THEOR. ISSUES IN ERGON. SCI., 2003, VOL. 4, NOS. 1–2, 113–131 Taylor & Francis Taylor & Francis Group

Neurophysiological measures of cognitive workload during human–computer interaction

ALAN GEVINS* and MICHAEL E. SMITH

San Francisco Brain Research Institute and SAM Technology, 425 Bush Street, 5th Floor, San Francisco, CA 94108, USA

Keywords: Human–computer interaction; adaptive automation; mental workload; working memory; attention; neurophysiological signals; EEG.

Perhaps the most basic issue in the study of cognitive workload is the problem of how to actually measure it. The electroencephalogram (EEG) continues to be the clinical method of choice for monitoring brain function in assessing sleep disorders, level of anaesthesia and epilepsy. This preference reflects the EEG's high sensitivity to variations in alertness and attention, the unimposing conditions under which it can be recorded, and the low cost of the technology it requires. These characteristics also suggest that EEG-based monitoring methods might provide a useful tool in ergonomics. This paper reviews a long-term programme of research aimed at developing cognitive workload monitoring methods based on EEG measures. This research programme began with basic studies of the way neuroelectric signals change in response to highly controlled variations in task demands. The results yielded from such studies provided a basis on which to develop appropriate signal processing methodologies to automatically differentiate mental effort-related changes in brain activity from artifactual contaminants and for gauging relative magnitudes of mental effort in different task conditions. These methods were then evaluated in the context of more naturalistic computer-based work. The results obtained from these studies provide initial evidence for the scientific and technical feasibility of using EEG-based methods for monitoring cognitive load during human–computer interaction.

1. Introduction

Although the EEG has limitations with respect to its use as a method for three-dimensional anatomical localization of neurofunctional systems, it has clear advantages relative to other neuroimaging techniques as a method for continuous monitoring of brain function. Indeed, it is often the method of choice for some clinical monitoring tasks. For example, continuous EEG monitoring is an essential tool in the diagnostic evaluation of epilepsy (Thompson and Ebersole 1999) and in the evaluation and treatment of sleep disorders (Carskadon and Rechtschaffen 1989). It is also coming to play an increasingly important role in neuro-intensive care unit monitoring (Vespa *et al.* 1999) and in gauging level-of-awareness during anesthesia (John *et al.* 2001, O'Connor *et al.* 2001).

For many years, efforts have also been under way to evaluate the extent to which the EEG might be useful as a monitoring modality in applied work contexts. To be useful in such settings, a monitoring method should be robust enough to be reliably measured under relatively unstructured task conditions, sensitive enough to

* Author for correspondence. e-mail: alan@eeg.com

Theoretical Issues in Ergonomics Science ISSN 1463–922X print/ISSN 1464–536X online © 2003 Taylor & Francis Ltd
http://www.tandf.co.uk/journals
DOI 10.1080/14639220210159717

consistently vary with some dimension of interest, unobtrusive enough to not interfere with operator performance and inexpensive enough to eventually be deployable outside of specialized laboratory environments. It should also have reasonably good time resolution to allow tracking of changes in mental status as complex behaviours unfold. The EEG appears to meet such requirements. Furthermore, the compactness of EEG technology also means that, unlike other functional neuroimaging modalities (which require massive machinery, large teams of technicians and complete immobilization of the subject), EEGs can even be collected from an ambulatory subject who is literally wearing the entire recording apparatus (Gilliam *et al.* 1999).

In recent years, we have been evaluating the potential of the EEG as a measure of cognitive workload, primarily in individuals working at computers. Modern, computer-based work environments demand sustained vigilance to multiple streams of information. Such conditions have the potential to exceed a human's limited capacity to attend to and analyse information; cognitive overload has, thus, long been recognized (Card *et al.* 1983, Kieras 1988, Olson and Olson 1990) to be an important source of performance errors during human–computer interaction. The potential for overload is particularly acute in unskilled users, where unfamiliar procedures are likely to require greater commitment of cognitive resources (Anderson and Boyle 1987, Carlson *et al.* 1989). The ability to continuously monitor cognitive workload might, thus, be valuable in task analysis research and in efforts to improve the usability of human–computer interfaces (Raskin 2000). Indeed, a central problem in interface design is to develop means to provide information to a user with minimum disruption and distraction (Cadiz *et al.* 2001).

Because the problem of cognitive overload is widely recognized, it has been the topic of extensive empirical attention. Ironically, perhaps the most basic issue in the study of cognitive workload is the problem of how to actually measure it. One possibility is to use the EEG to directly measure the brain's response to a particular set of task demands. An EEG-based measurement of cognitive workload could help characterize the success of efforts to design suitable interfaces and interaction protocols. Such a tool might also aid in the design of appropriate adaptive-automation strategies (Morrison and Gluckman 1994, Byrne and Parasuraman 1996, Parasuraman *et al.* 2000).

We have been taking a systematic approach towards developing EEG-based methods for addressing this problem. Our first efforts revolved around identification and characterization of the properties of EEG signals sensitive to variations in the difficulty of highly controlled cognitive tasks. We also evaluated methods for analysis of such signals that might be suitable for use in a continuous monitoring context. More recently, we have begun to generalize those methods to assess computer-based tasks that are more naturalistic in character. In the following, we review the progress of those efforts.

2. Brain signals sensitive to variations in mental effort

Our first objective in this programme of research was to attempt to better characterize the neurophysiological changes that accompany increases in cognitive workload and the allocation of mental effort. We have approached this issue in the context of EEG and event-related potential (ERP) studies of working memory (WM). WM can be construed as an outcome of the ability to control attention and sustain its focus on a particular active mental representation (or set of representations) in the face of distracting influences (Engle *et al.* 1999). In many ways, this notion is nearly synonymous with what we commonly understand as the ability to effortfully 'concentrate' on task performance. This ability plays an important role in comprehension, reasoning, planning and learning (Baddeley 1992). Indeed, the effortful use of active mental representations to guide performance appears critical to behavioural flexibility (Goldman-Rakic 1987, 1988) and measures of it tend to be positively correlated with performance on psychometric tests of cognitive ability and other indices of scholastic aptitude (Carpenter *et al.* 1990, Kyllonen and Christal 1990, Gevins and Smith 2000).

Most of our investigations related to the neurophysiological concomitants of WM have required subjects to perform controlled 'n-back' style tasks (Gevins *et al.* 1990, 1996, Gevins and Cutillo 1993) that demand sustained attention to a train of stimuli. In these tasks, the load imposed on WM varies, while perceptual and motor demands are kept relatively constant. For example, in a spatial variant of the n-back task we have often employed, stimuli are presented at different spatial positions on a computer monitor once every 4 or 5 s while the subject maintains a central fixation. Subjects must compare the spatial location of each stimulus with that of a previous stimulus, indicating whether a match criterion is met by making a key press response on a computer mouse or other device. In an easy, low load version of the task, subjects compare each stimulus to the first stimulus presented in each block of trials (0-back task). In a more difficult, higher load versions, subjects compare the position of the current stimulus with that presented one, two or even three trials previously (1-, 2-, or 3-back tasks). These require constant updating of the information stored in working memory on each trial, as well constant attention to new stimuli and maintenance of previously presented information. To be successful in such tasks when WM demands are high, subjects typically must make a significant and continuous mental effort. Similar n-back tasks have recently been adopted in many other laboratories as a means to activate WM networks in a controlled fashion in the context of conventional behavioural studies (McElree 2001), other electrophysiological studies (Ross and Segalowitz 2000, Wintink *et al.* 2001), studies of the effects of magnetic fields on cognitive function (Koivisto *et al.* 2000, Oliveri *et al.* 2001) and functional neuroimaging studies employing PET or fMRI methods (Jonides *et al.* 1993, Cohen *et al.* 1994, McCarthy *et al.* 1994, Braver *et al.* 1997, Jansma *et al.* 2000).

