 participants (7 female) was pre-processed to remove erroneous measures of pupil diameter arising from blinks, partial blinks, electromagnetic noise and artefacts resulting from tracking failure and camera joggle. Readings of 0 or near 0 were eliminated from the data to exclude blinks, partial blinks and tracking failure, and readings differing by more than +/- 0.1mm from the previous observation were excluded to reduce the influence of noise. The data then underwent 1-D wavelet decomposition using the orthogonal wavelet 'db4' from the Daubechies family of wavelets. The decomposition was achieved by convolving the signal with a high and low pass filter followed by downsampling by a factor of 2. Decomposition was performed using 5 iterations on each signal. The procedure produced a set of detail coefficients which were subjected to a minimax (hard) threshold to reduce noise, in which noise was presumed to be Gaussian white noise. Detail coefficients were then subjected to a threshold of 0.05 and coefficients above this value interpreted as showing high frequency discontinuous increases in pupil diameter. Numbers of these discontinuities, which have been found to correlate with cognitive processing [27], were used to generate an index consisting of the average no of discontinuities per second for each condition.

Results

EEG data were analysed with respect to two primary variables: frontal theta activity from the central area (Fz) and frontal asymmetry data. Theta activity at Fz was calculated using the dominant frequency (i.e. as personalised to each individual). The average power at the dominant theta frequency was calculated and submitted to analysis via ANOVA. The results revealed a significant trend [$F(6,12)=3.09$, $p<0.05$]. Post-hoc testing revealed that theta activity was significantly lower at baseline, the one-back and the six-back task compared to all other conditions ($p<0.05$).

Activity in the alpha bandwidth was also calculated with respect to the dominant frequency. Alpha power at the dominant frequency was calculated for all participants and converted via natural log prior to analysis. Asymmetry scores (left side minus right side) were calculated across three pairs of frontal sites on either side of the midline: AF3-AF4, F3-F4, FC3-FC4. Therefore, an increase of the asymmetry score is equated with greater activation of the left hemisphere. Each asymmetry score was analysed using an ANOVA model. There were no significant results for those asymmetry scores calculated with AF3-AF4 or F3-F4; however, the frontal-central sites (FC3-FC4) revealed a significant trend [$F(6,12)=2.57$, $p<0.05$]. Post-hoc testing revealed greater left-hemisphere activation (i.e. approach motivation) during all task conditions compared to baseline or the six-back condition ($p<0.05$). In other words, both the baseline (resting) condition and the six-back task were associated with greater levels of right hemispheric activation, which is associated with avoidance motivation. This finding is illustrated in Figure 1.

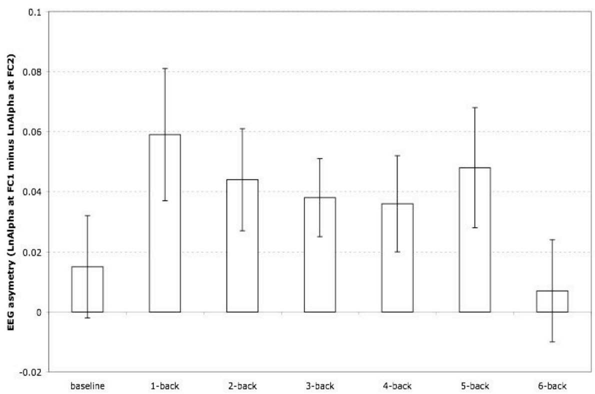

Figure 1. Frontal asymmetry scores (left side alpha power minus right side alpha power at dominant frequency) for FC3-FC4 across all six task demand conditions (N=18).

The measurement of systolic blood pressure has been associated with mental effort and task motivation. The analysis of this variable revealed a significant trend [$F(6,12)=13.01$, $p<0.01$]; however, post-hoc testing revealed only a significant difference between resting baseline and task conditions, i.e. the measure failed to distinguish between different levels of task demand.

An approximation of the Index of Cognitive Activity (ICA) was calculated for 14 participants based on changes in the pupil size. Specifically, the ICA captures short discontinuities in pupil size related to changes in mental workload. These data were subjected to ANOVA analysis, which revealed a significant difference due to experimental condition [$F(6,8)=7.26$, $p<0.05$]. Post-hoc testing revealed that the ICA was significantly lower than all working memory conditions, i.e. the ICA was not significantly sensitive to changes in working memory load (see Figure 2).

