ARL-RP-0581 ● OCT 2016

US Army Research Laboratory

Towards Serious Games for Improved BCI

by Brent J Lance, Jon Touryan, Yu-Kai Wang, Shao-Wei Lu, Chun-Hsiang Chuang, Peter Khooshabeh, Paul Sajda, Amar Marathe, Tzyy-Ping Jung, Chin-Teng Lin, and Kaleb McDowell

Reprinted from Handbook of Digital Games and Entertainment Technologies (Nakatsu R, Rauterberg M, Ciancarini P, editors). Singapore: Springer; c2015. p. 1–28. DOI: 10.1007/978-981-4560-52-8_4-1.

NOTICES

Disclaimers

ARL-RP-0581 ● OCT 2016

US Army Research Laboratory

Towards Serious Games for Improved BCI

by Brent J Lance, Jon Touryan, Peter Khooshabeh, Amar Marathe, and Kaleb McDowell
Human Research and Engineering Directorate, ARL

Yu-Kai Wang, Shao-Wei Lu, Chun-Hsiang Chuang, and Chin-Teng Lin
Brain Research Center, National Chiao Tung University, Taiwan

Paul Sajda
Department of Biomedical Engineering, Columbia University, New York, NY

Tzyy-Ping Jung
Swartz Center for Computational Neuroscience, University of California San Diego, San Diego, CA

Reprinted from Handbook of Digital Games and Entertainment Technologies (Nakatsu R, Rauterberg M, Ciancarini P, editors). Singapore: Springer; c2015. p. 1–28. DOI: 10.1007/978-981-4560-52-8_4-1.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
October 2016

2. REPORT TYPE
Reprint Report

3. DATES COVERED (From - To)
December 2014

4. TITLE AND SUBTITLE
Towards Serious Games for Improved BCI

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Brent J Lance, Jon Touryan, Yu-Kai Wang, Shao-Wei Lu, Chun-Hsiang Chuang, Peter Khooshabeh, Paul Sajda, Amar Marathe, Tzyy-Ping Jung, Chin-Teng Lin, and Kaleb McDowell

5d. PROJECT NUMBER
Translational Neuroscience

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
US Army Research Laboratory
ATTN: RDRL-HRF-C
Aberdeen Proving Ground, MD 21005

8. PERFORMING ORGANIZATION REPORT NUMBER
ARL-RP-0581

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
Reprinted from Handbook of Digital Games and Entertainment Technologies (Nakatsu R, Rauterberg M, Ciancarini P, editors). Singapore: Springer; c2015. p. 1–28. DOI: 10.1007/978-981-4560-52-8_4-1.

14. ABSTRACT

Brain-computer interface (BCI) technologies, or technologies that use online brain signal processing, have a great promise to improve human interactions with computers, their environment, and even other humans. Despite this promise, there are no current serious BCI technologies in widespread use, due to the lack of robustness in BCI technologies. The key neural aspect of this lack of robustness is human variability, which has two main components: (1) individual differences in neural signals and (2) intraindividual variability over time. In order to develop widespread BCI technologies, it will be necessary to address this lack of robustness. However, it is currently unknown how neural variability affects BCI performance. To accomplish these goals, it is essential to obtain data from large numbers of individuals using BCI technologies over considerable lengths of time. One promising method for this is through the use of BCI technologies embedded into games with a purpose (GWAP). GWAP are a game-based form of crowdsourcing which players choose to play for enjoyment and during which the player performs key tasks which cannot be automated but that are required to solve research questions. By embedding BCI paradigms in GWAP and recording neural and behavioral data, it should be possible to much more clearly understand the differences in neural signals between individuals and across different time scales, enabling the development of novel and increasingly robust adaptive BCI algorithms.

15. SUBJECT TERMS
robust brain-computer interaction, serious games, games with a purpose, Human Computation Game, EEG

16. SECURITY CLASSIFICATION OF:

a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES
Unclassified	Unclassified	Unclassified	UU	34

19a. NAME OF RESPONSIBLE PERSON
Brent J Lance

19b. TELEPHONE NUMBER (Include area code)
(410) 278-5943

Towards Serious Games for Improved BCI

Brent J. Lance, Jon Touryan, Yu-Kai Wang, Shao-Wei Lu, Chun-Hsiang Chuang, Peter Khooshabeh, Paul Sajda, Amar Marathe, Tzyy-Ping Jung, Chin-Teng Lin, and Kaleb McDowell

Contents

B.J. Lance (✉) • J. Touryan • A. Marathe • K. McDowell
Translational Neuroscience Branch, U.S. Army Research Laboratory, Aberdeen Proving Ground, Aberdeen, MD, USA
e-mail: brent.j.lance.civ@mail.mil; jonathan.o.touryan.civ@mail.mil; amar.marathe.civ@mail.mil; kaleb.g.mcdowell.civ@mail.mil

Y.-K. Wang • S.-W. Lu •
C.-H. Chuang • C.-T. Lin
Brain Research Center, National Chiao Tung University, Hsinchu, Taiwan
e-mail: yukaiwang@cs.nctu.edu.tw; swlucifer@mail.nctu.edu.tw; chchuang@mail.nctu.edu.tw; ctlin@mail.nctu.edu.tw

P. Khooshabeh
Cognitive Sciences Branch, U.S. Army Research Laboratory, Aberdeen Proving Ground, Aberdeen, MD, USA
e-mail: peter.khooshabehadeh2.civ@mail.mil

P. Sajda
Department of Biomedical Engineering, Columbia University, New York, NY, USA
e-mail: psajda@columbia.edu

T.-P. Jung
Swartz Center for Computational Neuroscience, University of California San Diego, San Diego, CA, USA
e-mail: jung@sccn.ucsd.edu

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-52-8_4-1

Abstract

Brain-computer interface (BCI) technologies, or technologies that use online brain signal processing, have a great promise to improve human interactions with computers, their environment, and even other humans. Despite this promise, there are no current serious BCI technologies in widespread use, due to the lack of robustness in BCI technologies. The key neural aspect of this lack of robustness is human variability, which has two main components: (1) individual differences in neural signals and (2) intraindividual variability over time. In order to develop widespread BCI technologies, it will be necessary to address this lack of robustness. However, it is currently unknown how neural variability affects BCI performance. To accomplish these goals, it is essential to obtain data from large numbers of individuals using BCI technologies over considerable lengths of time. One promising method for this is through the use of BCI technologies embedded into games with a purpose (GWAP). GWAP are a game-based form of crowdsourcing which players choose to play for enjoyment and during which the player performs key tasks which cannot be automated but that are required to solve research questions. By embedding BCI paradigms in GWAP and recording neural and behavioral data, it should be possible to much more clearly understand the differences in neural signals between individuals and across different time scales, enabling the development of novel and increasingly robust adaptive BCI algorithms.

Keywords

Games with a purpose (GWAP) • Human variability • Brain-computer interface (BCI) • Electroencephalography (EEG)

Introduction

Brain-computer interface (BCI) technologies, or technologies that use online brain signal processing, have a great promise to improve human interactions with computers, their environment, and even other humans (Lance et al. 2012). In particular, BCI technologies have a strong likelihood for enhancing or revolutionizing several fields, including training and education, medicine, and communication technologies. Broadly speaking, there are four classes of BCI systems that utilize neural data to estimate or predict behavior: active, reactive, passive, and hybrid (Lance et al. 2012; Wolpaw et al. 2002; Zander and Kothe 2011). Active BCIs rely on neural signals intentionally induced by the operator, typically after some degree of training or associative mapping. These volitionally generated patterns of brain activity are then used as an input modality to directly control a computer, prosthetic limb, or some specific function of a given technological system (e.g., sensorimotor rhythms for control of virtual movement (McFarland et al. 2010)). Reactive BCIs rely on neural responses elicited by an external stimulus or event, which are then used as an input modality to a technological system (e.g., P300 spellers (Krusienski et al. 2006) and for multimedia image triage and search systems (Sajda et al. 2010; Pohlmeyer et al. 2011)). Passive BCIs, including cognitive monitoring systems, utilize implicit or ongoing neural responses for the purposes of detecting an operator's current cognitive or affective state. Indicators of fatigue (Balasubramanian et al. 2011), mental workload (Mühl et al. 2014), and emotional state (Bos et al. 2010) can be used as supplemental information to adaptively enhance or augment performance of human-system interactions. Finally, hybrid BCIs may utilize any combination of active, reactive, and passive approaches. In addition, hybrid BCIs may integrate other signals reflective of operator state, including eye movement dynamics or behavioral performance, as well as incorporate context awareness information obtained from the environment or technological system itself (Zander and Jatzev 2012; Pfurtscheller et al. 2010; Jangraw et al. 2014).

This chapter is written from the viewpoint of having a strong interest in the development of serious (i.e., primarily nonentertainment in purpose), nonmedical BCI technologies that improve the performance of healthy individuals. As a result, this chapter focuses on BCI technologies that are noninvasive, i.e., based on sensors which are not within the user's body. While some medical BCI technologies are invasive, i.e., based on sensors which are actually within the user's body, we consider noninvasive methods to be a more practical and user-acceptable approach for nonmedical BCIs. In addition, this chapter focuses on mobile neuroimaging technologies, primarily EEG. While the ability and precision of measuring brain activity have been dramatically improved in recent years due to the advances in neuroimaging technology, including functional magnetic resonance imaging (fMRI), magnetoencephalography (MEG), and functional near-infrared spectroscopy (fNIRS); electroencephalography (EEG) remains the most promising near-term method for realizing broadly available BCI. For all classes of BCI, the majority of approaches, algorithms, and systems continue to utilize EEG as the neuroimaging modality of choice.