The stimulus-locked ERPs recorded in such conditions in themselves provide an intriguing picture of the transient, rapidly shifting, sub-second patterns of activation that characterize the neurofunctional networks that underlie task performance (Gevins and Cutillo 1993, Gevins *et al.* 1995, 1996, McEvoy *et al.* 1998, 2001). There has also been a long and productive history of experimentation with such measures as indices of task imposed cognitive workload (Isreal *et al.* 1980, Sirevaag *et al.* 1993, Humphrey and Kramer 1994, Wilson *et al.* 1994, Kramer *et al.* 1995, Kok 2001, Ullsperger *et al.* 2001). However, to compute such measures requires either that a primary task itself emit distinct and more-or-less regular stimuli that an ERP response can be reliably time-locked to (something lacking from most real-world activities), or that a task-irrelevant probe stimulus be added to the operator's work environment. In either case, the low signal-to-noise inherent in most single trial ERPs can often necessitate averaging the response over many similar events.

The spectral composition of the ongoing EEG also displays regular patterns of load-related modulation during n-back task performance. Some components of the EEG spectrum could have significant utility for continuous monitoring applications

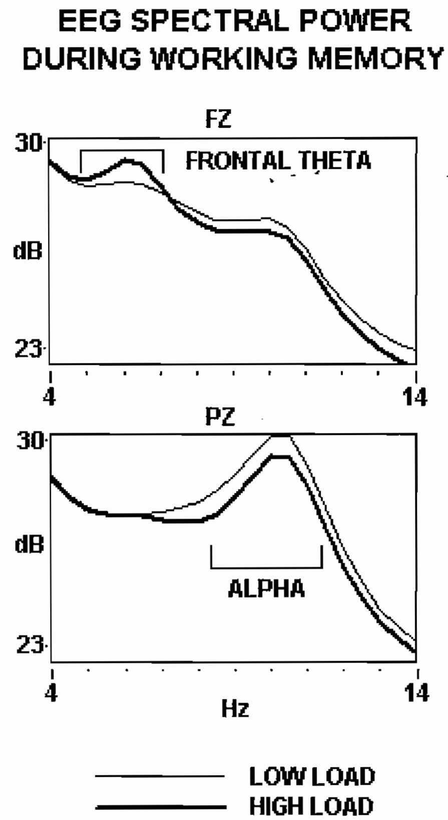

Figure 1. Effect of varying the difficulty of an n-back working memory task on the spectral power of EEG signals. The figure illustrates spectral power in dB of the EEG in 4–14 Hz range at frontal (Fz) and parietal (Pz) midline electrodes, averaged over all trials of the tasks and collapsed over 80 subjects. Data are from Gevins and Smith (2000).

and, in contrast to ERP indices, can be measured either independently of specific task events or in temporal conjunction to them. Because of these properties, we have focused our efforts to develop workload-monitoring methods on such spectral EEG measures. For example, figure 1 displays spectral power in the 4–14 Hz range at a frontal midline (Fz) and a parietal midline (Pz) scalp location computed from the continuous EEG during performance of low load (0-back) and moderately high load (2-back) versions of a spatial n-back task. The data represent the average response from a group of 80 subjects in a large study of individual differences in cognitive ability (Gevins and Smith 2000), and show significant differences in spectral power as a function of task load that vary between electrode locations and frequency bands.

More specifically, at the midline frontal site a 5–7 Hz or θ-band spectral peak is increased in power during the high load task relative to the low load task. This type of frontal midline τ-signal has frequently been reported to be enhanced in difficult, attention demanding tasks, particularly those requiring a sustained focus of concentration (Mizuki *et al.* 1980, Miyata *et al.* 1990, Yamamoto and Matsuoka 1990, Gundel and Wilson 1992, Gevins *et al.* 1997, 1998, Gevins and Smith 1999). Topographic analyses have indicated that this task loading-related θ-signal tends to have a sharply defined potential field with a focus in the anterior midline region of the scalp (Inouye *et al.* 1994, Gevins *et al.* 1997); such a restricted topography is unlikely to result from distributed generators in dorsolateral cortical regions. Instead, attempts to model the generating source of the frontal θ-rhythm from both EEG (Gevins *et al.* 1997) and magnetoencephalographic (Ishii *et al.* 1999) data have implicated the anterior cingulate cortex as a likely region of origin. This cortical region is thought to be part of an anterior brain network that is critical to attention control mechanisms and that is activated by the performance of complex cognitive tasks (Posner and Peterson 1990, Posner and Rothbart 1992). In a review of over 100 positron emission tomography (PET) activation studies that examined anterior cingulate cortex activity, Paus *et al.* (1998) found that the major source of variance that affected activation in this region was associated with changes in task difficulty. The EEG results are, thus, consistent with these views, implying that performance of tasks that require significant mental effort places high demands on frontal brain circuits involved with attention control.

In contrast, figure 1 also indicates that signals in the 8–12 Hz or α-band tend to be attenuated in the high load task relative to the low load task. This inverse relationship between task difficulty and α-power has been observed in many studies in which task difficulty has been systematically manipulated (Galin *et al.* 1978, Gundel and Wilson 1992, Gevins *et al.* 1997, 1998, Gevins and Smith 1999). Indeed, this task correlate of the α-rhythm has been recognized for over 70 years (Berger 1929). Because of this load-related attenuation, the magnitude of α-activity during cognitive tasks has been hypothesized to be inversely proportional to the fraction of cortical neurons recruited into a transient functional network for purposes of task performance (Gevins and Schaffer 1980, Pfurtscheller and Klimesch 1992, Mulholland 1995). This hypothesis is consistent with current understanding of the neural mechanisms underlying generation of the α-rhythm (reviewed in Smith *et al.* 2001). Convergent evidence for this view is also provided by observations of a negative correlation between α-power and regional brain activation as measured with PET (Larson *et al.* 1998, Sadato *et al.* 1998), and the frequent finding from neuroimaging studies of greater and more extensive brain activation during task performance when task difficulty increases (Baker *et al.* 1996, Bunge *et al.* 2000, Carpenter *et al.* 2000, Garavan *et al.* 2000).

In addition to signals in the θ- and α-bands, other spectral components of the EEG have also been reported to be sensitive to changes in effortful attention. These include slow wave activity in the δ- (< 3 Hz) band (McCallum *et al.* 1988, Rockstroh *et al.* 1989), high frequency activity in the β- (15–30 Hz) and γ- (30–50 Hz) bands (Sheer 1989) and rarely studied phenomenon such as the κ-rhythm that occurs around 8 Hz in a small percentage of subjects (Kennedy *et al.* 1948, Chapman *et al.* 1962). Since such phenomena were observed relatively infrequently in our series of studies on working memory, we will not discuss them further, but they may, nonetheless, ultimately prove useful in efforts to monitor cognitive workload using EEG measures.

3. Automating detection of mental effort-related changes in the EEG

As the data reviewed above indicate, spectral components of the EEG do in fact vary in a predictable fashion in response to variations in the cognitive demands of tasks. While this is a necessary condition for the development of an EEG-based monitor of cognitive workload, it is not sufficient. A number of other issues must also be addressed if such laboratory observations are to be transitioned into practical tools. Foremost among them is the problem of EEG artifact. That is, in addition to brain activity, signals recorded at the scalp include contaminating potentials from eye movements and blinks, muscle activity, head movements and other physiological and instrumental sources of artifact. Such contaminants can easily mask cognition-related EEG signals (Barlow 1986, Gevins *et al.* 1979a, b, c, 1980). In laboratory studies, human experts can be used to actively identify artifacts in raw data and eliminate any contaminated EEG segments, to insure that data used in analyses represent actual brain activity. For large amounts of data, this is an expensive, labour-intensive process which itself is both subjective and variable. To be practical in more routine applied contexts such decisions must be made algorithmically.