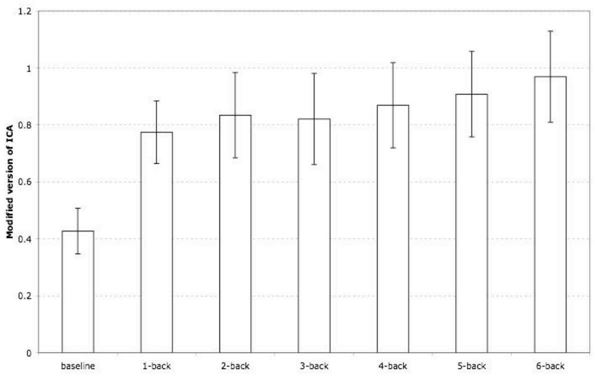

Figure 2. Mean score on modified Index of Cognitive Activity measure(N=14).

Experiment 2: Performance Feedback

Description of Study

34 participants (17 males and 17 females) formed 2 independent groups. A positive feedback group completed a working memory task and received pre arranged performance scores indicating a gradual improvement in performance over time. A negative feedback group completed the same task but received scores indicating a gradual decline in performance.

The memory task was computer based and was created using E-Studio software. It was developed from the 'n-back task' [24] . The version of the task used in this study was a 2 back task where participants continuously compared a currently presented stimulus to one seen 2 trials previously. Participants were presented with a 3x3 grid. On each trial, a green square appeared at one of the 9 grid locations for 1.75 seconds and was immediately followed by the next square. Participants were asked to respond on every trial by pressing 1 of 2 keyboard buttons to indicate that the location of the current square was either in the same location as the square seen 2 trials before (a match) or in a different location (a mis-match). The task was divided into 5 blocks, each of which contained 90 trials. Each block lasted just over 2.5 minutes and matches occurred on approximately 35% of trials.

In the experimental session of the study, participants were provided with false performance feedback as a percentage of overall accuracy at the end of each task block. Performance feedback was presented via a second computer placed adjacent to the memory task computer. Participants were misled to believe that performance data was being calculated in real-time by this second computer following each block of task activity. This illusion was achieved via a macro written in Microsoft Excel. The macro simulated a process of calculation and analysis and produced a chart to display performance accuracy. Each chart also included performance levels from any previous block/s which provided a visual representation of a gradual decline in performance for the negative feedback group and a gradual improvement in performance for the positive feedback group. Both groups received performance feedback of 60% after block 1 and both groups showed a cumulative decline or increase of 11% in total from block 1 to 5. For the negative feedback group, performance accuracy scores fell from 60% after block 1 to 56% after block 2, to 53% after block 3, to 52% after block 4 and finally reached 49% after block 5.

Blood pressure was recorded using a standard Dinamap with the pressure cuff placed over the brachial region of the participant's left arm. Initial screen and baseline readings were taken at the start of the experiment. Whilst participants worked on the memory task, 2 blood pressure readings were taken after approximately 20 seconds and 120 seconds from which an average was calculated.

EEG was recorded using active electrodes and sampled at 512Hz via a BioSemi system. Offline, EEG signals were corrected for ocular and physical artifacts and filtered using high and low band pass filters of 0.16Hz and 15Hz respectively. Artifact free epochs were then analysed via Fast Fourier Transform which yielded mean power in the alpha (8-12Hz) bandwidth. Alpha activity in the right hemisphere relative to homologous left hemisphere sites was calculated (ln [right] – ln [left]) to produce scores of alpha asymmetry for the following pairs of frontal sites: Fp2-1, Af4-3, F4-F3, FC2-FC1 and FC6-FC5. Theta activity was collected from frontal, central areas (Fz) as in the previous experiment.

Facial electromyographic activity (fEMG) was recorded to attain measures of muscle activity for the corrugator supercilii.

Results

EEG data: Two participants were excluded from this analysis due to technical problems with the data collection (one from each Feedback Group). The MANOVA analysis of EEG data revealed significant main effects for frontal asymmetry site, $F(4,26) = 5.70$, $p < .01$, and experimental condition, $F(1,29) = 4.05$, $p < .05$. The effect of experimental condition for EEG frontal asymmetry demonstrated that frontal asymmetry score (across all sites) was significant higher in the presence of performance feedback, i.e. higher level of activation in left hemispheric sites during feedback condition. There was no effect of feedback on levels of frontal theta activity.