BCI technologies have seen expanding usage primarily within the medical and entertainment domains (van Erp et al. 2012). For example, BCI technologies have been used in the development of medical technologies that provide a clinical population with a capability that they have lost, such as the BCI-controlled wheelchair (Vanacker et al. 2007) or P300 speller (Krusienski et al. 2006), which enables the disabled to slowly type via neural signals alone. They have also been applied to the development of entertainment technologies such as the Emotiv EPOC, a low-cost EEG headset designed to be used as a game controller, or the Force Trainer toy by Uncle Milton, which uses a single EEG electrode to control a fan and blow a light plastic ball through an obstacle course. Despite these successes, there are essentially no current BCI technologies in widespread use. This is because developing serious, nonmedical BCI technologies is a fundamentally different problem space than developing either medical or entertainment BCIs. For example, medical BCI technologies do not have to perform to the level of a healthy individual's motor control system to be useful. When comparing normal mouse movement to BCI-controlled mouse movement, Hochberg et al. (2006) found that users could reach a target on a screen 100 % of the time in under 1.3 s with a mouse but could only reach the target 75 % of the time in under 7 s using a BCI to control the mouse cursor. While this may be acceptable for some clinical populations, entertainment BCIs that perform poorly provide for frustrating gameplay (Dakan 2010), but have limited real-world effect. Serious, nonmedical BCI technologies will need to perform to a higher standard than either medical or entertainment BCI technologies and may also need to emphasize different approaches (see Box 1).

Box 1: A Brief Overview of Relevant BCI Paradigms

There are many different BCI paradigms, several of which are described elsewhere in this volume. The current most commonly used BCI paradigm is the motor imagery paradigm, where a subject attempts to directly control the movement of an effector by imagining the movement of an extremity such as a hand or a foot. We take the view that this paradigm is poorly suited to serious, nonmedical BCI technologies due to its poor performance when compared to existing control technologies (Hochberg et al. 2006). As a result, in this chapter, we focus on three different BCI paradigms: the steady-state visually evoked potential (SSVEP) paradigm, the rapid serial visual presentation (RSVP) paradigm, and a specific driving performance estimation paradigm. This section provides a brief overview of each of these paradigms, including analytical methods and potential applications.

The SSVEP paradigm is a reactive BCI paradigm and involves detecting which one of a number of objects the subject is attending to on a screen by rapidly flickering the objects at distinct frequencies and phases. While the subject's attention is focused on a specific object flickering at a specific frequency, a phase-locked increase in the spectral power of that frequency can be detected in the visual cortex. Even through using simple methods such as a fast Fourier transform (FFT) followed by classification, SSVEP can be detected fairly reliably, rapidly, and with minimal time spent training either the player or the classifier, due to its high signal-to-noise ratio (SNR). More advanced methods, such as canonical correlation analysis, allow for faster processing, a broader range of frequencies, and multiple stimuli at the same frequency but differing phases (Bin et al. 2009). One area in which SSVEP-based BCI technologies have been successfully used is in spelling applications that allow paralyzed patients to communicate.

The RSVP paradigm is also a reactive BCI paradigm, but it consists of subjects looking for targets in a rapidly presented stream of images. In RSVP, images are presented to a subject at a rate of about 5–10 Hz. These images

(continued)

either belong to a sparsely appearing target class or are distracter images which do not belong to that target class. When the subject sees an image in the target class, their brain will generate an evoked neural response which can be detected using a machine learning classifier. Successful approaches include ensembles of logistic regression classifiers or spatial filtering followed by binary classification. Generally, the target images can make up no more than 15–20 % of the total number of images. Otherwise, two things occur: First, the strength of the neural signal will begin to attenuate, and as a result, classifier performance will begin to degrade. Second, the subject will begin to increasingly make mistakes as a result of an effect known as the "attentional blink." What this means is that, if two target images occur too close together in time, the subject has a drastically increased probability of missing the second target. RSVP has two key strengths: first, the neural signal evoked when the subject sees a target can be recognized by a classifier with a high degree of accuracy. Second, the neural signal can be evoked by a variety of targets, leading to a potentially broad space of applications. One area in which RSVP has been successfully applied is in the domain of satellite image analysis. By using RSVP to pre-triage satellite images, targets such as airfields can be detected more rapidly than with manual search (Sajda et al. 2010). Figure 1 shows the neural response from an RSVP task. Within this neural response can be seen an SSVEP-like neural signal induced by the flashing of the images in the RSVP task.

While there are many passive BCI paradigms for neurally estimating performance at simulated driving tasks, we will be focusing on the paradigm described by Lin et al. (2005). In this task, the subject drives a car down a simulated road and attempts to stay in the lane through periodic but random "lane-deviation events," which consist of a lateral force that attempts to push the driver off of the road. The intended goal of this paradigm is to estimate when driving performance decrements are more likely to occur, potentially enabling them to be mitigated and thus decreasing the likelihood of accidents. These lane-deviation events provide a probe to evaluate the performance of the driver as their fatigue and attention fluctuate and as time-on-task increases.

In order for BCI technologies to achieve this higher standard of performance, they must be able to handle or adapt to human variability, both in terms of individual differences and intraindividual variability. Every individual is different, as are their neural signals, even those generated by individuals in similar mental states or being exposed to the same stimuli. In practice, what this means is that the feature extraction and classification methods underlying a BCI technology will perform very poorly unless they are calibrated or customized for each individual user. In fact, a percentage of the population may even be "BCI illiterate" in that BCI technologies cannot be effectively trained on the neural signals of those individuals

(Pfurtscheller et al. 2010). People also have significant intraindividual variability, which can lead to differences in neural signals within the same individual during similar mental states or in reaction to the same stimuli at different points in time (Wu et al. 2010). Therefore, static feature extraction and classification methods that are customized to the individual user will still have large fluctuations in performance over time. However, there are currently few BCI technologies that attempt to adapt to the neural states of the individual users.

For serious, nonmedical BCI technologies to be viable, it will be crucial to develop algorithms that can address both aspects of human variability by adapting to the individual user and to the changes in their neural signals at multiple time scales. To develop these algorithms, it will be necessary to study how individual differences and variability in neural response affect BCI performance, which will require data from large numbers of individuals using BCI technologies over considerable lengths of time. One promising method for obtaining this needed data is through the use of BCI technologies embedded into *games with a purpose* (GWAP). As defined by von Ahn (2006), GWAP are a game-based form of human computation which players choose to play for enjoyment. As a side effect of the gameplay, in GWAP the player performs key tasks which cannot be automated but that are required to solve research questions (von Ahn 2006). In this case, the player will interact with BCI paradigms embedded into the gameplay while recording neural and behavioral data. Through the analysis of this data, it should be possible to much more clearly understand the differences in neural signals between individuals and across different time scales.

However, the development of BCI-based GWAP to study the effects of individual differences and intraindividual variability on BCI performance faces significant challenges. First, among these is the fact that the individual differences and intraindividual variability being studied will have significant negative effects on the performance of the BCI paradigms embedded in the GWAP. The challenge arises from the fact that forcing the player to continually use BCI paradigms that do not function due to low performance will make the game unenjoyable and non-immersive. As a result, the GWAP will not attract a large player base that spends significant time with the game, and there will not be sufficient data to make the GWAP worthwhile. In this chapter, we argue that this challenge can be addressed by making the BCI paradigms peripheral to the main gameplay experience and sufficiently rewarding to the player.

The remainder of this chapter will review the research on individual differences and human variability and their effects on the robustness of BCI technologies, the potential value of GWAP for long-term and large-scale data collection, and the key challenges for developing GWAP for collecting large-scale and long-term BCI data.

Human Variability and BCI Performance

As previously mentioned, humans are highly variable. People display significant individual differences, with each individual generating distinct neural signals, even under highly similar conditions. People also show strong intraindividual variation,

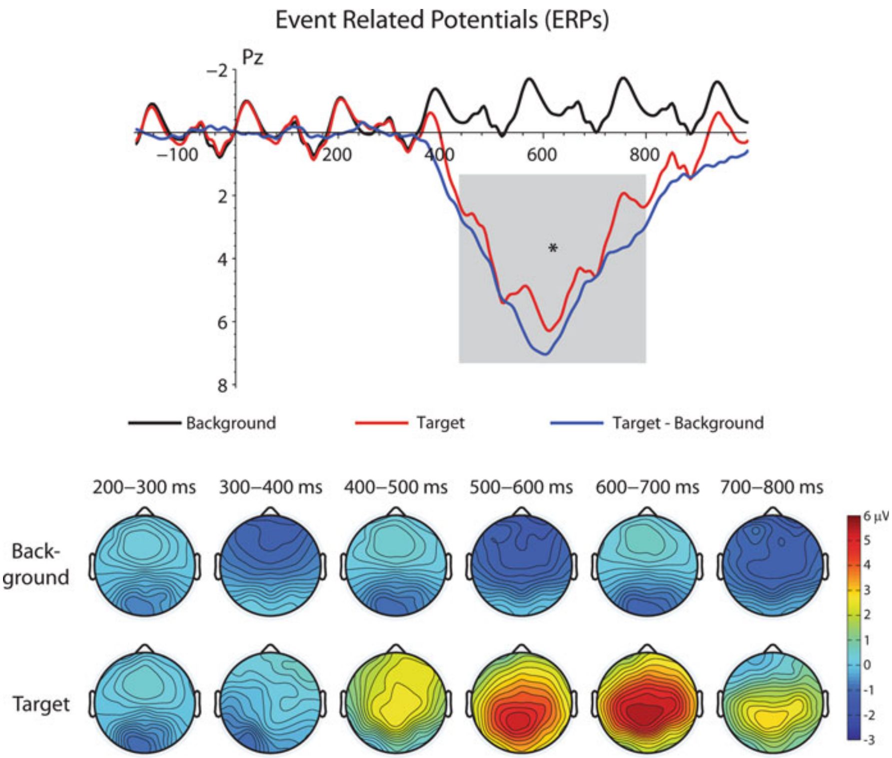

Fig. 1 Example neural signals from a 5-Hz RSVP task. In this task, a person is attempting to discriminate target images from background images in a continuous image stream. Target images are sparsely appearing images that the subject is tasked with identifying. Background images are commonly appearing images that the subject is tasked with ignoring. The three curves show 1 s of averaged neural response time-locked to image presentation. The *black line* shows neural response to a background image, the *red line* shows neural response to a target image, and the *blue line* shows the difference between the two. Slight but regular positive and negative deflections induced by image presentation which are similar to an SSVEP response can be seen in both the target and background neural responses. In addition, a large positive target-evoked inflection that peaks around 600 ms can be seen in the target neural response. The difference wave clarifies this target-evoked response by subtracting out the SSVEP-like response. The scalp maps show the location of the averaged neural responses as recorded on the scalp. Further description of this specific task and data can be found in Touryan et al. (2014)

with neural signals varying across similar conditions at multiple time scales. A hypothetical example of this can be seen in Fig. 2.