A great deal of research has been directed towards the problem of automated artifact detection. In previous work in our laboratory, we have developed and objectively evaluated several generations of automatic artifact detection algorithms. These include multi-criteria spectral detectors (Gevins *et al.* 1975, 1977), sharp transient waveform detectors (Gevins *et al.* 1976), and detectors using neural networks (Gevins and Morgan 1986, 1988). We have found that our most recent generations of detection algorithms perform about as well as the consensus of expert human judges. In a database of ~40 000 eye-movement, head/body movement and muscle artifacts, the algorithms successfully detected 98.3% of the artifacts, with a false detection rate of 2.9%, whereas the average expert human judge found 96.5% of the artifacts, with a 1.7% false detection rate. Thus, while further work on the topic is needed, it is reasonable to expect that the problem of automated artifact detection will not be an insurmountable barrier to the development of an EEG-based cognitive workload monitor.

A closely related problem is the fact that, in subjects actively performing tasks with significant perceptuomotor demands in a normal fashion, the incidence of data segments contaminated by artifacts can be high. As a result, it can be difficult to obtain enough artifact-free data segments for analysis. To minimize data loss, effective digital signal processing methods must also be developed to filter contaminants out of the EEG when possible. Our main approach to this problem has been to implement adaptive filtering methods to decontaminate artifacts from EEG signals (Du *et al.* 1994). We have found such methods to be effective at recovering most of the artifact contaminated data recorded in our typical laboratory studies of subjects working on computer-based tasks. A variety of other methods have been employed by different investigators in response to this problem, including such techniques as autoregressive modelling (Van den Berg-Lensssen *et al.* 1989), source modeling approaches (Berg and Scherg 1994) and independent components analysis (Jung *et al.* 2000). As with the problem of artifact detection, continued progress in this area suggests that, at least under some conditions and for some types of artifacts, decontamination strategies will evolve that will enable the automation of EEG processing for continuous monitoring applications.

Presuming, then, that automated pre-processing of the EEG can yield sufficient data for subsequent analyses, questions still remain as to whether the type of load-related changes in EEG signals can be measured in a reliable fashion in individual subjects and whether such measurements can be accomplished with a temporal granularity suitable for tracking complex behaviours. That is, in the experiments described above, changes in the θ- and α-bands in response to variations in WM load were demonstrated by collapsing over many minutes of data recorded from a subject at each load level, and then comparing the mean differences between load levels across groups of subjects using conventional parametric statistical tests. Under normal waking conditions, such task-related EEG measures have high test–re-test reliability when compared across a group of subjects measured during two sessions with 1 week between them (McEvoy *et al.* 2000). However, for the development of automated EEG analysis techniques suitable for monitoring applications, load-related changes in the EEG would ideally also be replicable when computed over short segments of data, and would need to have high enough signal-to-noise ratios to be measurable within such segments.

Prior work has demonstrated that multivariate combinations of EEG variables can be used to accurately discriminate between specific cognitive states (Gevins *et al.* 1979a, b, c, Wilson and Fisher 1995). Neural network-based pattern classification algorithms trained on data from individual subjects could also be used to automatically discriminate data recorded during different load levels of versions of the type of n-back WM task described above. For example, we performed an experiment (Gevins *et al.* 1998) in which eight subjects performed both spatial and verbal versions of 3-, 2- and 1-back WM tasks on test sessions conducted on different days. For each single trial of data in each subject, spectral power estimates were computed in the θ- and α-bands for each electrode site. Pattern recognition was performed with the classic Joseph-Viglione neural network algorithm (Joseph 1961, Viglione 1970, Gevins 1980, Gevins and Morgan 1986, 1988). This algorithm iteratively generates and evaluates two-layered feed-forward neural networks from the set of signal features, automatically identifying small sub-sets of features that produce the best classification of examples from the sample of data set aside for training. The resulting classifier networks were then cross-validated on the remaining data not included in the training sample.

Utilizing these procedures, we found that test data segments from 3-back vs 1-back load levels were discriminated with over 95% ($p < 0.001$) accuracy. Over 80% ($p < 0.05$) of test data segments associated with a 2-back load could also be discriminated from data segments in the 3-back or 1-back task loads. Such results provide initial evidence that, at least for these types of tasks, it is possible to develop algorithms capable of discriminating different cognitive workload levels with a high degree of accuracy. Not surprisingly, they also indicated that relatively large differences in cognitive workload are easier to detect than smaller differences, and that there is an inherent trade-off between the accuracy of classifier performance and the temporal length of the data segments being classified.

High levels of accurate classification were also achieved when applying networks trained with data from one day to data from another day and when applying networks trained with data from one task (e.g. spatial WM) to data from another task (e.g. verbal WM). We also attempted to develop networks trained with data from a group of subjects to data from new subjects. Such generic networks were found on average to yield statistically significant classification results when discriminating the 1-back from the 3-back task load conditions, but their accuracy was much reduced from that achievable with subject specific networks. On the one hand, such results indicate that there is a fair amount of commonality across days, tasks and subjects in the particular set of EEG frequency-band measures that are sensitive to increases in cognitive workload. Such commonalities can be exploited in efforts to design efficient sensor montages and signal processing methods. Nonetheless, they also indicate that, to achieve optimal performance using EEG-based cognitive load monitoring methods, it will likely be necessary to calibrate algorithms to accommodate individual differences. Such conclusions are also consistent with the observation that patterns of task-related EEG changes vary in conjunction with individual differences in cognitive ability and cognitive style (Gevins and Smith 2000).

Finally, it is also worthwhile to briefly mention another potential direction for such methods. In examining changes in neurophysiological correlates of n-back task performance over time, we found substantial changes in both the magnitude and the topography of task-related modulation of EEG activity in the α- and θ-bands between the time when naïve subjects were first learning to perform versions of the n-back WM and after they had developed some skill at it (Gevins *et al.* 1997, Smith *et al.* 1999). To further analyse these data, we used neural network methods analogous to those described above in an attempt to discriminate the task-related EEG signals recorded in the 3-back task when subjects were first learning the task from signals recorded during skilled performance. Across subjects, we were able to obtain very high levels of classification accuracy (range: 96–100%, $p < 0.001$) for discriminating naïve from practiced states. Such results imply that the WM demands of task performance changed with practice and that such changes could be automatically detected with EEG methods. Since overload of attentional or WM capacity has been found to be a limiting factor in the early stages of procedural skill acquisition (Woltz 1988, Kyllonen and Shute 1989), minimizing the potential of such overload is an important design guideline for the development of intelligent tutoring systems (Anderson and Boyle 1987, Carlson *et al.* 1989). The data described above, thus, suggest that it might be possible to utilize information provided by such monitoring methods to adapt a computer-aided instruction protocol to the cognitive constraints and skill levels of individual students.

4. Extension of neurophysiology-based workload monitoring methods to 'naturalistic' HCI

The types of results described above provide evidence for the basic feasibility of using EEG-based methods for unobtrusively monitoring cognitive task load in individuals engaged in computer-based work. However, the n-back WM task made minimal demands on perceptual and motor systems and it only required that a subject's effort be focused on a single repetitive activity. The ability to reliably measure cognitive load in individual subjects under such constrained circumstances might in itself be useful. For example, it has been applied to the problem of assessing the effect of environmental stressors on cognitive functions (Gevins and Smith 1999). Even so, in more naturalistic work environments, task demands are usually less structured and mental resources often must be divided between competing activities, raising questions as to whether results obtained with the n-back task could generalize to contexts that are more realistic.

