International Journal of Human-Computer Interaction

ISSN: 1044-7318 (Print) 1532-7590 (Online) Journal homepage: https://www.tandfonline.com/loi/hihc20

Brain–Computer Interfaces for Multimodal Interaction: A Survey and Principles

Hayrettin Gürkök & Anton Nijholt

To cite this article: Hayrettin Gürkök & Anton Nijholt (2012) Brain–Computer Interfaces for Multimodal Interaction: A Survey and Principles, International Journal of Human-Computer Interaction, 28:5, 292-307, DOI: 10.1080/10447318.2011.582022

To link to this article: https://doi.org/10.1080/10447318.2011.582022

Published online: 09 Apr 2012.

Intl. Journal of Human–Computer Interaction, 28: 292–307, 2012

ISSN: 1044-7318 print / 1532-7590 online
DOI: 10.1080/10447318.2011.582022

Brain–Computer Interfaces for Multimodal Interaction: A Survey and Principles

Hayrettin Gürkök and Anton Nijholt
University of Twente, Enschede, The Netherlands

For decades, brain–computer interfaces (BCIs) have been used for restoring the communication and mobility of disabled people through applications such as spellers, web browsers, and wheelchair controls. In parallel to advances in computational intelligence and the production of consumer BCI products, BCIs have recently started to be considered as alternative modalities in human–computer interaction (HCI). One of the popular topics in HCI is multimodal interaction (MMI), which deals with combining multiple modalities in order to provide powerful, flexible, adaptable, and natural interfaces. This article discusses the situation of BCI as a modality within MMI research. State-of-the-art, real-time multimodal BCI applications are surveyed in order to demonstrate how BCI can be helpful as a modality in MMI. It is shown that multimodal use of BCIs can improve error handling, task performance, and user experience and that they can broaden the user spectrum. The techniques for employing BCI in MMI are described, and the experimental and technical challenges with some guidelines to overcome these are shown. Issues in input fusion, output fission, integration architectures, and data collection are covered.

1. INTRODUCTION

The intelligent computers of today are able to perceive their environment using sensor technologies and to respond with the help of advanced decision-making algorithms. They welcome us into an elevator or a photo booth, and they accompany us in our pockets or on our clothes. Considering the amount of interaction we enter into with these pervasive machines, we need natural, intuitive user interfaces that understand or anticipate our intentions and react to make our lives easier. Thus, we should be able to interact with computers in the same way that we do with humans. In other words, human–computer interaction (HCI) should carry the characteristics of human–human interaction (HHI).

HHI relies on concurrent use and perception of behavioral signals (cues) such as speaking, moving, gazing, and gesturing, which convey various messages (communicative intentions). We show our approval by a thumbs-up, perhaps accompanied by speech, a wink, or a nod. To describe an object, we talk about it, at the same time moving our hands to explain its different features such as its size or shape. While we are sending our signals, our conversation partner receives them through his multiple senses; he listens to us and watches our gestures. For a humanlike interaction, the interfaces of the modern HCI offer multiple sensing (input) and response (output) modalities for interacting with computers. Within this interaction style, called multimodal interaction (MMI), computers hear us via the microphone, see through the camera, and even feel through haptic devices. In return, they give feedback in the form of an embodied conversational agent (Cassell, Sullivan, Prevost, & Churchill, 2000) or through a tactile device.

Although computers can mimic some human senses, there are situations in which they need to possess better sensing abilities than humans. There are times that we, consciously or not, conceal our mental or emotional states. Some people are just not comfortable with expressing themselves overtly, or they deliberately suppress their behavioral cues as in the case of bluffing. Moreover, in the absence of a human conversation partner, the cues may become subtle or may even vanish. In expressing our intentions, we are also not always explicit. This is perhaps because we are so tired that we do not want to move, or our hands are occupied so that we cannot use them, or we are physically disabled. Still, we expect computers to understand our implicit emotions, difficulties, and intentions.

Computers cannot read our minds, but brain-computer interfaces (BCIs) can infer our mental/emotional states and intentions by interpreting our brain signals. The classical application domains of BCI are the restoration of the mobility and communication of the disabled individuals through a brain-controlled wheelchair (Leeb et al., 2007) or a spelling device (Sellers & Donchin, 2006), and the rehabilitation of

We gratefully acknowledge the support of the BrainGain Smart Mix Programme of the Netherlands Ministry of Economic Affairs and the Netherlands Ministry of Education, Culture and Science. We thank the two anonymous reviewers whose comments helped greatly to improve the paper and L. Packwood for giving the paper its final shape by proofreading it.