In this section, we will provide a review of human variability, including both individual differences and intraindividual variability, and how this variability affects BCI performance. We will focus on three distinct examples, showing how both individual differences and intraindividual variability affect the underlying neural signals that BCI technologies rely upon. First, we will discuss the effects of individual differences and intraindividual variation on a three-image-class RSVP

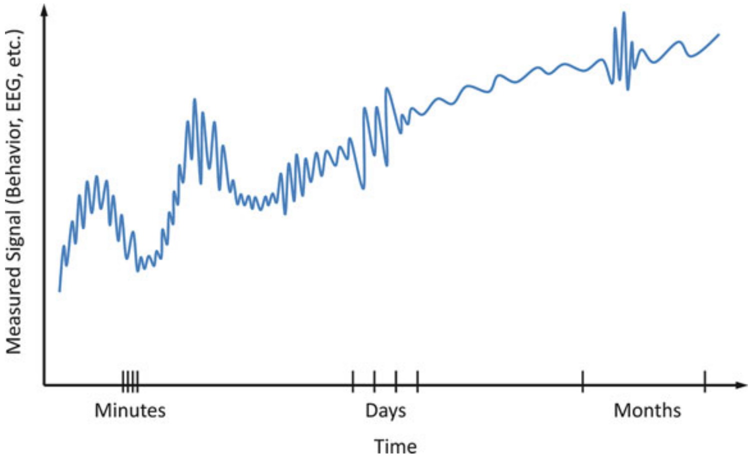

Fig. 2 Hypothetical example of intraindividual variability over time. Variability occurs in the course of minutes and hours, days and weeks, and even over months. Some of this fluctuation is internally driven, such as short-term fluctuations in attentional focus, while some of it is in response to changes in the external environment. Some of this variability is cyclic and predictable, such as circadian rhythm and sleep cycles; most of it is not. The cumulative result of this variability is neural signals that are nonstationary (i.e., the statistical properties of the signal change over time) and non-ergodic (i.e., the signal may not return to previously seen states), negatively impacting BCI performance

task (see Box 1). We will then discuss how the prior state of an individual affects their neural signals in a binary RSVP task and how this effect is modulated by individual differences. Finally, we will give an example of how individual differences and intraindividual variability affect the estimation of task performance from neural data.

Human Variability Example #1: Evoked Potentials

A specific example of both individual differences and intraindividual variability from three different people performing an RSVP task (see Box 1) with three categories of image stimuli, target, nontarget, and background images, can be seen in Fig. 3. In this task, target images are sparsely appearing images that the subject is tasked with identifying. Nontarget images are sparsely appearing images that are visually similar to the targets but which the subject is tasked with ignoring. Background images are commonly appearing images that the subject is also tasked with ignoring. In Fig. 3, the top rasters show the single-trial evoked neural responses time-locked to target images. Each horizontal row of the raster contains the single-trial neural response to one target appearance, and each vertical column contains one sample of EEG data. The color represents the amplitude of the neural response, with blue pixels representing lower-amplitude EEG recordings, red pixels representing higher amplitude, and green and yellow representing intermediate

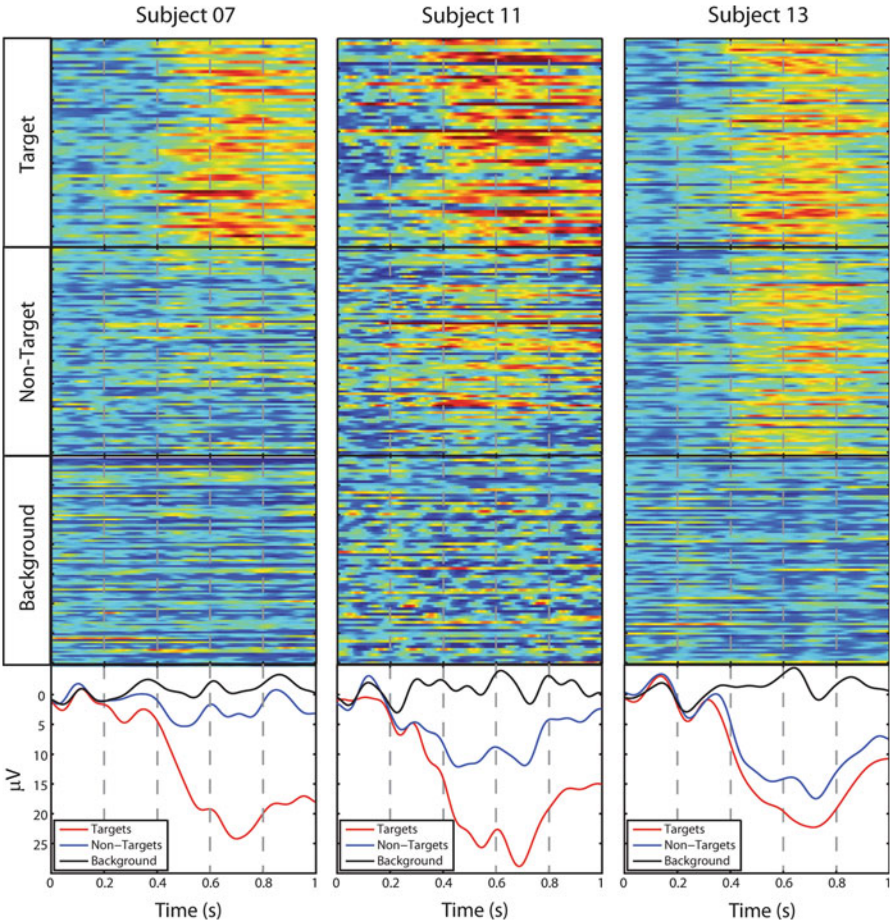

Fig. 3 An example of human variability in both single-trial evoked neural responses and averaged ERPs from an RSVP task. Each column contains the data from a different subject. Top three rows show rasters of the single-trial evoked neural responses time-locked to target, nontarget, and background images, respectively. Each horizontal row of the raster contains the single-trial neural response to one target appearance, and each vertical column contains one sample of EEG data. *Blue* pixels represent lower-amplitude EEG recordings, *red* pixels represent higher amplitude, and *green* and *yellow* represent intermediate amplitude. Bottom row shows the averaged event-related potentials (*ERPs*) from the three preceding rasters

amplitude. Similarly, the next two rasters show the neural responses to nontarget and background images, respectively.

Finally, the bottom images show the averaged event-related potentials (ERPs) from the three preceding rasters. ERPs are a traditional method for handling signal variability in EEG by averaging the EEG signal that occurs immediately after a stimulus (Geisler et al. 1958). Given enough trials where the data is time-locked to stimulus onset, the averaging will smooth out the variation that is not phase-locked to the stimulus and leave only the common phase-locked neural response. The non-phase-locked variation removed by this approach is a combination of intraindividual variation and the low signal-to-noise ratio (SNR) of EEG.

In Fig. 3, intraindividual variability in the recorded neural activity can be seen in the row-to-row differences in both the peak amplitude of the neural signal and the poststimulus latency of that peak within each raster. This row-to-row variability causes the noisy appearance of the rasters. Clear individual differences between the amplitude, latency, and even the amount of variability in the neural signals can also be seen in the different subjects' rasters. For example, Subject 11 has considerably more variability in their rasters than does Subject 13. The averaged ERPs at the bottom of Fig. 3 also reveal a very interesting pattern of individual differences in the neural response to the nontarget images. Specifically, it shows that for some subjects (such as Subject 07), the neural response to the sparse nontarget images was indistinguishable from the response to the sparse target images. For other subjects (Subject 13), the neural response to the nontarget stimuli was indistinguishable from the background images, while other subjects (Subject 11) were between these end points.

Human Variability Example #2: Effect of Prior State on Evoked Neural Signals

As the state of an individual changes due to intraindividual variability, this state affects their neural signals in a way that is modulated by individual differences. The concept of a "state" in this instance is intentionally ambiguous. Consider fatigue as an example of state change. Over time, individuals will have levels of decreased and increased fatigue. Different levels of fatigue have clear effects on an individual's neural signals (Gevins et al. 1977). However, from the perspective of BCI development, coping with these effects is extremely difficult for several reasons. First, fatigue is a very loaded term, potentially being used to mean many different but interrelated states, including sleep deprivation, physical exhaustion, time on task, visual fatigue, mental fatigue, boredom, and inattention (Matthews and Desmond 2002; Jung et al. 1997). While each of these different states will affect the neural response to stimuli, the effects are often concurrent and not clearly distinguishable. In addition, they will have differing effects on different individuals and will even affect the same person differently based on additional factors such as the stressfulness or consequences of their current situation.

Further, BCI research tends to use a limited number of approaches to address fatigue. The most common approach is avoidance. Most BCI research is done with experimental sessions that are sufficiently short and far apart in order to minimize the effects of task-related fatigue. When this does not occur, another approach often used is to discard the data recorded while the subject was fatigued and focus exclusively on the period of time when the subject was alert. Approaches like these are used to minimize the effects of fatigue because the alternative is BCI systems that fail to perform reliably. As a result, the actual quantitative effect of fatigue on BCI algorithms and technologies is poorly understood.