Recent studies have demonstrated that more complicated forms of human–computer interaction (such as videogame play) produce mental effort-related modulation of the EEG that is similar to that observed during n-back tasks (Pellouchoud *et al.* 1999, Smith *et al.* 1999). This implies that it might be possible to extend EEG-based multivariate methods for monitoring task load to such circumstances. To evaluate this possibility we performed a subsequent study (Smith *et al.* 2001) in which the EEG was recorded while subjects performed the Multi-Attribute Task Battery (MATB; Comstock and Arnegard 1992). The MATB is a personal computer-based multi-tasking environment that simulates some of the activities a pilot might be required to perform. It has been used in several prior studies of mental workload and adaptive automation (Parasuraman *et al.* 1993, 1996, Fournier *et al.* 1999). The data collected during performance of the MATB were used to test whether it is possible to derive combinations of EEG features that can be used for indexing task loading during a relatively complex form of human–computer interaction.

The MATB task included four concurrently performed sub-tasks in separate windows on a computer screen (for graphic depictions of the MATB visual display, see Fournier *et al.* (1999) and Molloy and Parasuraman (1996)). These included a 'systems monitoring task' that required the operator to monitor and respond to simulated warning lights and gauges, a 'resource management task' in which fuel levels in two tanks had to be maintained at a certain level, a 'communications task' that involved receiving audio messages and making frequency adjustments on virtual radios, and a compensatory tracking task that simulated manual control of aircraft position. Manipulating the difficulty of each sub-task served to vary load; such manipulations were made in a between blocks fashion. Subjects learned to perform low-, medium- and high-load (LL, ML and HL) versions of the tasks. For comparison purposes, they also performed a 'passive watching' (PW) condition in which they observed the tasks unfolding without actively performing them.

Subjects engaged in extensive training on the tasks on one day, and then returned to the laboratory on a subsequent day for testing. On the test day, subjects performed multiple 5 min blocks of each task difficulty level. Behavioural and subjective workload ratings provided evidence that, on average, workload did indeed increase in a monotonic fashion across the PW, LL, ML and HL task conditions. This increase in workload was associated with systematic changes in the EEG. In particular, as in the prior study of workload changes in the n-back task paradigm, frontal θ-band activity tended to increase with increasing task difficulty, whereas α-band activity tended to decrease (figure 2). Such results indicated that the workload manipulations were successful, and that spectral features in the θ- and α-range might be useful in attempting to automatically monitor changes in workload with EEG measures.

Separate blocks of data were, thus, used to derive and then independently validate subject-specific, EEG-based, multivariate cognitive workload functions. In contrast to the two-class pattern detection functions that were employed to discriminate between different task load levels in the prior study, we evaluated a different technique that results in a single subject-specific function that produces a continuous index of cognitive workload and, hence, could be applied to data collected at each difficulty level of the task. In this procedure, the EEG data was first decomposed into short windows and a set of spectral power estimates of activity in the θ- and α-frequency ranges was extracted from each window. A unique multivariate function was then defined for each subject that maximized the statistical distance or divergence (Tou and Gonzalez 1974) between a small sample of data from low and high task load conditions. To cross-validate the function it was tested on new data segments from the same subject. Across subjects (figure 3), mean task load index values were found to increase systematically with increasing task difficulty, and differed significantly between the different versions of the task (Smith *et al.* 2001). These results provide encouraging initial evidence that EEG measures can indeed provide a modality for measuring cognitive workload during more complex forms of computer interaction. Although complex, the signal processing and pattern classification algorithms employed in this study were for real time implementation. A prototype

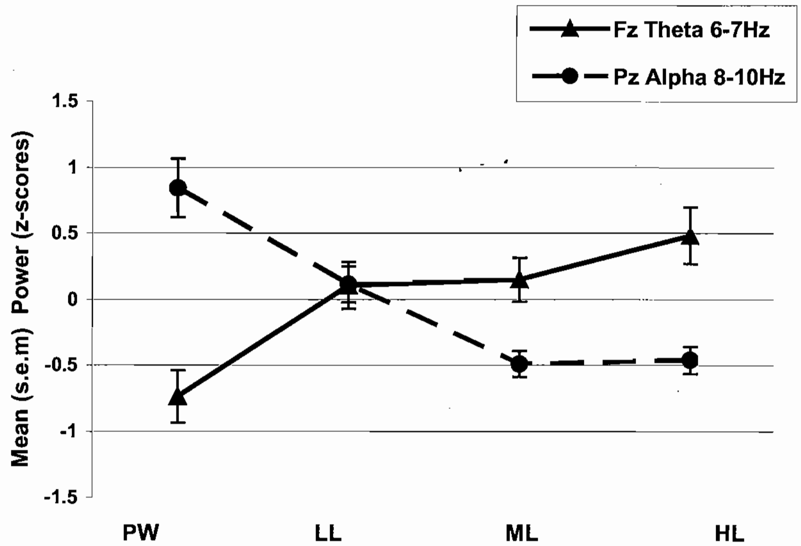

Figure 2. Mean ($n = 16$) EEG power for the frontal θ- and parietal α-EEG signals during performance of the MATB flight simulation task. Data have been normalized within each subject. Data are presented for each of four task versions (PW = passive watch, L = low load, ML = moderate load, HL = high load). Normalized spectral power for frontal θ and parietal α are plotted. τ-power increases from the PW to the HL task version, whereas θ-power decreases from the PW to the HL task version. Data are from Smith *et al.* (2001).

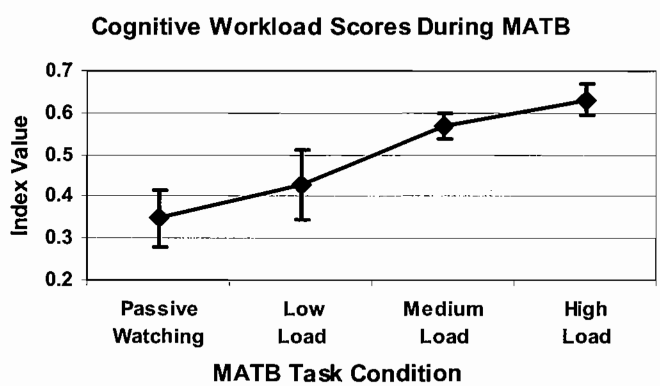

Figure 3. Mean and SEM ($n = 16$) EEG-based cognitive workload index values during performance of the MATB flight simulation task. Data are presented for each of four task versions (PW = passive watch, LL = low load, ML = moderate load, HL = high load). Average cognitive workload index scores increased monotonically with increasing task difficulty. Data are from Smith *et al.* (2001b).

online system running on a circa 1997 personal computer performed the requisite calculations on-line and provided an updated estimate of cognitive workload at 4 s intervals while subjects were engaged in task performance.

To further evaluate the utility of the approach described above as a tool for research on human–computer interaction, we also performed a small exploratory study that involved more naturalistic computer tasks. In this experiment (Smith and

Gevins, unpublished observations), EEG data was first recorded from subjects as they performed multiple series of trials of easy (0-back) and difficult (3-back) n-back tasks, as well as when they rested quietly, in order to establish a baseline view of the response of their brain to increased cognitive load. Data was then recorded while subjects performed more common computer-based tasks that were performed under time pressure and that were more-or-less intellectually demanding. These more naturalistic activities required subjects to perform word processing, take a computer-based aptitude test, and search for information on the web. The word processing task required subjects to correct as many misspellings and grammatical errors as they could in the time allotted, working on a lengthy text sample using a popular word processing program. The aptitude test was a practice version of the Computer-Adaptive GMAT® test. Subjects were asked to solve as many 'data sufficiency' problems as possible in the time allotted; such problems make a high demand on logical and quantitative reasoning skills and require significant mental effort to complete in a timely fashion. The web-searching task required subjects to use a popular web browser and search engine to find as many answers as possible in the time allotted to a list of trivia questions provided by the experimenter. For example, subjects were required to use the browser and search engine to 'convert 98.6 degrees Fahrenheit into degrees Kelvin', 'find the population of the 94105 area code in the 1990 US Census' and 'find the monthly mortgage payment on a $349 000, 30 year mortgage with a 7.5% interest rate'. Each type of task was structured such that subjects would be unlikely to be able to complete it in the time allotted.