Systolic Blood Pressure (SBP): The ANOVA model conducted on SBP data revealed a significant main effect for experimental condition, $F(1,30) = 4.82$, $p < .05$, i.e. SBP was significantly higher during Feedback [$\underline{M} = 115.78$] compared to the No Feedback condition

[M = 112.71]. The same model also revealed significant interactions between Feedback Group x Task Block, $F(4,27) = 3.55$, $p < .05$, and Feedback Group x Experimental Condition x Task Block, $F(4,27) = 3.20$, $p < .05$. For the positive feedback group, mean SBP was significantly higher at Task Block 5, $t(15) = 3.26$, during the Feedback condition compared to the No Feedback Condition (Figure 3).

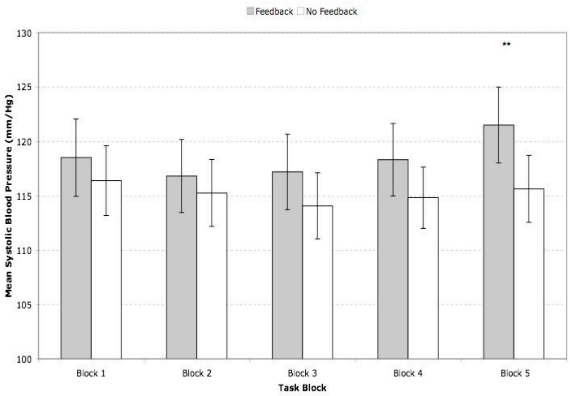

Figure 3. Mean Systolic Blood Pressure (mm/Hg) for Positive Feedback Group compared across both experimental conditions (N=16).

Corrugator Activity: The corrugator data were subjected to ANOVA model with an eyes open baseline included as an additional cell in the Task Block factor. This analysis revealed no significant effects. The trend of the data was to increase in presence of Feedback and this trend was particularly prominent during Task Block 5.

Discussion & Conclusions

Explanation of findings

Two experiments were conducted to identify the sensitivity of psychophysiological variables to the manipulation of task engagement. In the first experiment, engagement was manipulated by systematically increasing task difficulty. It was anticipated that the high level of working memory load at the 5- and 6-back versions of the task would cause participants to disengage. However, there was evidence from subjective measures of workload (NASA-TLX) that a point of overload was not reached, i.e. mean TLX score at 6-back task = approx. 6.5 on a 10-point scale. A reduction of frontal theta and an increase of right hemispheric frontal activity (Figure 1) was observed at maximum task demand. These data indicated that our participants were reducing levels of mental effort and shifting motivational orientation towards avoidance. In other words, they were withdrawing from the task. The pupilometry data from the ICA did not yield a statistically significant trend, however, a trend was observed of increasing cognitive demand (Figure 2).

These findings beg a question about volitional vs. mandatory responses to task demand in the psychophysiological realm. The positive linear relationship between task demand and ICA illustrated in Figure 2 contradicts the quadratic pattern that characterised both frontal theta and EEG frontal asymmetry (Figure 1). We may speculate that the ICA represents a response to perceived task demand, regardless of engagement, whereas the quadratic trend describes a self-regulated process of energy mobilisation. With respect to the latter, the initial level of low task demand (e.g. 1-back task) failed to increase frontal theta, which increased rapidly for 2-, 3- and 4-back versions of the task, before falling during the highest levels of task demand. The trend for frontal asymmetry was slightly different (Figure 1); exposure to the task led to increased approach motivation (at the 1-back task), which declined as task difficulty increased (indicating avoidance motivation) with the exception of a marked increase at the 5-back version of the task.

The second experiment attempted to manipulate task engagement in two ways. First, it was anticipated that performance feedback inevitably increases task engagement as the quality of one's own performance is rendered more salient. By providing repeated exposure to both positive and negative feedback, we anticipated different patterns of mental effort investment; specifically, we expected positive feedback to reduce effort investment (as participants received the impression that performance was consistently improving).

It was hypothesised that the presentation of negative feedback would initially mobilise high levels of effort, leading to disengagement towards the latter periods of the task as prompted by repeated exposure to negative feedback. The first hypothesis was supported by the frontal asymmetry data; participants exhibited higher left frontal activation during the feedback condition (regardless of whether feedback was positive or negative). The only psychophysiological response to the direction of feedback was found with respect to systolic blood pressure. This variable is associated with sympathetic activation of the autonomic nervous system, i.e. increased activation. Whilst systolic blood pressure did not respond to different levels of task demand during the first experiment, this variable exhibited a broadly linear increase in response to feedback of positive performance (Figure 3). This pattern was unexpected but was interpreted in the following way; contrary to expectations, positive feedback increased participants' appraisal of their own capability, which motivated these individual to both aspire towards higher levels of performance and increase mental effort mobilisation. The absence of the opposite trend in the presence of negative feedback was puzzling; perhaps negative feedback had no impact on any psychophysiological indicators of effort because the task was quite abstract and there were no negative consequences of task failure