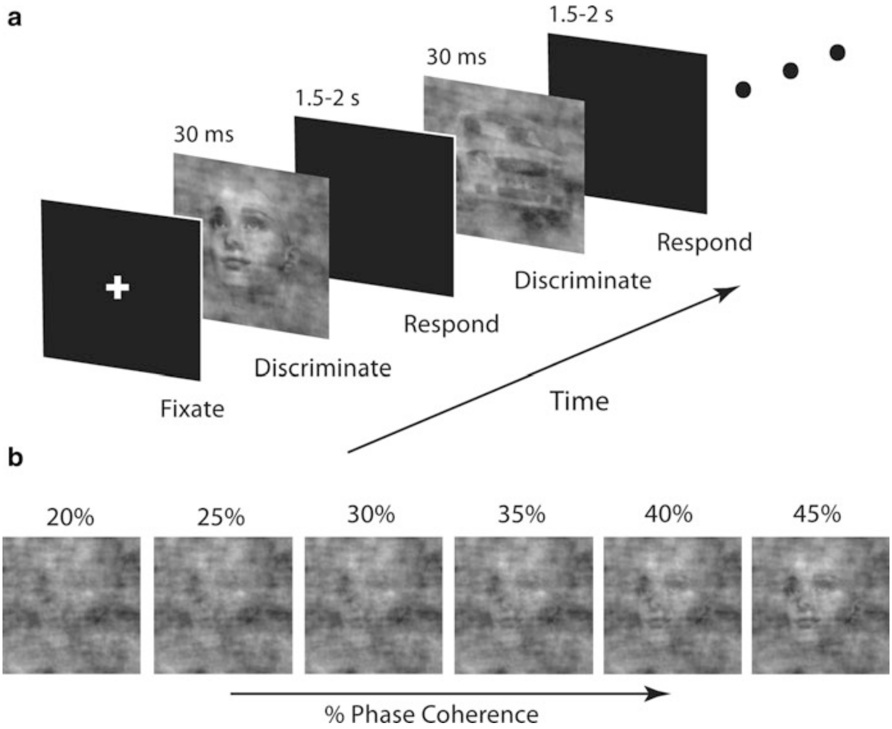

Fig. 4 Pre-stimulus cognitive state modulates sensory encoding of rapidly presented visual stimuli. Rapid serial visual presentation (*RSVP*) task in which subjects must classify the category of an object in an image corrupted by visual noise. The difficulty of the task is manipulated by controlling the noise level, through manipulating the images' phase coherence. Shown are three-phase coherence levels, with a lower coherence resulting in more noise in the image

Fatigue is not the only state that a person can be in which affects their neural signals. Many other factors of intraindividual variability, including emotion, stress, and prior experience, can modify neural signals and thereby affect BCI performance. It is not at all clear how these factors all affect neural signals independently and even less understood how they are interrelated. The interrelation and complexity of these factors mean that it can often be easier to consider abstract states, defined as changes in neural signal during a task.

Consider the example shown in Fig. 4, a modified RSVP task where the objective is to classify rapidly presented images masked with visual noise. In this task, the image is "flashed" for 30 ms, and the subject must respond with a button press to determine if the image was from category 1 (face) or from category 2 (car). The goal of the BCI experiment is to classify whether or not the user correctly identified the category of the image, using only the neural data. This is done by using a windowed logistic regression to identify "discriminating components" in the post-image presentation neural data that can be used to identify the image class

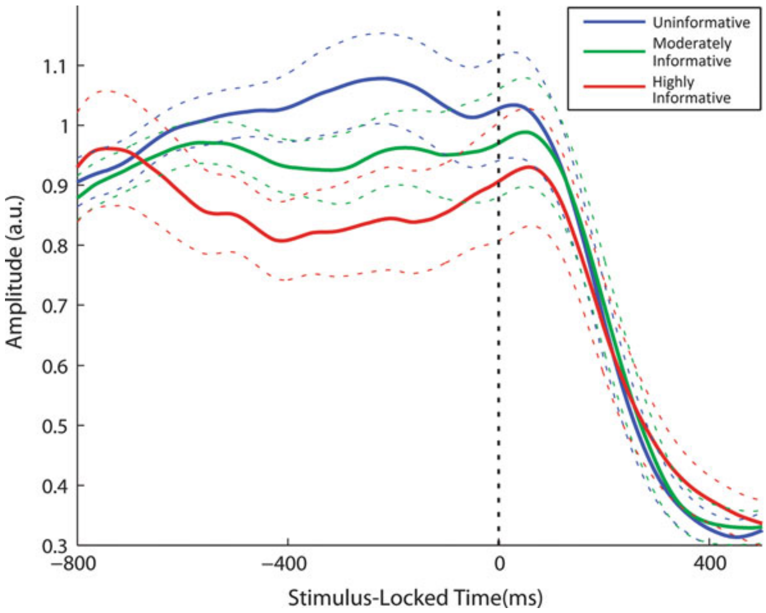

Fig. 5 Peak-to-peak amplitude of alpha power, grouped by the strength of the early poststimulus component. Greater strength of the early poststimulus component is associated with lower pre-stimulus alpha amplitude. Note that shortly after image presentation, alpha power drops sharply and is no longer useful for predicting targets. Instead, it becomes necessary to rely on the discriminating components identified using logistic regression (Reprinted from NeuroImage, 87, Lou et al., Prestimulus Alpha power predicts fidelity of sensory encoding in perceptual decision making, 242–251, Copyright (2014), with permission from Elsevier)

at a statistically significant level (Lou et al. 2014). This process identified two primary discriminating components: one that is early (appearing approximately 170 ms after the image was displayed) and one that is late (appearing approximately 350 ms after image presentation).

Interestingly, the strength of these post-image presentation discriminating components is directly affected by the pre-image presentation state of the subject, as defined by the oscillations that occur in the alpha frequency band (7–13 Hz) prior to the display of the stimulus image. These oscillations may reflect idling processes in the brain, with occipital (visual) alpha potentially serving as an index of attention. High alpha power is associated with idling and a resulting reduction in visual attention, while low alpha power represents engagement with the visual stimulus. However, the spectral power of EEG frequencies is highly variable over time. Thus, a subject might have different behavioral and neural responses evoked by the exact same visual stimulus, depending in part on their pre-stimulus alpha state. Figure 5 illustrates the variability of pre-stimulus alpha oscillations as a function of the strength of the early poststimulus discriminating component. Trials where the

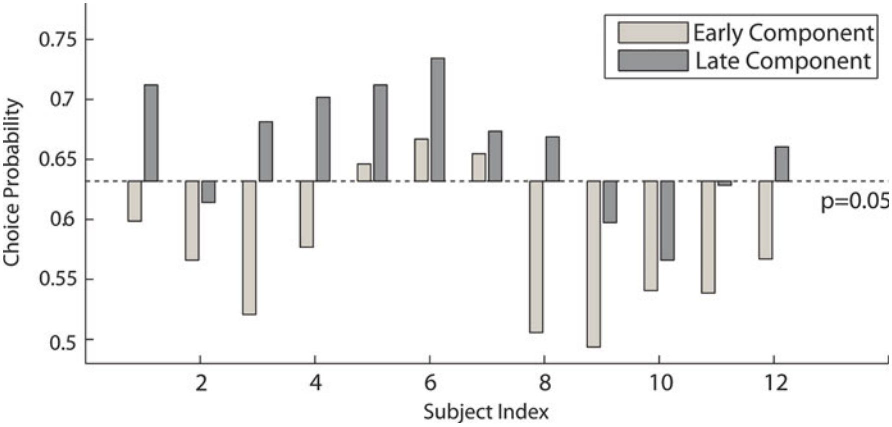

Fig. 6 Decoding of EEG results in two poststimulus evoked components that predict the subjects' choice: an early component and a late component. Shown is the decoding accuracy (chance is choice probability of 0.5) across 12 subjects for the two components. The early component, occurring around 170 ms poststimulus, is less of a predictor of the decision than the late component, which occurs around 350 ms poststimulus (Reprinted from NeuroImage, 87, Lou et al., Prestimulus Alpha power predicts fidelity of sensory encoding in perceptual decision making, 242–251, Copyright (2014), with permission from Elsevier)

pre-stimulus alpha amplitude is low will tend to result in highly informative poststimulus evoked responses (red curve in Fig. 5). Conversely, higher pre-stimulus alpha results in less informative evoked responses (blue curve in Fig. 5). This relationship between the pre-stimulus alpha and the evoked components is true for the early component but not the late component.

Figure 6 shows that, in addition to being affected by the state of the subject, the information content of the two discriminating components also varies substantially across the subjects. While the late component is more informative than the early component for every subject, for some subjects (e.g., S06), both components are significantly predictive. For other subjects (e.g., S09) neither is, and for some subjects (e.g., S01) only the late component significantly predicts performance.

Human Variability Example #3: Estimating Behavioral Performance

For the third example, consider a passive BCI that attempts to estimate performance at a task, based solely on neural signals (see Box 1). Intuitively, it seems like developing a system that can estimate behavioral performance at a task from neural signals should be straightforward. After all, the neural signals are the underlying driver that causes the behavioral task performance. However, due to the effects of human variability; the behavioral task performance, the neural data being used to estimate the behavior, and the relationship between the two are all widely variable, both within and across individuals. As a result of this variability, robustly predicting task performance using neural signals remains a promising but elusive goal.

Figure 7a demonstrates the variability of behavioral performance by showing the change in average relative reaction times to a lateral force across 80 90-min simulated driving sessions. This behavioral reaction time was smoothed using a causal 90-s bell-shaped moving average filter to eliminate variance at cycle lengths shorter than 1–2 min. The results follow a well-known trend of vigilance data, with three distinct states clearly visible in the data: initial near-perfect (relative RT $= 1$) performance begins to decay after about 1 min. Thereafter, reaction time to lane-deviation event rate rises steadily until $\sim$10 min into the task, after which it remains more or less stable near three to four times the optimal RT. While this group mean shows a strong time-on-task effect in this sustained attention task, the fluctuations in task performance during a single session could largely differ from the group trend and from those of other sessions.

Figure 7b shows that the behavior was highly variable both by individual and over time. For example, some subjects had poor performance in the first half of the session (e.g., S07), while others had exactly the opposite (e.g., S44). Similarly, the relationship between neural activity and behavior was also highly variable across subjects and over time. Figure 7c shows the correlations between EEG log power spectrum and behavioral performance at the simulated driving task at each EEG frequency. Some subjects had higher correlations between neural signals and behavior in the lower frequency bands (e.g., S49), while, for others, the higher frequency bands were more informative (e.g., S75). In addition, as shown in Fig. 7d, the relationship between EEG spectra and the behavioral reaction time varies as the subject performs the task. While increases in both the theta (4–7 Hz) and alpha (8–15 Hz) bands are generally predictive of periods of slower reaction time, they do not seem to predict just how slow the reaction time will be and, thus, how problematic the period of slow reaction times is for the driving task.

Across all three of these examples, both individual differences and intraindividual variability significantly affect the neural signals of BCI users. In terms of implications for robust and reliable BCI, this means that systems must be able to be calibrated to the individual user and must be able to adapt to the variability that occurs in both the individual's neural signals and task performance. These are known problems in many areas of science and engineering, and several techniques have been developed in order to address them, often by attempting to capitalize on the similarities between and across individuals.