The same basic analysis procedure described above that was applied to the EEG data recorded during MATB performance was also employed in this study. More specifically, a personalized continuous index of cognitive workload was first developed for each subject from a calibration set of data. In this case, the calibration data used to create the subject-specific index of cognitive workload included samples of the subject's 0-back and 3-back WM task EEG data. The resulting function was then applied to windowed samples of that subject's data from the quiet resting condition, from samples of 0-back and 3-back data not included in the calibration data set, and from samples of data during performance of the various naturalistic types of computer-based work.

A summary of the results from these analyses, averaged across data segments within each task condition and compared between conditions, is presented in figure 4. These comparisons indicated that the cognitive load index performed in a predictable fashion. That is, the condition in which the subject was asked to sit quietly and passively view a blank screen produced an average EEG-based cognitive workload around the zero point of the scale. Average index values during 0-back task performance were slightly higher than those during the resting condition, and average index values during the 3-back task were significantly higher than those recorded either during the 0-back WM task or during the resting state. All three naturalistic tasks produced workload index values slightly higher than that obtained in the 3-back task, which might be expected given that the n-back tasks had been practiced and were repetitive in nature, whereas the other tasks were novel and required the use of strategies of information gathering, reasoning and responding that were less stereotyped in form. Among the naturalistic tasks, the highest levels of cognitive workload were recorded during the computerized aptitude-testing task—the condition that was also subjectively experienced as the most difficult.

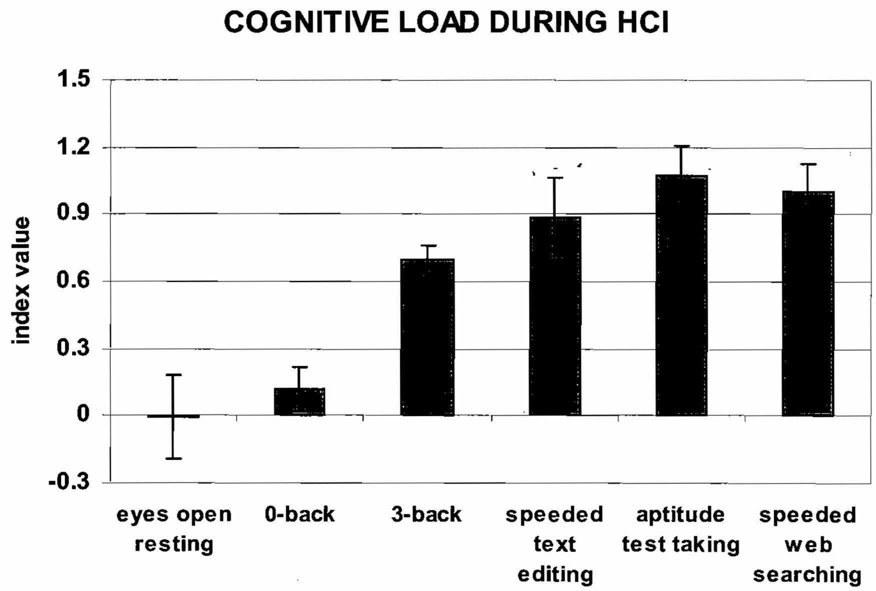

Figure 4. Mean and SEM ($n = 7$) EEG-based cognitive workload index values during resting conditions, easy and difficult versions of the n-back WM tasks, and a few naturalistic types of computer-based work (see text for full description of tasks and procedure). The data represent average index values over the course of each type of task. The easy WM and resting conditions produced significantly lower values than the more difficult WM condition or during the naturalistic tasks.

This pattern of results is interesting, not only because it conforms with *a priori* expectations about how workload would vary among the different tasks, but also because it provides data relevant to the issue of how the workload measure is affected by differences in perceptuomotor demands across conditions. Since in the n-back tasks, stimuli and motor demands are kept constant between the 0-back and 3-back load levels, the observed EEG differences in those conditions are clearly closely related to differences in the amounts of mental work demanded by the two task variants rather than other factors. However, in the study of MATB task performance described above, the source of variation in the index is somewhat less clear. On the one hand, performance and subjective measures unambiguously indicated that the mental effort required to perform the high load version of the MATB was substantially greater than that required by the low load (or passive watching) versions. On the other hand, the perceptuomotor requirements in the high load version were also substantially greater than those imposed by the other version. In this latter experiment, such confounds was less of a concern. Indeed, both the text editing task and the web searching task required more effortful visual search and more active physical responding than the aptitude test, whereas the aptitude test had little reading and less responding and instead required a great deal of thinking and mental evaluation of possibilities. Thus, the fact that the average cognitive workload values during performance of the aptitude test were higher than those observed in the other tasks provides convergent support for the notion that the subject-specific indices were more closely tracking variations in mental demands rather than variations in perceptuomotor demands in these instances.

5. Conclusions

In summary, the results reviewed above indicate that the EEG changes in a highly predictable way in response to sustained changes in task load and associated changes in the mental effort required for task performance. It appears that such changes can be automatically detected and measured using algorithms that combine parameters of the EEG power spectra into multivariate functions. Such methods can be effective both in gauging the variations in cognitive workload imposed by highly controlled laboratory tasks and in monitoring differences in the mental effort required to perform tasks that more closely resemble those that an individual might encounter in a real-world work environment.

The results presented would benefit from further replication, and are in need of significant refinement. For example, the data presented herein collapsed cognitive activity over 'whole-tasks', that is, the data were collapsed over many minutes of sustained performance. However, cognitive workload indices were calculated over data segments that were frequently updated and were of much shorter duration. Future work will need to try to identify how such momentary measures of cognitive workload vary with specific intra-task events. One possibility for using EEG spectral measures to evaluate cognitive load in response to specific task events is to employ 'event-related desynchonization' (ERD) methods that compare post-stimulus power to a pre-stimulus baseline measure and use degree of change in the spectra as a load measure. However, past efforts to evaluate such measurements in somewhat naturalistic tasks (in fact, during the MATB) have found that, although useful in single tasks, the ERD is insensitive to workload variations in multi-tasking contexts (Fournier *et al.* 1999). This failure likely reflects the fact that, in such contexts, workload is likely to be relatively high even in the pre-stimulus baseline period, and so any stimulus related change in the EEG spectra is likely to be fairly small. It is an open question whether other types of EEG-based methods might be more fruitfully applied in such circumstances.

Another area of future refinement is related to the current unitary nature of the cognitive workload measures. That is, some views of the structure of the mental resources that can be allocated to task performance posit a relative independence of the resources involved with cognitive processes and those involved with perceptual processing and motor expression. Future development of such methods should, thus, explore the possibility of developing somewhat orthogonal physiological indices that can differentiate between the loading of one or another type of neural resource system. While the need for such future refinements is clear, the current results, nonetheless, provide compelling initial evidence for the feasibility of creating EEG-based technologies for monitoring cognitive workload during human–computer interaction.

Acknowledgements

This research was supported by several agencies of the US Government including the Air Force Office of Scientific Research, the National Science Foundation and the National Aeronautics and Space Administration. We thank Drs Harrison Leong, Robert Du and Linda McEvoy for their contributions to the studies reviewed herein.

References

ANDERSON, J. R. and BOYLE, S. F. 1987, Cognitive principles in the design of computer tutors, in P. Morris (ed.), *Modelling Cognition* (London: John Wiley & Sons Ltd), 93–133.

BADDELEY, A. 1992, Working Memory, *Science*, **255**, 556–559.