Address correspondence to Hayrettin Gürkök, Human Media Interaction, University of Twente, P.O. Box 217, 7500 AE, Enschede, The Netherlands. E-mail: h.gurkok@utwente.nl

people with disorders such as attention deficit hyperactivity disorder (Gevensleben et al., 2009). With the emerging portable and usable signal acquisition hardware as well as robust data processing and artifact removal techniques, BCI has started to be considered as an HCI modality for nondisabled users as well. Some nonmedical BCI applications include games (Plass-Oude Bos et al., 2010), attention monitors (Jackson & Mappus, 2010) and tools for mobile devices (Campbell et al., 2010; Wang, Wang, & Jung, 2011).

Although there are a substantial number of BCI applications for nondisabled users, the majority of them are unimodal (i.e., BCI is the only modality). As we later report in this article, there are also a small number of multimodal BCI applications as well, but these were developed mostly for research purposes and assessed in controlled experiments. To take part in an HHI-like HCI and be used at home by regular users, BCI needs to be employed in MMI. Therefore, many prototype multimodal BCI applications should be developed to evaluate the joint advantages of BCI and MMI, and guidelines should be proposed to the challenges encountered.

The purpose of this article is twofold. First, we survey the state-of-the-art multimodal BCI applications to demonstrate how BCI can be helpful as a modality in MMI. Second, we describe the techniques for employing BCI in MMI and show the experimental and technical challenges with some guidelines to overcome these.

The article is organized as follows. In section 2, we summarise the fundamental principles of MMI and BCI. First we define MMI and its components. Then we give an overview of brain activity measurement and interpretation methods followed by BCI interaction paradigms. Section 3 provides a range of multimodal BCI applications grouped according to their major benefits and illustrates their possible real-life uses. In section 4, we describe the techniques for using BCI in MMI and discuss the challenges involved while providing some guidelines to cope with these challenges. We cover the issues of input fusion, output fission, integration architectures, and data collection. Finally, in section 5, we conclude by stressing the essential aspects discussed throughout the article and point to future research directions.

2. BACKGROUND

Before going into any discussion about the employment of MMI and BCI together, we find it useful to give an overview of each separately. We first define MMI and provide the MMI framework. Next we describe the principles of BCI and give examples of the latest applications.

2.1. Definition and Mechanism of Multimodal Interaction

A multimodal user interface (MUI) could simply be regarded as an interface that allows its users to provide input through two or more different modalities. However, we understand “modality” not only as a means of input but as an output pathway as well. For better understanding and structuring of this concept, here we adopt the definition from Oviatt (2008, p. 414): “Multimodal systems process two or more combined user input modes -such as speech, pen, touch, manual gestures, gaze, and head and body movements- in a coordinated manner with multimedia system output.” Thus, in MMI one major task is to perceive the human input (which is nonconventional, thus other than keyboard and mouse), whereas the other one is to output reaction accordingly.

In congruence with this definition, W3C identifies six components of an MMI framework (see Figure 1; Larson, Raman, & Raggett, 2003). The input and output components can be

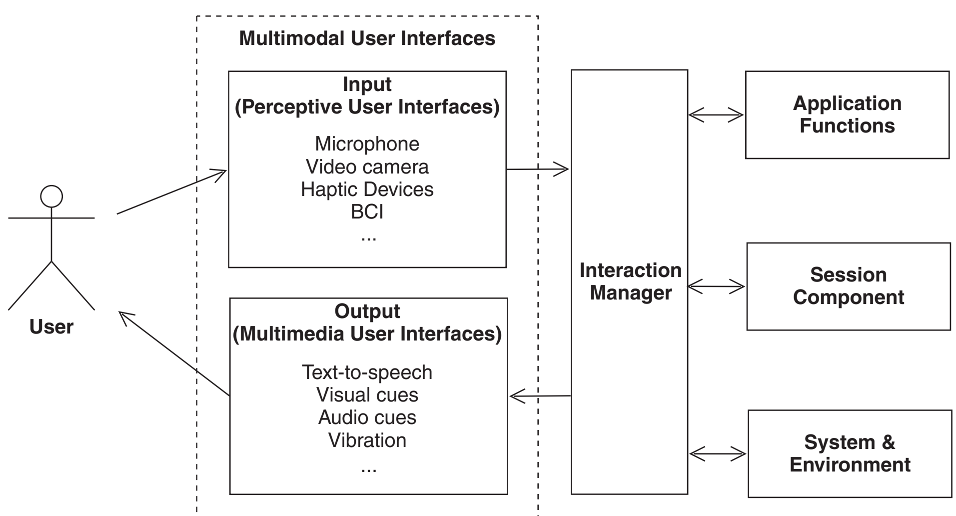

FIG. 1. Multimodal interaction framework based on the original model of W3C (Larson et al., 2003). *Note.* The user provides input using perceptive user interfaces while receiving output through multimedia user interfaces. The communication between the perceptive and multimedia user interfaces is facilitated by the interaction manager in cooperation with application functions, the session component, and the system and environment component.