These include machine learning approaches such as transfer learning, active learning, online learning, deep learning, and data mining approaches such as collaborative filtering. Transfer learning is a class of machine learning methods where data, features, or models from one domain are used to make predictions in another domain and which can be used to transfer information learned from one subject to help understand data collected from another subject or the same subject at a later time (Pan and Yang 2010). Active learning methods select the most informative data for labeling and incorporation into a model and can be used to more rapidly calibrate systems to individual users (Settles 2010). Online and adaptive

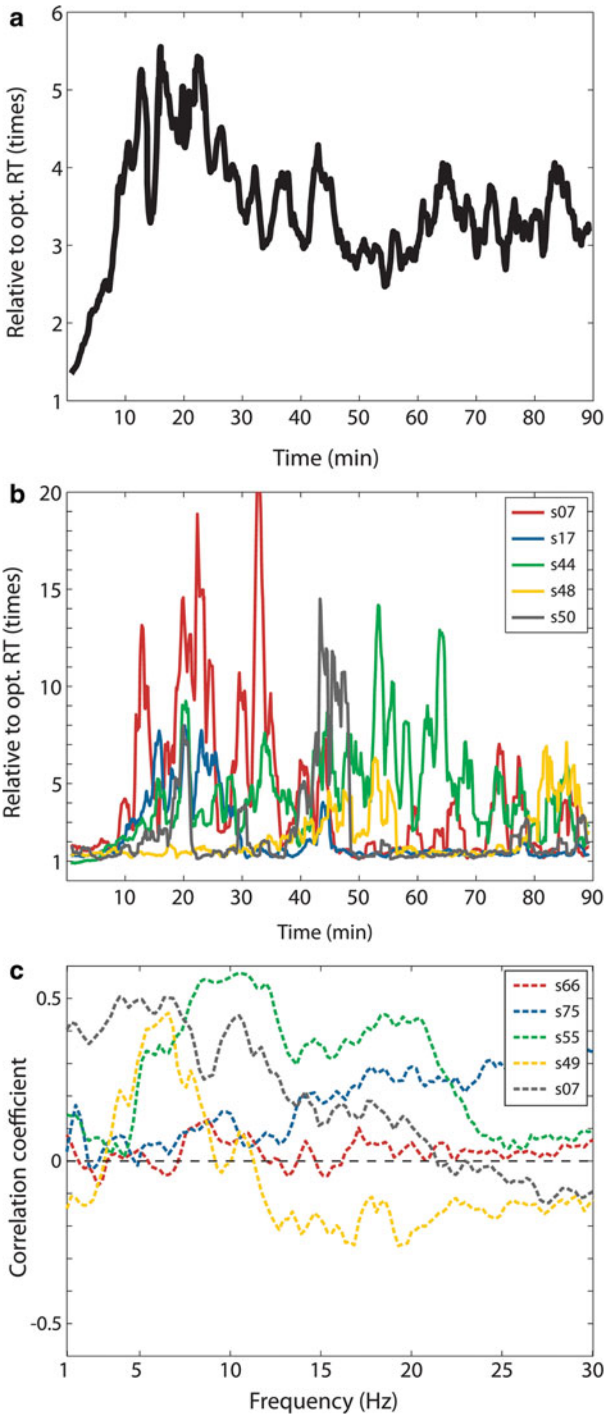

Fig. 7 (continued)

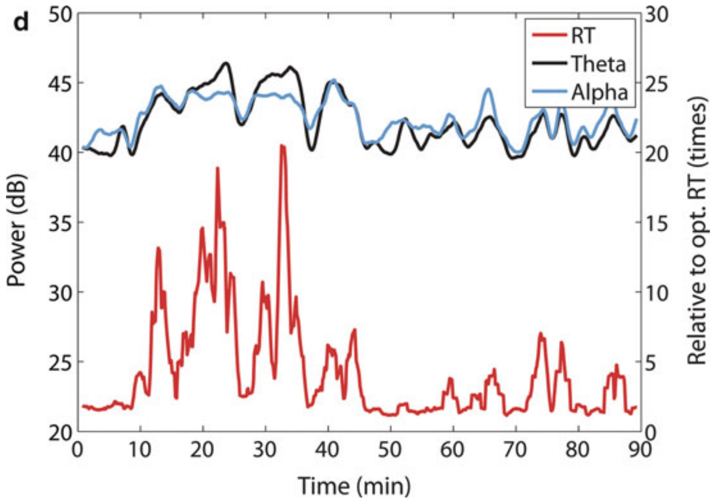

Fig. 7 (**a**) Group mean relative reaction time to lane-deviation events averaged across 80 90-min simulated driving experiments. (**b**) Relative reaction times to lane-deviation events across five example subjects, showing the differing variability of individual behavior. (**c**) Correlation between EEG spectral power and reaction time for five individual simulated driving sessions. Correlations are computed separately for 120 EEG frequencies between 0.98 and 30 Hz. (**d**) Reaction time, theta band power, and alpha band power over the course of one experimental session for subject S07

learning methods continually update as the system is used, attempting to adapt to the user. Deep Learning methods attempt to identify the most meaningful representation of the data and develop models based on that representation. Collaborative filtering approaches attempt to identify the data in a data set that is most relevant to a specific individual based on that individual's similarities and differences from a set of other known individuals.

One key aspect that unites all of these distinct methodologies is that their performance tends to increase as the amount of data available to train and infer upon increases. To date, however, the amount of data that is available from individuals using BCI technologies is highly limited, particularly with regard to serious, nonmedical BCI technologies. We argue that using these approaches to develop robust and reliable BCIs will require large amounts of data from large numbers of individuals using BCIs over long periods of time. In order to collect this large-scale and long-term data, it will be necessary to embed BCI paradigms and technologies into systems that people either need to use or will want to use. As a result, we argue that it should be possible to embed BCI paradigms into a video game in order to obtain the required data. Games have previously been utilized for both long-term and large-scale data collection. Often, these results are achieved through the use of *games with a purpose* or GWAP (Box 2).

Box 2: The Use of Games for Large-Scale and Long-Term Data Collection

Perhaps the earliest example of a successful large-scale GWAP is the ESP game, also known as the Google Image Labeler game (von Ahn 2006). In this game, two randomly selected players would be shown the same set of images. Their task was to provide as many labels for each image as possible within a short period of time, without communicating with the other player. If the labels provided by the two players matched, then both players received points. Players had profiles with persistent scores, and Google maintained high-score lists for the game. The image labels provided by these players served as the ground truth labels for searching the images with Google Image Search.

Another early example of large-scale data collection using a game was the deployment of the tracking real-time user experience or TRUE system (Kim et al. 2008), by Bungie Studios during the development of Halo 3. The TRUE system logged large quantities of data, including user interactions, contextual information, and attitudinal data from users who playtested Halo 3. While not technically a game with a purpose, the data collected using the TRUE system allowed identification of problematic and frustrating areas in the game, as well as improved fairness in multiplayer matches, resulting in an overall more enjoyable gaming experience (Thompson 2007).

A more recent game with a purpose is the Airport Scanner game for Android and iPhone, which simulates the task of searching for contraband using an airport X-ray scanner. To date, anonymous data from thousands of users performing millions of scans have been collected from players of the game. Researchers have used this data to explore questions about how well the airport baggage screening task can be performed. In particular, they found that the less often a target appeared, the less likely it was to be seen by the players. This was true even when looking at the high scoring subset of players (Mitroff and Biggs 2013). Interestingly, this low-probability, high-consequence detection problem is exactly what Transportation Security Administration (TSA) screeners encounter on a daily basis but is very hard to replicate in a relatively short laboratory experiment.

Games have also been used to study performance at cognitive tasks. In particular, Lumosity provides a web-based platform for simple games which are intended to be used for training various aspects of cognition. Recently, Lumosity has made the data collected by individuals playing these games available to researchers. This data set is the largest existing set of cognitive performance data, consisting of over 35 million users playing 600 million game trials. Researchers hope to use this data to obtain new insights into human cognition and how people think. Initial analyses reveal that aging and lifestyle factors such as amount of sleep taken and alcohol consumed significantly correlate with task performance (Sternberg et al. 2013).

(continued)

The final example here is the Foldit GWAP. Foldit is a collaborative online multiplayer protein-folding game (Cooper et al. 2010). Foldit players generate complex protein folds through directly manipulating protein structures and working with automated protein structure prediction technologies. Protein folding is an extremely complex problem, with even simple proteins having over 1000 degrees of freedom. It is also a problem that can have potentially high real-world payoff in the understanding of disease and in the development of novel drugs. One of the key successes of Foldit was in identifying the crystal structure of the retroviral protease (a key component of a retrovirus that influences viral maturation and proliferation) of the Mason-Pfizer monkey virus, which causes AIDS in monkeys. This structure had previously proved to be elusive to uncover through standard methods, and it is hoped that understanding this structure will lead to new antiviral treatments (Khatib et al. 2011).

Challenges in Developing GWAP for BCI

However, there is one primary challenge that must be overcome to develop a BCI-based GWAP. This is the challenge of developing an entertaining and engaging game such that large numbers of people will choose to play the game and use the BCI paradigms within the game in the face of the fact that the technical capability for robust and reliable BCIs is currently lacking. As a result, any BCI-based GWAP run the risk of either overemphasizing the BCI paradigms and being too frustrating and unpleasant to play or overemphasizing the gameplay and not providing viable data. Developing viable GWAP that incorporate BCI will require carefully balancing three key factors: the game must be *engaging* such that people will play for significant periods of time; BCI paradigms must be seamlessly *integrated* with the game while keeping gameplay entertaining and ensuring that sufficient experimental control is maintained; and sufficient information must be recorded to *contextualize* the resulting data so that it can be usefully interpreted. A hypothetical game that attempts to address all three of these challenges is shown in Fig. 8.