BAKER, S. C., ROGERS, R. D., OWEN, A. M., FRITH, C. D., DOLAN, R. J., FRACKOWIAK, R. S. and ROBBINS, T. W. 1996, Neural systems engaged by planning: a PET study of the Tower of London task, *Neuropsychologia*, **34**, 515–526.

BARLOW, J. S. 1986, Artifact processing rejection and reduction in EEG data processing, in F. H. Lopes da Silva, W. Storm van Leeuwen and A. Remond (eds), *Handbook of Electroencephalography and Clinical Neurophysiology Vol 2* (Amsterdam: Elsevier), 15–65.

BERG, P. and SCHERG, M. 1994, A multiple source approach to the correction of eye artifacts, *Electroencephalography and Clinical Neurophysiology*, **90**, 229–241.

BERGER, H. 1929, Uber das Elektroenzephalogramm des Menschen, *Archives of Psychiatry*, **87**, 527–570.

BRAVER, T. S., COHEN, J. D., NYSTROM, L. E., JONIDES, J., SMITH, E. E. and NOLL, D. C. 1997, A parametric study of prefrontal cortex involvement in human working memory, *Neuroimage*, **5**, 49–62.

BUNGE, S. A., KLINGBERG, T., JACOBSEN, R. B. and GABRIELI, J. D. 2000, A resource model of the neural basis of executive working memory, *Proceedings of the National Academy of Sciences (USA)*, **97**, 3573–3578.

BYRNE, E. A. and PARASURAMAN, R. 1996, Psychophysiology and adaptive automation, *Biological Psychology*, **42**, 249–268.

CADIZ, J. J., VENOLIA, G. D., JANCKE, G. and GUPTA, A., 2001, *Sideshow: Providing peripheral awareness of important Information* (MSR-TR-2001-83, Redmond, WA: Microsoft Research, Microsoft Corporation).

CARD, S. K., MORAN, T. P. and NEWELL, A. 1983, *The Psychology of Human-Computer Interaction* (Hillsdale, NJ: Lawrence Erlbaum Associates, Inc).

CARLSON, R. A., SULLIVAN, M. A. and SCHNEIDER, W. 1989, Practice and working memory effects in building procedural skill, *Journal of Experimental Psychology: Learning, Memory and Cognition*, **3**, 517–526.

CARPENTER, P. A., JUST, M. A. and REICHLE, E. D. 2000, Working memory and executive function: evidence from neuroimaging, *Current Opinion in Neurobiology*, **10**, 195–199.

CARPENTER, P. A., JUST, M. A. and SHELL, P. 1990, What one intelligence test measures: a theoretical account of the processing in the Raven Progressive Matrices Test, *Psychological Review*, **97**, 404–431.

CARSKADON, M. A. and RECHTSCHAFFEN, A. 1989, Monitoring and staging human sleep, in M. H. Kryger, T. Roth and W. C. Dement (eds), *Principles and Practice of Sleep Medicine*, 2nd edn (Philadelphia: W.B. Saunders & Co), 943–960.

CHAPMAN, R. M., ARMINGTON, J. C. and BRAGDEN, H. R. 1962, A quantitative survey of kappa and alpha EEG activity, *Electroencephalography and Clinical Neurophysiology*, **14**, 858–868.

COHEN, J. D., FORMAN, S. D., BRAVER, T. S., CASEY, B. J., SERVAN-SCHREIBER, D. and NOLL, D. C. 1994, Activation of prefrontal cortex in a non-spatial working memory task with functional MRI, *Human Brain Mapping*, **1**, 293–304.

COMSTOCK, J. R. and ARNEGARD, R. J. 1992, *The Multi-Attribute Task Battery for Human Operator Workload and Strategic Behavior Research* (104174: NASA Technical Memorandum).

DU, W., LEONG, H. M. and GEVINS, A. S. 1994, Ocular artifact minimization by adaptive filtering, *Proceedings of the Seventh IEEE SP Workshop on Statistical Signal and Array Processing*, Quebec City, Canada, 433–436.

ENGLE, R. W., TUHOLSKI, S. and KANE, M. 1999, Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence and functions of the prefrontal cortex, in A. Miyake and P. Shah (eds), *Models of Working Memory* (Cambridge: Cambridge University Press), 102–134.

FOURNIER, L. R., WILSON, G. F. and SWAIN, C. R. 1999, Electrophysiological, behavioral, and subjective indexes of workload when performing multiple tasks: manipulations of task difficulty and training, *International Journal of Psychophysiology*, **31**, 129–145.

Galin, D., Johnstone, J. and Herron, J. 1978, Effects of task difficulty on EEG measures of cerebral engagement, *Neuropsychologia*, **16,** 461–472.

Garavan, H., Ross, T. J., Li, S. and Stein, E. A. 2000, A parametric manipulation of central executive functioning, *Cerebral Cortex*, **10,** 585–592.

Gevins, A. and Cutillo, B. 1993, Spatiotemporal dynamics of component processes in human working memory, *Electroencephalography and Clinical Neurophysiology*, **87,** 128–143.

Gevins, A. and Morgan, N. 1986, Classifier-directed signal processing in brain research, *IEEE Transactions on Biomedical Engineering*, **33,** 1058–1064.

Gevins, A. and Smith, M. E. 1999, Detecting transient cognitive impairment with EEG pattern recognition methods, *Aviation Space and Environmental Medicine*, **70,** 1018–1024.

Gevins, A. and Smith, M. E. 2000, Neurophysiological measures of working memory and individual differences in cognitive ability and cognitive style, *Cerebral Cortex*, **10,** 829–839.

Gevins, A., Leong, H., Smith, M. E., Le, J. and Du, R. 1995, Mapping cognitive brain function with modern high-resolution electroencephalography, *Trends in Neurosciences*, **18,** 429–436.

Gevins, A., Smith, M. E., Leong, H., McEvoy, L., Whitfield, S., Du, R. and Rush, G. 1998, Monitoring working memory load during computer-based tasks with EEG pattern recognition methods, *Human Factors*, **40,** 79–91.

Gevins, A., Smith, M. E., McEvoy, L. and Yu, D. 1997, High-resolution EEG mapping of cortical activation related to working memory: effects of task difficulty, type of processing, and practice, *Cerebral Cortex*, **7,** 374–385.

Gevins, A. S. 1980, Pattern recognition of brain electrical potentials, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **2,** 383–404.

Gevins, A. S. and Morgan, N. H. 1988, Applications of neural-network (NN) signal processing in brain research, *IEEE Transactions on Acoustics, Speech, and Signal Processing*, **36,** 1152–1161.

Gevins, A. S. and Schaffer, R. E. 1980, A critical review of electroencephalographic EEG correlates of higher cortical functions, *CRC Critical Reviews in Bioengineering*, **4,** 113–164.

Gevins, A. S., Bressler, S. L., Cutillo, B. A., Illes, J., Miller, J. C., Stern, J. and Jex, H. R. 1990, Effects of prolonged mental work on functional brain topography, *Electroencephalography and Clinical Neurophysiology*, **76,** 339–350.

Gevins, A. S., Doyle, J. C., Schaffer, R. E., Callaway, E. and Yeager, C. 1980, Lateralized cognitive processes and the electroencephalogram, *Science*, **207,** 1005–1008.

Gevins, A. S., Smith, M. E., Le, J., Leong, H., Bennett, J., Martin, N., McEvoy, L., Du, R. and Whitfield, S. 1996, High resolution evoked potential imaging of the cortical dynamics of human working memory, *Electroencephalography and Clinical Neurophysiology*, **98,** 327-348.

Gevins, A. S., Yeager, C. L., Diamond, S. L., Spire, J. P., Zeitlin, G. M. and Gevins, A. H. 1975, Automated analysis of the electrical activity of the human brain (EEG): a progress report, *Proceedings of the Institute of Electrical and Electronics Engineers*, **63,** 1382–1399.