viewed as perceptive user interfaces and multimedia user interfaces, respectively, as suggested by Turk and Robertson (2000), although they situate these interfaces and MUIs differently under the concept of perceptual user interfaces. *Perceptive user interfaces* are those that add perceptual (i.e., sensing) capabilities to the computer. This can be achieved through a combination of techniques such as speech, gesture or handwriting recognition. *Multimedia user interfaces*, on the other hand, provide feedback to the user through different media formats such as audio, animation, automated speech, and text. The communication between these two interfaces is maintained by the interaction manager in cooperation with application functions, the session component, and the system and environment component. Within this framework, BCI is a perceptive user interface and therefore belongs to the input component.

In this section, we limit our discussion on MMI to the basic mechanism. We provide MMI examples for various scenarios with and without BCIs in section 3. In section 4 we describe the principles of MMI, namely, input fusion, output fission, synchronization, and data collection in general while showing the challenges they pose to BCI.

2.2. Principles of Brain–Computer Interfaces

In this article we represent a BCI as a system with three procedural components that outputs supporting actions according to human intention or mental/emotional state derived through brain activity (see Figure 2). First, a user's brain activity is acquired and quantified as a signal (see the next section). Then, based on the neuromechanisms stemming from the neurological functioning of the brain, the signal is interpreted to obtain knowledge on user state or intention (see the Interpreting Brain Activity (Neuromechanisms) section). Finally, this knowledge is employed in an application to satisfy the user's need (see the Interacting via Brain Activity section). Next, we briefly describe each component. For an extensive overview of BCIs we refer the reader to the review by Wolpaw, Birbaumer, McFarland, Pfurtscheller, and Vaughan (2002).

Acquiring (measuring) brain activity (Imaging Modalities). The first experiments on acquiring (measuring) human brain activity date back to the 1920s. Berger (1929) was the first to publish the results of electroencephalography (EEG) experiments on humans (translated version available by Gloor, 1969). EEG is a technique for acquiring the electrical activity of the brain from the scalp by use of electrodes. Since Berger's first experiments, not only have EEG recordings become prevalent but other acquisition techniques relying on electrical, magnetic, and hemodynamic (blood movement) response of the brain have also emerged.

Brain activity acquisition methods (also known as imaging modalities in neuroscience, but we do not refer to them as modalities so as not to cause any conflict with the HCI definition of modality) can be categorized according to the manner of deployment as being invasive or noninvasive (see Table 1). Invasive methods are implemented either by placing electrodes on the surface of the cortex (electrocorticography, or ECoG) or by implanting them inside the cortex (multielectrode array, or MEA). These methods provide high dimensionality and signal quality. However, they require surgery for deployment and extreme care for stability and against possible infections. Therefore, they are applied only on people with severe disabilities for whom these are the only ways of redressing the disability. Noninvasive methods measure the activity from the scalp, so do not carry the same risks as invasive methods. Thus, they are used more frequently in human research. Among noninvasive methods, magnetoencephalography (MEG) and functional magnetic resonance imaging (fMRI) are immobile machines and require good shielding from the environment so they are bound to controlled laboratory environments. On the other hand, EEG and near-infrared spectroscopy (NIRS) are portable, easily deployable, and relatively inexpensive devices. Their wireless implementations are also feasible, making them even more convenient to use. Therefore, they are more suitable for HCI research. Despite many advantages, the noninvasive methods are prone to artifacts from intense movements of the body, and the measurement quality is inferior in comparison to the invasive methods.

EEG, ECoG, MEA, and MEG measure the activity of the fast dendritic currents in a large population of brain cells. Thus, the recordings of the measurements have low latency (i.e., high temporal resolution). fMRI and NIRS measure the blood oxygenation in the brain, which is a much slower correlate of the brain activity. Therefore they offer lower temporal resolution. Noninvasive methods provide lower spatial resolution in comparison to invasive ones due to spatial mixing of electrical activity generated by different cortical areas and passive conductance of these signals through brain tissue, bone, and skin (van Gerven et al., 2009). Among noninvasive methods, fMRI has relatively higher spatial resolution, as it can sample the activity of deep brain structures. For a detailed description of

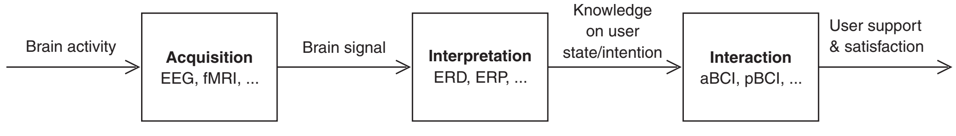

FIG. 2. Three-component brain–computer interface (BCI) model. *Note.* Brain activity is acquired and quantified as a signal. Then, the signal is interpreted to obtain knowledge on user state or intention. Finally, this knowledge is employed in an application to satisfy the user's need. Since the users' brain activity affects the BCI input as well as gets affected by the BCI output, the components of the BCI model actually form a loop.