Developing Engaging Gameplay

In order to obtain the amount of data required to successfully address the issue of robustness and reliability in BCI, we will require the truly large-scale and long-term usage that only comes from intrinsic motivation, i.e., subjects play the game because they enjoy it. While there are an increasing number of BCI games available, none of them have obtained a large and dedicated audience sufficient to provide the needed large-scale and long-term data. One of the key reasons for

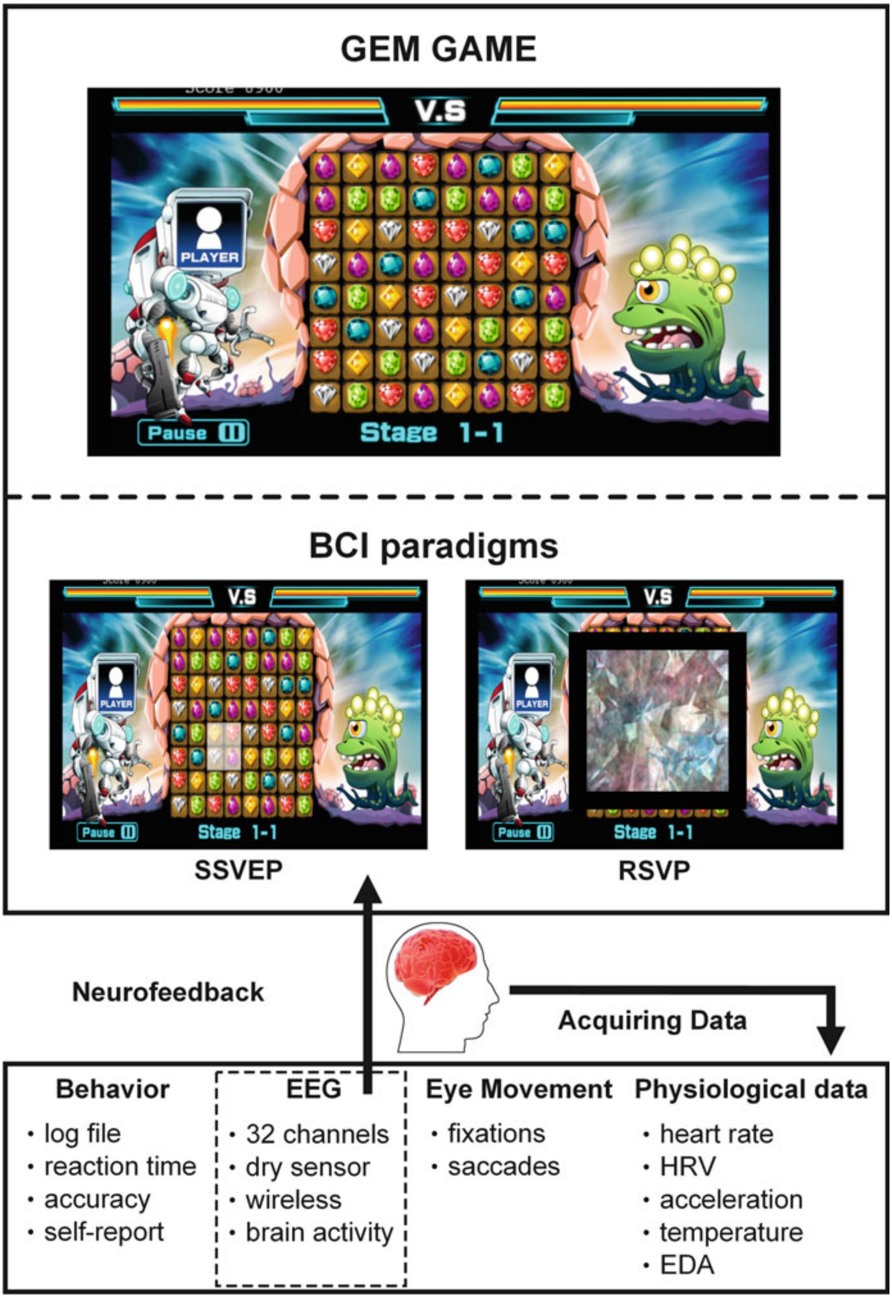

Fig. 8 A hypothetical game with a purpose for collecting long-term and large-scale BCI data. This hypothetical game is based upon an existing gameplay model which is known to be engaging, that of an opposed gem-matching game, similar to the Puzzle Quest series. The game includes multiple BCI paradigms that are integrated both mechanically and artistically with the game and records multimodal physiological and behavioral data for contextualization

this is that BCI games often focus on the BCI usage as the central aspect of the experience (Marshall et al. 2013).

Many of these games are focused on directly controlling the movement of objects on the screen through neural signals, either through the player directly imagining the movement of the object or through modulation of the spectral power within the alpha frequency band by attempting to feel more “relaxed” or “stressed.” As the BCI paradigm is usually the central focus of the game, further aspects of gameplay and game design tend to be very limited. While these games can initially be highly entertaining, the novelty of the BCI wears off in a matter of minutes. Once this occurs, it becomes clear that BCI technologies currently lack the robustness and reliability required to be a primary human interface device, leading to gameplay which is frustrating and unenjoyable (Dakan 2010). In fact, if the user gets frustrated by a BCI with subpar functional performance, this frustration itself affects their neural signals and can further degrade functionality.

For BCI-based GWAP to be successful, the focus of the design should be on ensuring that the game is sufficiently engaging and enjoyable that players will be intrinsically motivated to play the game. Understanding the user experience is a crucial aspect of game design, and previous research shows a key relationship between engagement, user experiences, and increased play time (Hassenzahl and Tractinsky 2006). We argue that focusing the game design on engagement, entertainment, and immersion first gains several advantages. It will make it possible to create a game that can be fun for hours, making the long-term data collection much less painful on the subjects and thereby decreasing the likelihood of subjects dropping out of the experiment. It will also be much more likely to obtain a large player base that will willingly play for free, enabling large-scale data collection. One way to ease the development of a sufficiently engaging game is to utilize an existing model of gameplay.

Integrating BCIs with Gameplay: The Free-To-Play (F2P) Model

There is a large body of previous research focused on increasing engagement and immersion during gameplay, and there are many aspects of designing a game that is engaging or even addictive. One example of the criticality of engaging experiences is the development of free-to-play (F2P) games. F2P games are distributed online for free and make money by providing in-game items, areas, or capabilities which can be purchased. As a result, the game only makes money by drawing large numbers of people in and getting them engaged over the long term. There are lessons to be taken from the design of these F2P games on how to embed BCI usage within a game in such a way that both BCI usage and long-term gameplay are encouraged (Luban 2011; Luton 2013).

Under the F2P model, the use of the BCI can be viewed as the “price” that is paid to play the game. As a result, it should be treated similarly to purchasing upgrades in a standard F2P game. Thus, unlike the design of the core gameplay, which should emphasize engagement, the integration of the BCI paradigms should be focused on making the experience as painless and as rewarding for the player as possible. By this, we mean that the player should not have to go out of their way in order to activate the BCI paradigm and should receive a significant gameplay benefit for doing so. However, due to the robustness of current state-of-the-art BCI technologies, it is likely that the BCI paradigm will fail to work a significant percentage of the time. Therefore, the game must be designed such that it is possible for the player to circumvent the use of the BCI in the game, while still providing the required research data.

Consider how this could be done with the three BCI paradigms described in Box 1. For example, the RSVP paradigm is very similar to reward “slot machines” that are commonly used in F2P games. In these “slot machines,” a small number of highly valuable in-game rewards are interspersed with a large number of low-value rewards that are rapidly displayed to the player, similar to a wheel on a slot machine. The player attempts to obtain a high-value reward by selecting it, which is difficult to accomplish due to the speed of the display. By simply adding a behavioral selection to the RSVP task and providing an in-game reward for selecting a target, it is possible to mirror a successful mechanism that commonly appears in F2P games. However, while this addresses the issue of integrating the RSVP paradigm into the game, it may not provide sufficient experimental control, depending on the goals of the experiment. For example, if the goal is to study the effects of changes in states such as fatigue on BCI performance, then one key problem that can arise is that of performance saturation. If the task is too easy, then all players will always perform it with near-perfect accuracy. When this happens, the fluctuations in behavior that would indicate low-performing states will simply not appear in the data, making it impossible to address the intended question. One possibility is to mask the targets and nontargets to make them more difficult to distinguish. Figure 8 shows a faceted masking, which fits artistically with the overall theme of the game, being used for this purpose.

There are many potential ways in which an SSVEP paradigm could be implemented in a game. Several targets could be displayed on-screen simultaneously, all flickering at different frequencies. The player could then use the SSVEP paradigm to select a desired reward or a target for an attack. However, in order to minimize the potential for frustration, the player must have the capability to circumvent a potentially nonfunctioning BCI. For example, the player could be allowed to click on the desired target. Unfortunately, while this integrates the SSVEP paradigm into the game, it does not provide sufficient experimental control. Instead, what is likely to happen is that the player will rapidly click the desired target without attending to it for the 1–2 s window that most SSVEP algorithms require. As a result, the neural data recorded will be insufficient for later analysis. Alternatively, the SSVEP paradigm could be interacted with through a mechanic where the player receives an award based on how quickly they click on the target but only once the target stops flickering. This would incentivize the player to pay close attention to the flickering target in order to maximize reward and ensure that the desired neural data of the player performing the SSVEP task would still be recorded.

Integrating the third BCI paradigm in Box 1, a passive behavioral performance estimation paradigm, is perhaps the simplest of the three BCI paradigms. As a passive BCI paradigm, no additional interaction with the player is needed, although if a specific task is being studied then that task could be embedded. Instead, what is required is extensive logging of the player's interaction with the game, as well as metrics and heuristics for estimating the quality of the player's moves. This data can then be used to study the relationship between neural activity and player behavior.

It is important to emphasize that the specific issues of maintaining experimental control when integrating the BCI paradigms with gameplay will vary based on the overall research goals being addressed. In addition, many of these issues, such as performance saturation, are not unique to any single paradigm but are relevant across many potential paradigms.

Finally, in order to encourage long-term gameplay, successful games provide for the player to experience task-reward feedback loops with multiple avenues for incremental progression. Example avenues for progression include difficulty, reward, capabilities, game mechanics, virtual space, and narrative space. Each of these also presents potential avenues for BCI integration. For example, there are usually various levels or difficulties in a game, and players will generally complete challenges in sequentially increasing difficulty (Brown and Cairns 2004). BCI paradigms can similarly be scaled in terms of difficulty. For example, both SSVEP and RSVP can be made more difficult by increasing the number of targets and distracters. This can also be tied to narrative progression, with difficulty or types of targets in a BCI also connected to sequential narrative episodes.