Gevins, A. S., Yeager, C. L., Diamond, S. L., Spire, J. P., Zeitlin, G. M. and Gevins, A. H. 1976, Sharp-transient analysis and thresholded linear coherence spectra of paroxysmal EEGs, in P. Kellaway and I. Petersen (eds), *Quantitative Analytic Studies in Epilepsy* (New York: Raven Press), 463–481.

Gevins, A. S., Yeager, C. L., Zeitlin, G. M., Ancoli, S. and Dedon, M. 1977, On-line computer rejection of EEG artifact, *Electroencephalography and Clinical Neurophysiology*, **42,** 267–274.

Gevins, A. S., Zeitlin, G. M., Doyle, J. C., Schaffer, R. E. and Callaway, E. 1979a, EEG patterns during 'cognitive' tasks. II. Analysis of controlled tasks, *Electroencephalography and Clinical Neurophysiology*, **47,** 704–710.

Gevins, A. S., Zeitlin, G. M., Doyle, J. C., Yingling, C. D., Schaffer, R. E., Callaway, E. and Yeager, C. L. 1979b, Electroencephalogram correlates of higher cortical functions, *Science*, **203,** 665–668.

Gevins, A. S., Zeitlin, G. M., Yingling, C. D., Doyle, J. C., Dedon, M. F., Schaffer, R. E., Roumasset, J. T. and Yeager, C. L. 1979c, EEG patterns during 'cognitive' tasks. I. Methodology and analysis of complex behaviors, *Electroencephalography and Clinical Neurophysiology*, **47,** 693–703.

Gilliam, F., Kuzniecky, R. and Faught, E. 1999, Ambulatory EEG monitoring, *Journal of Clinical Neurophysiology*, **16,** 111–115.

Goldman-Rakic, P. 1987, Circuitry of primate prefrontal cortex and regulation of behavior by representational memory, in F. Plum and V. Mountcastle (ed.), *Handbook of Physiology, The Nervous System—Higher Functions of the Brain*, 1st edn, Vol. 5 (Bethesda, MD: American Physiological Society), 373–417.

Goldman-Rakic, P. 1988, Topography of cognition: parallel distributed networks in primate association cortex, *Annual Review of Neuroscience*, **11,** 137–156.

Gundel, A. and Wilson, G. F. 1992, Topographical changes in the ongoing EEG related to the difficulty of mental tasks, *Brain Topography*, **5,** 17–25.

Humphrey, D. and Kramer, A. F. 1994, Toward a psychophysiological assessment of dynamic changes in mental workload, *Human Factors*, **36,** 3–26.

Inouye, T., Shinosaki, K., Iyama, A., Matsumoto, Y., Toi, S. and Ishihara, T. 1994, Potential flow of frontal midline theta activity during a mental task in the human electroencephalogram, *Neuroscience Letters*, **169,** 145–148.

Ishii, R., Shinosaki, K., Ukai, S., Inouye, T., Ishihara, T., Yoshimine, T., Hirabuki, N., Asada, H., Kihara, T., Robinson, S. E. and Takeda, M. 1999, Medial prefrontal cortex generates frontal midline theta rhythm, *Neuroreport*, **10,** 675–679.

Isreal, J. B., Wickens, C. D., Chesney, G. L. and Donchin, E. 1980, The event-related brain potential as an index of display-monitoring workload, *Human Factors*, **22,** 211–224.

Jansma, J. M., Ramsey, N. F., Coppola, R. and Kahn, R. S. 2000, Specific versus nonspecific brain activity in a parametric n-back task, *Neuroimage*, **12,** 688–697.

John, E. R., Prichep, L. S., Kox, W., Valdes-Sosa, P., Bosch-Bayard, J., Aubert, E., Tom, M., diMichele, F. and Gugino, L. D. 2001, Invariant reversible QEEG effects of anesthetics, *Consciousness and Cognition*, **10,** 165–183.

Jonides, J., Smith, E. E., Koeppe, R. A., Awh, E., Minoshima, S. and Mintun, M. 1993, Spatial working memory in humans as revealed by PET, *Nature*, **363,** 623–625.

Joseph, R. D. 1961, Contributions of perceptron theory, unpublished PhD thesis, Cornell University, Ithaca, New York.

Jung, T. P., Makeig, S., Humphries, C., Lee, T. W., McKeown, M. J., Iragui, V. and Sejnowski, T. J. 2000, Removing electroencephalographic artifacts by blind source separation, *Psychophysiology*, **37,** 163–178.

Kennedy, J. L., Gottsdanker, R. M., Arinington, J. C. and Gray, F. E. 1948, A new electroencephalogram associated with thinking, *Science*, **108,** 527.

Kieras, D. E. 1988, Towards a practical GOMS model methodology for user interface design, in M. Helander (ed.), *The Handbook of Human–Computer Interaction* (Amsterdam: North-Holland), 135–158.

Koivisto, M., Krause, C. M., Revonsuo, A., Laine, M. and Hamalainen, H. 2000, The effects of electromagnetic field emitted by GSM phones on working memory, *Neuroreport*, **11,** 1641–1643.

Kok, A. 2001, On the utility of P3 amplitude as a measure of processing capacity, *Psychophysiology*, **38,** 557–577.

Kramer, A. F., Trejo, L. J. and Humphrey, D. 1995, Assessment of mental workload with task-irrelevant auditory probes, *Biological Psychology*, **40,** 83–100.

Kyllonen, P. C. and Christal, R. E. 1990, Reasoning ability is little more than working memory capacity?!, *Intelligence*, **14,** 389–433.

Kyllonen, P. C. and Shute, V. J. 1989, A taxonomy of learning skills, in P. L. Ackerman (ed.), *Learning and Individual Differences* (New York: Freeman), 117–163.

Larson, C. L., Davidson, R. J., Abercrombie, H. C., Ward, R. T., Schaefer, S. M., Jackson, D. C., Holden, J. E. and Perlman, S. B. 1998, Relations between PET-derived measures of thalamic glucose metabolism and EEG alpha power, *Psychophysiology*, **35,** 162–169.

McCallum, W. C., Cooper, R. and Pocock, P. V. 1988, Brain slow potential and ERP changes associated with operator load in a visual tracking task, *Electroencephalography and clinical Neurophysiology*, **69,** 453–468.

McCarthy, G., Blamire, A. M., Puce, A., Nobre, A. C., Bloch, G., Hyder, F., Goldman-Rakic, P. and Shulman, R. G. 1994, Functional magnetic resonance imaging of human prefrontal cortex activation during a spatial working memory task, *Proceedings of the National Academy of Science (USA)*, **91,** 8690–8694.

McElree, B. 2001, Working memory and focal attention, *Journal of Experimental Psychology: Learning, Memory and Cognition*, **27,** 817–835.

McEvoy, L. K., Pellouchoud, E., Smith, M. E. and Gevins, A. 2001, Neurophysiological signals of working memory in normal aging, *Cognitive Brain Research*, **11,** 363–376.

McEvoy, L. K., Smith, M. E. and Gevins, A. 1998, Dynamic cortical networks of verbal and spatial working memory: effects of memory load and task practice, *Cerebral Cortex*, **8,** 563–574.

McEvoy, L. K., Smith, M. E. and Gevins, A. 2000, Test-retest reliability of cognitive EEG, *Clinical Neurophysiology*, **111,** 457–463.

Miyata, Y., Tanaka, Y. and Hono, T. 1990, Long term observation on Fm-theta during mental effort, *Neuroscience*, **16,** 145–148.

Mizuki, Y., Tanaka, M., Iogaki, H., Nishijima, H. and Inanaga, K. 1980, Periodic appearances of theta rhythm in the frontal midline area during performance of a mental task, *Electroencephalography and Clinical Neurophysiology*, **49,** 345–351.