Implications for Physiological Computing

What conclusions can be drawn from these laboratory studies for the development of physiological computing (PC) systems? In the first instance, the pattern of EEG data from experiment one point to the feasibility of capturing task engagement as a volitional response to task demand. This may be particularly important for applications such as computer games, which emphasise both autonomy and different levels of task demand. It is proposed that theta activity at frontal-central sites and frontal asymmetry are investigated as real-time variables to be integrated into the biocybernetic loop. Both variables demonstrated a sufficient degree of sensitivity to justify follow-up work. Further research must also explore individualised algorithms using neural net approaches [43, 44].

It should be noted that both EEG variables failed to show any sensitivity to positive vs. negative performance feedback during the second experiment. Therefore, these EEG variables seemed to respond primarily to engagement in the context of cognitive load. On the other hand, systolic blood pressure, which demonstrated a sensitivity to performance feedback, failed to distinguish between different levels of cognitive load in the first experiment. This pattern of results demonstrates the multidimensional nature of task engagement - different categories of measures may exhibit sensitivity to specific aspects of the concept. In this case, EEG variables respond to disengagement due to cognitive load whilst changes in systolic blood pressure reacted to changes in goal-setting behaviour, i.e. a desire to achieve at a higher level.

The relationship between physiology and psychology may be described as ‘many-to-one’ in the case of task engagement [41]; data from several physiological sources are required to successfully capture this dimension. The data from both experiments demonstrate the sensitivity of certain variables to different levels of task load or performance feedback. But the crucial distinction for the development of PC systems is the discrimination between rising engagement, sustained engagement and sustained disengagement. Systems that are designed to adaptively respond to changes in engagement need to assess: (1) how to facilitate rising levels of task engagement, and (2) how to counteract periods where the user may become disengaged from the task. With respect to our data, systolic blood pressure would appear to be a candidate for (1) whereas the EEG variables were sensitive to (2).

From the perspective of system design, it is not simply a question of selecting the correct variable to represent engagement, there is also the issue of sensitivity to the specific aspect of task engagement that is central to the application. For designers of adaptive automation applications, it is important to protect safety-critical performance; therefore, the ability to detect and predict task disengagement is a top priority. If the PC approach is applied to an automatic tutoring system, detection of sustained or rising engagement becomes just as important because learning software should be designed to engross and inspire users, and to sustain these positive states via real-time adaptation.

It is important for designers to have a clear idea about the level of discrimination that the system must achieve in order to provide appropriate levels of adaptation. For some systems, detecting two categories of engagement will suffice (high vs. low engagement); other systems may require more fine-tuned levels of discrimination (high vs. high/med vs. med vs. med/low vs. low). As the number of possible categories increases, the quantitative distance between each category declines, which will lead to higher false positives or misses, so the designer must consider this trade-off to optimise the performance of the system as a whole. Much depends on the adaptive capability of the system under development, PC systems that are capable of only one kind of adaptation (e.g. present help vs. no help presentation) only require a two-category classification. Systems with several levels of adaptive capability (e.g. present four different categories of help information) will require a psychophysiological algorithm that can discriminate four levels of task engagement [2].

From the perspective of building PC systems, it is obvious that psychophysiological variables offer significant advantages for representing user states. These measures are covert, passive and highly sensitive, but this level of sensitivity is double-edged. Psychophysiological variables are sensitive to a wide range of possible influences from physical artifacts (moving the body) to environmental factors (room temperature) to diurnal influences (time of day), the effects of caffeine and food, exercise, personality, mood etc. If system designers wish to harness the sensitivity of psychophysiology, this double-edged property must be appreciated. One could resolve the problem by monitoring confounding variables in order to model and isolate their influence on the psycho-physiological inference that is central to PC systems. Alternatively, designers could seek context via another route by considering psychophysiological changes in the same data space as other categories of variable, i.e. a multimodal approach [45]. This approach would combine psychophysiological changes with behavioural markers, such as posture [46] and facial expression. Psychophysiological changes could also be assessed in relation to measures of task performance [47]. One could combine markers from several categories (psychophysiological, behavioural, performance) in order to discern the level of task engagement via a process of triangulation. The danger with this approach is how to handle divergence/disagreement between the different categories of data.