Contextualizing Data

Data recording and analysis is a key issue for F2P games, GWAP, and BCI implementation. In F2P games, games are often launched in a minimally playable state. Data from the game is then continuously monitored to determine key pieces of information such as what aspects of the game keep players engaged, what causes players to quit, and what drives players to perform in-game purchases. The game is then continually refined and updated based on this information (Luton 2013). When considering GWAP, in addition to monitoring this information, the experimental questions at the center of the GWAP must also be answered. As a result, additional data relevant to the experimental questions must also be recorded and analyzed (von Ahn and Dabbish 2008). Similarly, the analysis of the EEG data used for current BCI technologies requires a large amount of contextualization (Zander and Jatzev 2012). In order to estimate what the recorded neural signals may mean, it is necessary to know as much as possible about what the player is doing and how they are doing it.

As such, EEG data is not the only type of data that would need to be recorded. One key type of data needed is the information from the game itself, which includes player actions, game events, the times these occur, and contextual information from the state of the game that will serve as a basis for interpreting the other streams of data. Alternative sources of physiological data can also be used to contextualize the

EEG data and game data as well. For example, eye tracking technology provides the gaze direction in monitor coordinates and can be used to identify what aspect of the game the player is focusing on at any given time. While this and other physiological recording methods such as heart rate or electrodermal activity generally require laboratory-grade hardware, it is becoming increasingly possible to approximate these methods using common digital cameras. For example, pupil diameter, which can serve as an approximation for physiological arousal (Partala and Surakka 2003), can be detected using OpenCV, an open-source computer vision library (Schwarz et al. 2012). Similarly, heart rate can be detected using a webcam (Poh et al. 2011).

In order to uncover the correlates of longer-term state change, it will also be necessary to record self-report data from game players. Many of the correlates of this longer-term variation, such as sleep history and stress, cannot be derived from behavioral or physiological data while the subject is playing the game. However, some of these can be approximated using self-report (Buysse et al. 1989; Brantley et al. 1987).

One key challenge for recording this multimodal data is that the synchronization across the data streams must be highly accurate. As the quality of the synchronization decreases, it becomes increasingly difficult to associate the recorded neural signals with any external stimuli or behavior. For example, the amplitude of an averaged ERP is rapidly attenuated as synchronization becomes increasingly noisy. While other data streams may not degrade as rapidly with poor synchronization as EEG does, it will still become extremely difficult to make sense of inaccurately synchronized data during analysis.

Usable EEG Technologies

In addition to the challenges of developing an engaging game and integrating BCI paradigms into that game, there are also clear engineering challenges that need to be met in order to develop BCI-based GWAP which can obtain a large player base. The primary challenge is in the development of electroencephalography (EEG) recording devices which are low-cost, easy to use, and provide high signal quality (McDowell et al. 2013; Liao et al. 2012). While a BCI-based GWAP may still be of value without these devices, it will be limited in focus to single-session and longitudinal experiments with small subject populations.

One key usability issue is that current EEG technologies require a conductive NaCl-based gel or saline solution between the electrodes and skull to decrease the impedance and improve signal-to-noise. These gels and solutions serve as a bottleneck to using the BCI applications by requiring a time-consuming pregame application and postgame cleanup. Several research groups are developing dry EEG sensor technologies that would eliminate the need for this conductive gel while maintaining the reliable measurement of EEG signals (Sullivan et al. 2008; Grozea et al. 2011; Liao et al. 2014).

Likewise, traditional EEG systems transmit the measured brain activity through a cable which is connected between the EEG cap and computer. This greatly limits usability, and the swaying of the wires when the player moves their head causes additional artifact in the EEG signal. Wireless transmission using a protocol such as Bluetooth or Wi-Fi can overcome this limitation (Liao et al. 2014) and is easily extendable to tablet and mobile devices. While greatly adding to the usability of the EEG technology, using these wireless protocols may require additional power, raise bandwidth and packet loss issues, and increase the difficulty of properly synchronizing the EEG data.

While there are existing examples of commercially available dry-electrode, wireless EEG technologies (Sullivan et al. 2008), these systems tend to be too uncomfortable for long-term wear and to lack sufficient sensor density over a broad enough scalp distribution to examine different brain areas associated with a variety of cognitive functions. However, distinct types of BCI applications may relate to these cognitive functions and require collecting data from several different brain areas. For example, a motor imagery paradigm, where a BCI user attempts to control an object by imagining either its movement or their own, is usually based on the brain activity recorded from the motor cortex (Billinger et al. 2013). On the other hand, the data acquired from motor areas may be not fully appropriate for either SSVEP or RSVP paradigms. Instead, for these two types of BCI applications, the activity in the visual areas is usually measured and analyzed (Lin et al. 2014). For most research applications, the trend is to record the brain activity from the entire brain. While this may not be needed for BCI applications, it is unlikely that only one or two electrodes would be sufficient, although it is an open question as to how many electrodes are needed.

Conclusion

Performance issues in BCI technologies arise in a large part from human variability, including both individual differences and intraindividual variability, although there are other causes. In fact, in nearly every area of BCI, both individual differences and intraindividual variability have strong effects on the BCI users' neural data. This is not unique to BCI. There are existing approaches to deal with individual differences and intraindividual variability which may drastically improve the performance of BCI technologies by detecting and adapting to the individual user and their current mental state. However, using these approaches will require large amounts of data from people using BCI technologies over long periods of time.

We propose that one method for obtaining this data is through the development of BCI-based games with a purpose (GWAP) that not only are focused on obtaining the required BCI data but are also engaging and entertaining enough that large numbers of users will willingly play them. As a result, these BCI games which are expressly designed for entertainment and engagement have the potential to revolutionize research into a broad range of serious BCI applications.

There are critical challenges that must be overcome to develop BCI-based GWAP. The most pressing challenge is that of developing engaging BCI-based games when BCI technologies have such low levels of performance. As a result, developing these GWAP will require carefully balancing three key factors: developing an engaging and entertaining game, seamlessly integrating BCI paradigms with the game, and properly contextualizing the resulting data to enable useful interpretation and analysis. However, in order for data from a large number of users to be obtained, EEG systems that are comfortable, easy-to-use, low-cost, and that provide high-quality data will be required. A BCI-based GWAP would still be of value without this EEG technology, as currently available EEG systems could be used to obtain data from fewer participants over long periods of time. As a result, whether these EEG technologies come in the near future or are longer-term technologies, the future development of serious BCI applications for nonmedical applications would benefit greatly from large-scale and long-term collection of data from individuals performing BCI-relevant tasks in GWAP.

Recommended Reading

V. Balasubramanian, K. Adalarasu, A. Gupta, EEG based analysis of cognitive fatigue during simulated driving. Int. J. Ind. Syst. Eng. **7**(2), 135–149 (2011). doi:10.1504/IJISE.2011.038563

M. Billinger, C. Brunner, G.R. Müller-Putz, Single-trial connectivity estimation for classification of motor imagery data. J. Neural Eng. **10**(4), 046006 (2013). doi:10.1088/1741-2560/10/4/046006

G. Bin, X. Gao, Y. Zheng, B. Hong, S. Gao, An online multi-channel SSVEP-based brain–computer interface using a canonical correlation analysis method. J. Neural Eng. **6**(4), 046002 (2009). doi:10.1088/1741-2560/6/4/046002

D.P.-O. Bos, B. Reuderink, B. van de Laar, H. Gürkök, C. Mühl, M. Poel, A. Nijholt, D. Heylen, Brain-computer interfacing and games, in *Brain-Computer Interfaces*, ed. by D.S. Tan, A. Nijholt. Human-computer interaction series (Springer, London, 2010), pp. 149–178. doi:10.1007/978-1-84996-272-8_10

P.J. Brantley, C.D. Waggoner, G.N. Jones, N.B. Rappaport, A daily stress inventory: development, reliability, and validity. J. Behav. Med. **10**(1), 61–73 (1987)

E. Brown, P. Cairns, A grounded investigation of game immersion, in *CHI'04 Extended Abstracts on Human Factors in Computing Systems (CHI EA'04)* (ACM, New York, 2004), pp. 1297–1300. doi:10.1145/985921.986048

D.J. Buysse, C.F. Reynolds III, T.H. Monk, S.R. Berman, D.J. Kupfer, The Pittsburgh sleep quality index: a new instrument for psychiatric practice and research. Psychiatry Res. **28**(2), 193–213 (1989)

S. Cooper, F. Khatib, A. Treuille, J. Barbero, J. Lee, M. Beenen, A. Leaver-Fay, D. Baker, Z. Popović, F. Players, Predicting protein structures with a multiplayer online game. Nature **466**(7307), 756–760 (2010). doi:10.1038/nature09304

R. Dakan, Review: Emotiv EPOC, tough thoughts on the new mind-reading controller. Joystiq (2010). 27 Jan. http://www.joystiq.com/2010/01/27/review-emotiv-epoc-tough-thoughts-on-the-new-mind-reading-cont/

C.D. Geisler, L.S. Frishkopf, W.A. Rosenblith, Extracranial responses to acoustic clicks in man. Science **128**(3333), 1210–1211 (1958). doi:10.1126/science.128.3333.1210

A.S. Gevins, G.M. Zeitlin, S. Ancoli, C.L. Yeager, Computer rejection of EEG artifact. II. Contamination by drowsiness. Electroencephalogr. Clin. Neurophysiol. **43**(1), 31–42 (1977)

C. Grozea, C.D. Voinescu, S. Fazli, Bristle-sensors – low-cost flexible passive dry EEG electrodes for neurofeedback and BCI applications. J. Neural Eng. **8**(2), 025008 (2011)

M. Hassenzahl, N. Tractinsky, User experience – a research agenda. Behav. Inform. Technol. **25**(2), 91–97 (2006). doi:10.1080/01449290500330331

L.R. Hochberg, M.D. Serruya, G.M. Friehs, J.A. Mukand, M. Saleh, A.H. Caplan, A. Branner, D. Chen, R.D. Penn, J.P. Donoghue, Neuronal ensemble control of prosthetic devices by a human with tetraplegia. Nature **442**(7099), 164–171 (2006). doi:10.1038/nature04970

D.C. Jangraw, J. Wang, B.J. Lance, S.-F. Chang, P. Sajda, Neurally and ocularly informed graph-based models for searching 3D environments. J. Neural Eng. **11**(4), 046003 (2014). doi:10.1088/1741-2560/11/4/046003