Molloy, R. and Parasuraman, R. 1996, Monitoring an automated system for a single failure: vigilance and task complexity effects, *Human Factors*, **38,** 311–322.

Morrison, J. G. and Gluckman, J. P. 1994, Definitions and prospective guidelines for the application of automation, in M. Mouloua and R. Parasuraman (eds), *Human performance in automated systems: Current research and trends* (Hillsdale, NJ: Lawrence Erlbaum Associates), 256–263.

Mulholland, T. 1995, Human EEG, behavioral stillness and biofeedback, *International Journal of Psychology*, **19,** 263–279.

O'Connor, M. F., Daves, S. M., Tung, A., Cook, R. I., Thisted, R. and Appelbaum, J. 2001, BIS monitoring to prevent awareness during general anesthesia, *Anesthesiology*, **94,** 520–522.

Oliveri, M., Turriziani, P., Carlesimo, G. A., Koch, G., Tomaiuolo, F., Panella, M. and Caltagirone, C. 2001, Parieto-frontal interactions in visual-object and visual-spatial working memory: evidence for transcranial magnetic stimulation, *Cerebral Cortex*, **11,** 606–618.

Olson, J. R. and Olson, G. M. 1990, The growth of cognitive modeling in human-computer interaction since GOMS, *Human-Computer Interaction*, **5,** 221-265.

Parasuraman, R., Molloy, R. and Singh, I. L. 1993, Performance consequences of automation-induced 'complacency', *International Journal of Aviation Psychology*, **3,** 1–23.

Parasuraman, R., Mouloua, M. and Molloy, R. 1996, Effects of adaptive task allocation on monitoring of automated systems, *Human Factors*, **38,** 665–679.

Parasuraman, R., Sheridan, T. B. and Wickens, C. D. 2000, A model for types and levels of human interaction with automation, *IEEE Transactions Systems, Man, and Cybernetics-Part A: Systems and Humans*, **30,** 286–297.

Paus, T., Koski, L., Caramanos, Z. and Westbury, C. 1998, Regional differences in the effects of task difficulty and motor output on blood flow response in the human anterior cingulate cortex: a review of 107 PET activation studies, *Neuroreport*, **9,** R37–R47.

Pellouchoud, E., Smith, M. E., McEvoy, L. and Gevins, A. 1999, Mental effort-related EEG modulation during video-game play: comparison between juvenile subjects with epilepsy and normal control subjects, *Epilepsia*, **40** (Suppl 4), 38–43.

Pfurtscheller, G. and Klimesch, W. 1992, Functional topography during a visuoverbal judgment task studied with event-related desynchronization mapping, *Journal of Clinical Neurophysiology*, **9,** 120–131.

Posner, M. I. and Peterson, S. E. 1990, The attention system of the human brain, *Annual Review of Neuroscience*, **13,** 25–42.

POSNER, M. I. and ROTHBART, M. K. 1992, Attentional mechanisms and conscious experience, in A. D. Milner and M. D. Rugg (eds), *The Neuropsychology of Consciousness* (San Diego: Academic Press), 91–111.

RASKIN, J. 2000, *Humane Interface: New Directions for Designing Interactive Systems* (Boston, MA: Addison-Wesley).

ROCKSTROH, B., ELBERT, T., CANAVAN, A., LUTZENBERGER, W. and BIRBAUMER, N. 1989, *Slow cortical potentials and behavior* (Baltimore: Urban & Schwarzenberg).

ROSS, P. and SEGALOWITZ, S. J. 2000, An EEG coherence test of the frontal dorsal versus ventral hypothesis of n-back working memory, *Brain and Cognition*, **43,** 375–379.

SADATO, N., NAKAMURA, S., OOHASHI, T., NISHINA, E., FUWAMOTO, Y., WAKI, A. and YONEKURA, Y. 1998, Neural networks for generation and suppression of alpha rhythm: a PET study, *Neuroreport*, **30,** 893–897.

SHEER, D. E. 1989, Sensory and cognitive 40 Hz event-related potentials, in E. Basar and T. H. Bullock (eds), *Brain Dynamics*, Vol. 2 (Berlin: Springer), 339–374.

SIREVAAG, E. J., KRAMER, A. F. and WICKENS, C. D. 1993, Assessment of pilot performance and mental workload in rotary wing aircraft, *Ergonomics*, **36,** 1121–1140.

SMITH, M. E., GEVINS, A., BROWN, H., KARNIK, A. and DU, R. 2001, Monitoring task load with multivariate EEG measures during complex forms of human computer interaction, *Human Factors*, **43,** 366–380.

SMITH, M. E., MCEVOY, L. K. and GEVINS, A. 1999, Neurophysiological indices of strategy development and skill acquisition, *Brain Research Cognitive Brain Research*, **7,** 389–404.

THOMPSON, J. L. and EBERSOLE, J. S. 1999, Longterm inpatient audiovisual scalp EEG monitoring, *Journal of Clinical Neurophysiology*, **16,** 91–99.

TOU, J. T. and GONZALEZ, R. C. 1974, *Pattern Recognition Principles* (Reading, MA: Addison-Wesley Publishing Co).

ULLSPERGER, P., FREUDE, G. and ERDMANN, U. 2001, Auditory probe sensitivity to mental workload changes—an event-related potential study, *International Journal of Psychophysiology*, **40,** 201–209.

VAN DEN BERG-LENSSSEN, M. M., BRUNIA, C. H. and BLOM, J. A. 1989, Correction of ocular artifacts in EEGs using an autoregressive model to describe the EEG: a pilot study, *Electroencephalography and Clinical Neurophysiology*, **73,** 72–83.

VESPA, P., NENOV, V. and NUWER, M. R. 1999, Continuous EEG monitoring in the intensive care unit: early findings and clinical efficacy, *Journal of Clinical Neurophysiology*, **16,** 1–13.

VIGLIONE, S. S. 1970, Applications of pattern recognition technology, in J. M. Mendel and K. S. Fu (eds), *Adaptive Learning and Pattern Recognition Systems* (New York: Academic Press), 115–161.

WILSON, G. F. and FISHER, F. 1995, Cognitive task classification based upon topographic EEG data, *Biological Psychology*, **40,** 239–250.

WILSON, G. F., FULLENKAMP, B. S. and DAVIS, I. 1994, Evoked potential, cardiac, blink, and respiration measures of pilot workload in air-to-ground missions, *Aviation, Space and Environmental Medicine*, **65,** 100–105.

WINTINK, A. J., SEGALOWITZ, S. J. and CUDMORE, L. J. 2001, Task complexity and habituation effects on frontal P300 topography, *Brain and Cognition*, **46,** 307–311.

WOLTZ, D. J. 1988, An investigation of the role of working memory in procedural skill acquisition, *Journal of Experimental Psychology: General*, **117,** 319–331.

YAMAMOTO, S. and MATSUOKA, S. 1990, Topographic EEG study of visual display terminal VDT performance with special reference to frontal midline theta waves, *Brain Topography*, **2,** 257–267.

About the authors

Alan Gevins is the founder and Executive Director of the San Francisco Brain Research Institute and the founder and President of SAM Technology, Inc., both in San Francisco. He is internationally known for developing algorithms and systems for analysing human brain function, and for basic science studies of human neurocognitive functions. He is the author of over 125 scientific papers and of 16 US patents.

Michael E. Smith is a cognitive neuroscientist who specializes in basic and applied research on the neural systems mediating human attention and memory. He has authored over 50 scientific papers on related topics. He holds an undergraduate degree from the University of Michigan, a PhD from the University of California, Los Angeles, and an MBA from the University of California, Berkeley. He currently directs the research department of SAM Technology, Inc.