To conclude, task engagement is an important psychological dimension for the development of physiological computing systems. It is also a complex dimension incorporating aspects of cognition with self-regulatory activities such as goal-setting and motivation.

Laboratory experiments have been described to identify candidate variables such as EEG frontal asymmetry and systolic blood pressure. The next step is to evaluate these variables in the context of a computerised task in the field.

References

1. Allanson, J. and S.H. Fairclough, A research agenda for physiological computing. Interacting With Computers, 2004. 16: p. 857-878.
2. Fairclough, S.H., Fundamentals of Physiological Computing. Interacting With Computers, 2009. 21: p. 133-145.
3. Gilleade, K.M., A. Dix, and J. Allanson. Affective videogames and modes of affective gaming: assist me, challenge me, emote me. in Proceedings of DiGRA 2005. 2005.
4. Fairclough, S.H. Psychophysiological inference and physiological computer games. in ACE Workshop - Brainplay'07: Brain-Computer Interfaces and Games. 2007.
5. Freeman, F.G., et al., Evaluation of a psychophysiologically controlled adaptive automation system, using performance on a tracking system. Applied Psychophysiology and Biofeedback, 2000. 25(2): p. 103-115.
6. Wilson, G.F. and C.A. Russell, Operator functional state classification using multiple psychophysiological features in an air traffic control task. Human Factors, 2003. 45(3): p. 381-389.
7. Picard, R.W., E. Vyzas, and J. Healey, Towards machine emotional intelligence: analysis of affective physiological state. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001. 23(10): p. 1175-1191.
8. Allison, B.Z., E.W. Wolpaw, and J.R. Wolpaw, Brain-computer interface systems: progress and prospects. Expert Review of Medical Devices, 2007. 4(4): p. 463-474.
9. Fairclough, S.H., BCI and Physiological Computing: Similarities, Differences and Intuitive Control, in Workshop on BCI and Computer Games: CHI'08. 2008: Florence.
10. Pope, A.T., E.H. Bogart, and D.S. Bartolome, Biocybernetic system evaluates indices of operator engagement in automated task. Biological Psychology, 1995. 40: p. 187-195.
11. Hettinger, L.J., et al., Neuroadaptive technologies: applying neuroergonomics to the design of advanced interfaces. Theoretical Issues in Ergonomic Science, 2003. 4(1-2): p. 220-237.
12. Norman, D.A., The Design of Future Things. 2007, New York: Basic Books.
13. Klein, G., et al., Ten challenges for making automation a "team player" in joint human-agent activity. IEEE Intelligent Systems, 2004. 19(6): p. 91-95.
14. Pantic, M., et al., Human computing and machine understanding of human behaviour: a survey, in Artificial Intelligence for Human Computing, T. Huang, et al., Editors. 2007, Springer. p. 47-71.
15. Prinzel, L.J., Research on Hazardous States of Awareness and Physiological Factors in Aerospace Operations. 2002, NASA: Hampton, Virginia.
16. Rani, P., et al., Anxiety detecting robotic system - towards implicit human-robot collaboration. Robotica, 2004. 22: p. 85-95.
17. Kapoor, A., W. Burleson, and R.W. Picard, Automatic prediction of frustration. International Journal of Human-Computer Studies, 2007. 65: p. 724-736.
18. Mandryk, R.L. and M.S. Atkins, A fuzzy physiological approach for continuously modeling emotion during interaction with play technologies. International Journal of Human-Computer Studies, 2007. 65: p. 329-347.
19. Matthews, G., et al., Fundamental dimensions of subjective state in performance settings: Task engagement, distress and worry. Emotion, 2002. 2(4): p. 315-340.
20. Scerbo, M.W., F.G. Freeman, and P.J. Mikulka, A brain-based system for adaptive automation. Theoretical Issues in Ergonomic Science, 2003. 4(1-2): p. 200-219.
21. Picard, R.W., et al., Affective learning - a manifesto. BT Technology Journal, 2004. 22(4): p. 253-269.
22. Burleson, W. and R.W. Picard. Affective agents: sustaining motivation to learn through failure and a state of "stuck". in Workshop on Social and Emotional Intelligence in Learning Environments. 2004.
23. Gevins, A. and M.E. Smith, Neurophysiological measures of cognitive workload during human-computer interaction. Theoretical Issues in Ergonomic Science, 2003. 4(1-2): p. 113-121.
24. Gevins, A., et al., Monitoring working memory load during computer-based tasks with EEG pattern recognition models. Human Factors, 1998. 40(1): p. 79-91.
25. Beatty, J., Task-evoked pupillary responses, processing load and the structure of processing resources. Psychological Bulletin, 1982. 91(2): p. 276-292.
26. Beatty, J., Phasic not tonic pupillary responses vary with auditory vigilance performance. Psychophysiology, 1982. 19(2): p. 167-172.
27. Marshall, S.P., C.L. Davis, and S.R. Knust, The index of cognitive activity: estimating cognitive effort from pupil dilation. 2004, EyeTracking Inc.: San Diego, CA.
28. Wright, R.A. and J.D. Dill, Blood pressure responses and incentive appraisals as a function of perceived ability and objective task demand. Psychophysiology, 1993. 30: p. 152-160.
29. Wright, R.A. and A. Dismukes, Cardiovascular effects of experimentally induced efficacy (ability) appraisals at low and high levels of avoidant task demand. Psychophysiology, 1995. 32: p. 172-176.
30. Richter, M. and G.H.E. Gendolla, Incentive value, unclear task difficulty, and cardiovascular reactivity in active coping. International Journal of Psychophysiology, 2007. 63(3): p. 294-301.
31. Richter, P. and G.H.E. Gendolla, Incentive effects on cardiovascular reactivity in active coping with unclear task difficulty. International Journal of Psychophysiology, 2006. 61: p. 216-225.
32. Davidson, R.J., Anterior electrophysiological asymmetries, emotion and depression: conceptual and methodological conundrums. Psychophysiology, 1998. 35: p. 607-614.