T.-P. Jung, S. Makeig, M. Stensmo, T.J. Sejnowski, Estimating alertness from the EEG power spectrum. IEEE Trans. Biomed. Eng. **44**(1), 60–69 (1997). doi:10.1109/10.553713

F. Khatib, F. DiMaio, S. Cooper, M. Kazmierczyk, M. Gilski, S. Krzywda, H. Zabranska et al., Crystal structure of a monomeric retroviral protease solved by protein folding game players. Nat. Struct. Mol. Biol. **18**(10), 1175–1177 (2011)

J.H. Kim, D.V. Gunn, E. Schuh, B. Phillips, R.J. Pagulayan, D. Wixon, Tracking real-time user experience (TRUE): a comprehensive instrumentation solution for complex systems, in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (ACM, 2008), pp. 443–452. http://dl.acm.org/citation.cfm?id=1357126

D.J. Krusienski, E.W. Sellers, F. Cabestaing, S. Bayoudh, D.J. McFarland, T.M. Vaughan, J.R. Wolpaw, A comparison of classification techniques for the P300 speller. J. Neural Eng. **3**, 299–305 (2006). doi:10.1088/1741-2560/3/4/007

B.J. Lance, S.E. Kerick, A.J. Ries, K.S. Oie, K. McDowell, Brain-computer interface technologies in the coming decades. Proc. IEEE **100**(Special Centennial Issue), 1585–1599 (2012). doi:10.1109/JPROC.2012.2184830

L.-D. Liao, C.-T. Lin, K. McDowell, A.E. Wickenden, K. Gramann, T.-P. Jung, L.-W. Ko, J.-Y. Chang, Biosensor technologies for augmented brain–computer interfaces in the next decades. Proc. IEEE **100**(Special Centennial Issue), 1553–1566 (2012). doi:10.1109/JPROC.2012.2184829

L.-D. Liao, S.-L. Wu, C.-H. Liou, S.-W. Lu, S.-A. Chen, S.-F. Chen, L.-W. Ko, C.-T. Lin, A novel 16-channel wireless system for electroencephalography measurements with dry spring-loaded sensors. IEEE Trans. Instrum. Meas. **63**(6), 1545–1555 (2014). doi:10.1109/TIM.2013.2293222

C.-T. Lin, R.-C. Wu, S.-F. Liang, W.-H. Chao, Y.-J. Chen, T.-P. Jung, EEG-based drowsiness estimation for safety driving using independent component analysis. IEEE Trans. Circuits Syst. Regul. Pap. **52**(12), 2726–2738 (2005). doi:10.1109/TCSI.2005.857555

Y.-P. Lin, Y. Wang, T.-P. Jung, Assessing the feasibility of online SSVEP decoding in human walking using a consumer EEG headset. J. Neuroeng. Rehabil. **11**(1), 119 (2014). doi:10.1186/1743-0003-11-119

B. Lou, Y. Li, M.G. Philiastides, P. Sajda, Prestimulus alpha power predicts fidelity of sensory encoding in perceptual decision making. Neuroimage **87**, 242–251 (2014). doi:10.1016/j.neuroimage.2013.10.041

P. Luban. The design of free-to-play games. Gamasutra (2011). 22 Nov. http://www.gamasutra.com/view/feature/134920/the_design_of_freetoplay_games_.php

W. Luton, *Free-to-Play: Making Money From Games You Give Away*, 1st edn. (New Riders, 2013)

D. Marshall, D. Coyle, S. Wilson, M. Callaghan, Games, gameplay, and BCI: the state of the art. IEEE Trans. Comput. Intell. AI Games **5**(2), 82–99 (2013). doi:10.1109/TCIAIG.2013.2263555

G. Matthews, P.A. Desmond, Task-induced fatigue states and simulated driving performance. Q. J. Exp. Psychol. A Hum. Exp. Psychol. **55**(2), 659–686 (2002). doi:10.1080/02724980143000505

K. McDowell, C.-T. Lin, K.S. Oie, J. Tzyy-Ping, S. Gordon, K.W. Whitaker, S.-Y. Li, S.-W. Lu, W.D. Hairston, Real-world neuroimaging technologies. IEEE Access **1**, 131–149 (2013). doi:10.1109/ACCESS.2013.2260791

D.J. McFarland, W.A. Sarnacki, J.R. Wolpaw, Electroencephalographic (EEG) control of three-dimensional movement. J. Neural Eng. **7**(3), 036007 (2010). doi:10.1088/1741-2560/7/3/036007

S.R. Mitroff, A.T. Biggs, Mitroff, Biggs, The ultra-rare-item effect visual search for exceedingly rare items is highly susceptible to error. Psychol. Sci. (2013). doi:10.1177/0956797613504221

C. Mühl, C. Jeunet, F. Lotte, EEG-based workload estimation across affective contexts. Neuroprosthetics **8**, 114 (2014). doi:10.3389/fnins.2014.00114

S.J. Pan, Q. Yang, A survey on transfer learning. IEEE Trans. Knowl. Data Eng. **22**(10), 1345–1359 (2010)

T. Partala, V. Surakka, Pupil size variation as an indication of affective processing. Int. J. Hum. Comput. Stud. **59**(1), 185–198 (2003)

G. Pfurtscheller, B.Z. Allison, C. Brunner, G. Bauernfeind, T. Solis-Escalante, R. Scherer, T.O. Zander, G. Mueller-Putz, C. Neuper, N. Birbaumer, The hybrid BCI. Front. Neurosci. **4** (2010). doi:10.3389/fnpro.2010.00003

M.-Z. Poh, D.J. McDuff, R.W. Picard, Advancements in noncontact, multiparameter physiological measurements using a webcam. IEEE Trans. Biomed. Eng. **58**(1), 7–11 (2011)

E.A. Pohlmeyer, J. Wang, D.C. Jangraw, B. Lou, S.-F. Chang, P. Sajda, Closing the loop in cortically-coupled computer vision: a brain–computer interface for searching image databases. J. Neural Eng. **8**(3), 036025 (2011)

P. Sajda, E. Pohlmeyer, J. Wang, L.C. Parra, C. Christoforou, J. Dmochowski, B. Hanna, C. Bahlmann, M.K. Singh, S.-F. Chang, In a blink of an Eye and a switch of a transistor: cortically coupled computer vision. Proc. IEEE **98**(3), 462–478 (2010). doi:10.1109/JPROC.2009.2038406

L. Schwarz, Humberto Remigio Gamba, Fabio Cabral Pacheco, Rodrigo Belisario Ramos, Miguel Antonio Sovierzoski, Pupil and iris detection in dynamic pupillometry using the OpenCV library, in *2012 5th International Congress on Image and Signal Processing* (*CISP*) (2012), pp. 211–215. doi: 10.1109/CISP.2012.6469846

B. Settles, in *Active Learning Literature Survey,* vol. 52 (University of Wisconsin, Madison, 2010), pp. 55–66

D.A. Sternberg, K. Ballard, J.L. Hardy, K. Benjamin, P.M. Doraiswamy, M. Scanlon, The largest human cognitive performance dataset reveals insights into the effects of lifestyle factors and aging. Front. Hum. Neurosci. **7**, 292 (2013). doi:10.3389/fnhum.2013.00292

T.J. Sullivan, S.R. Deiss, T.-P. Jung, G. Cauwenberghs, A brain-machine interface using dry-contact, low-noise EEG sensors, in *Circuits and Systems, 2008. ISCAS 2008. IEEE International Symposium on* (IEEE, 2008), pp. 1986–1989. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4541835

C. Thompson. Halo 3: how Microsoft labs invented a new science of play. http://archive.wired.com/gaming/virtualworlds/magazine/15-09/ff_halo?currentPage=all Mag. **15**(9) (2007). http://josquin.cti.depaul.edu/~rburke/courses/f07/gam224/doc/science_of_play.pdf

J. Touryan, G. Apker, B.J. Lance, S.E. Kerick, A.J. Ries, K. McDowell, Estimating endogenous changes in task performance from EEG. Neuroprosthetics **8**, 155 (2014). doi:10.3389/fnins.2014.00155

J. Van Erp, F. Lotte, M. Tangermann, Brain-computer interfaces: beyond medical applications. Computer **45**(4), 26–34 (2012). doi:10.1109/MC.2012.107

G. Vanacker, R. del José, E.L. Millán, P.W. Ferrez, F.G. Moles, J. Philips, H. Van Brussel, M. Nuttin, Context-based filtering for assisted brain-actuated wheelchair driving. Comput. Intell. Neurosci. **2007**, 3 (2007). doi:10.1155/2007/25130

L. Von Ahn, Games with a purpose. Computer **39**(6), 92–94 (2006). doi:10.1109/MC.2006.196

L. Von Ahn, L. Dabbish, Designing games with a purpose. Commun. ACM **51**(8), 58–67 (2008). doi:10.1145/1378704.1378719

J.R. Wolpaw, N. Birbaumer, D.J. McFarland, G. Pfurtscheller, T.M. Vaughan, Brain–computer interfaces for communication and control. Clin. Neurophysiol. **113**(6), 767–791 (2002). doi:10.1016/S1388-2457(02)00057-3

D.D. Wu, C.G. Courtney, B.J. Lance, S.S. Narayanan, M.E. Dawson, K.S. Oie, T.D. Parsons, Optimal arousal identification and classification for affective computing using physiological signals: virtual reality stroop task. IEEE Trans. Affect. Comput. **1**(2), 109–118 (2010). doi:10.1109/T-AFFC.2010.12

T.O. Zander, S. Jatzev, Context-aware brain–computer interfaces: exploring the information space of user, technical system and environment. J. Neural Eng. **9**(1), 016003 (2012). doi:10.1088/1741-2560/9/1/016003

T.O. Zander, C. Kothe, Towards passive brain–computer interfaces: applying brain–computer interface technology to human–machine systems in general. J. Neural Eng. **8**(2), 025005 (2011). doi:10.1088/1741-2560/8/2/025005

1 DEFENSE TECHNICAL
(PDF) INFORMATION CTR
DTIC OCA

2 DIRECTOR
(PDF) US ARMY RESEARCH LAB
RDRL CIO L
IMAL HRA MAIL & RECORDS
MGMT

1 GOVT PRINTG OFC
(PDF) A MALHOTRA

1 DIR USARL
(PDF) RDRL HRF C
B LANCE

INTENTIONALLY LEFT BLANK.