33. Davidson, R.J., What does the prefrontal cortex "do" in affect: perspectives on frontal EEG asymmetry research. Biological Psychology, 2004. 67: p. 219-233.
34. Pizzagelli, D., et al., Frontal brain asymmetry and reward responsiveness. Psychological Science, 2005. 16(10): p. 805-813.
35. Miller, A. and A.J. Tomarken, Task-dependent changes in frontal brain asymmetry: effects of incentive cues, outcome expectancies and motor responses. Psychophysiology, 2001. 38: p. 500-511.
36. Sobotka, S.S., R.J. Davidson, and J.A. Senulis, Anterior brain electrical asymmetries in response to reward and punishment. Electroencephalography and Clinical Neurophysiology, 1992. 83: p. 236-247.
37. Coan, J.A. and J.J.B. Allen, Frontal EEG asymmetry and the behavioural activation and inhibition systems. Psychophysiology, 2003. 40: p. 106-114.
38. Harmon-Jones, E. and J.J.B. Allen, Behavioural activation sensitivity and resting frontal EEG asymmetry: covariation of putative indicators related to risk for mood disorders. Journal of Abnormal Psychology, 1997. 106(1): p. 159-163.
39. Sutton, S.K. and R.J. Davidson, Prefrontal brain asymmetry: a biological substrate of the behavioural activation and inhibition system. Psychological Science, 1997. 8(3): p. 204-210.
40. Fairclough, S.H. and L. Venables, Prediction of subjective states from psychophysiology: a multivariate approach. Biological Psychology, 2006. 71: p. 100-110.
41. Cacioppo, J.T. and L.G. Tassinary, Inferring psychological significance from physiological signals. American Psychologist, 1990. 45(1): p. 16-28.
42. Coan, J.A. and J.J.B. Allen, Frontal EEG asymmetry as a moderator and mediator of emotion. Biological Psychology, 2004. 67: p. 7-49.
43. Wilson, G.F., J.D. Lambert, and C.A. Russell. Performance Enhancement with Real-Time Physiologically Controlled Adaptive Aiding. in Human Factors and Ergonomics Society Annual Meeting. 2007. Baltimore, Maryland: HFES.
44. Wilson, G.F. and C.A. Russell, Performance enhancement in an uninhabited air vehicle task using psychophysiologically determined adaptive aiding. Human Factors, 2007. 49(6): p. 1005-1018.
45. Pantic, M. and L.J.M. Rothkrantz, Towards an affect-sensitive multimodal human-computer interaction. Proceedings of the IEEE, 2003. 91(9): p. 1370-1390.
46. Ahn, H., et al. Stoop to conquer: posture and affect interact to influence computer users' persistence. in Second International Conference on Affective Computing and Intelligent Interaction. 2007. Lisbon, Portugal.
47. Fairclough, S.H., Psychophysiological inference and physiological computing games, in BrainPlay07: Brain-Computer Interfaces and Computer Games. 2007: Salzburg.