TABLE 1
Properties of Brain Activity Measurement (Acquisition) Methods (Imaging Modalities)

	EEG	MEG	NIRS	fMRI	ECoG	MEA
Deployment	Noninvasive	Noninvasive	Noninvasive	Noninvasive	Invasive	Invasive
Measured activity	Electrical	Magnetic	Hemodynamic	Hemodynamic	Electrical	Electrical
Temporal resolution	Good	Good	Low	Low	High	High
Spatial resolution	Low	Low	Low	Good	Good	High
Portability	High	Low	High	Low	High	High
Cost	Low	High	Low	High	High	High

Note. EEG = electroencephalography; MEG = magnetoencephalography; NIRS = near infrared spectroscopy; fMRI = functional magnetic resonance imaging; ECoG = electrocorticography; MEA = microelectrode (or multielectrode) array.

the recording methods, the reader should see Kübler and Müller (2007), Lebedev and Nicolelis (2006), and van Gerven et al. (2009).

Interpreting brain activity (Neuromechanisms). Once the brain activity is acquired as a signal, the next step is to interpret its content. In doing this, we benefit from neuromechanisms, which signify certain changes in the signal with respect to an event. The event can be a voluntary action such as moving a hand or looking at something as well as an involuntary reaction to a stimulus or an error. In this section we briefly cover the most commonly employed neuromechanisms.

The brain maintains an ongoing (rhythmic) activity in absence of an external or internal intervention. These rhythms are identified by the frequency and brain location they occur at. Two closely related neuromechanisms, *event related desynchronisation* (ERD) and *event related synchronisation* (ERS; Pfurtscheller & Lopes da Silva, 1999; often referred to together as ERD/ERS) are the suppression and enhancement of the rhythmic brain activities respectively in relation to an event. By observing the signal amplitude in certain frequencies measured at specific parts of the brain, we can infer the underlying brain activity. As an example, the rolandic μ rhythm oscillates 9–13 Hz in the sensorimotor area. It is desynchronised during execution, preparation, or imagination of motor actions. So, by analyzing the amplitude of the signal recorded from the sensorimotor area between 9–13 Hz, it is possible to understand when a person executes or imagines executing a motor action, such as a hand, foot, or tongue movement (Pfurtscheller, Brunner, Schlögl, & Lopes da Silva, 2006). If in an application certain motor actions are matched to some commands, then one can control the application without any device or even actual movement. Scherer et al. (2007) used motor imagery to navigate in a virtual environment (VE) and execute certain commands in Google Earth. Another example is the alpha rhythm oscillating 8–13 Hz in the posterior region. It is blocked or attenuated by attention, especially visual, and mental effort so it has been associated with physical relaxation and relative mental inactivity (Deuschl & Eisen, 1999). Plass-Oude Bos et al. (2010) used parietal alpha power in the game World of Warcraft to switch the player avatar between an elf and a bear according to the player's relaxedness.

Another family of neuromechanisms is the *event-related potentials* (ERPs). These are called event related fields in the case of magnetic activity measurement, but for brevity we refer to them as potentials from now on. ERPs are short amplitude deflections in the brain signal, time-locked to an event. That is, they are expected at a fixed positive or negative latency with respect to a particular event. Thus, by observing the amplitude at this fixed latency, we can infer a person's reaction or intention. Various ERPs have been employed in BCI applications; we introduce the most commonly used ones next. ERPs are identified by the triggering event, direction of deflection, observed location, and latency. For the purpose of this article, we emphasize only the triggering event for each ERP and describe example applications. We encourage the reader to refer to Luck (2005) and Fabiani, Gratton, and Federmeier (2007) for a complete overview. A commonly used potential of the brain, P300, occurs after being exposed to a task-relevant stimulus (Farwell & Donchin, 1988). This makes P300 suitable for detecting intention through selection tasks. Edlinger, Holzner, Groenegress, Guger, and Slater (2009) used P300 to select and control items in a virtual apartment, whereas Campbell et al. (2010) used it to select and dial contacts on a real mobile phone. Intentions can also be inferred through the *readiness potential* (RP, also known as the *Bereitschaftspotential*), which precedes voluntary motor movements (Shibasaki & Hallett, 2006). Krepki, Blankertz, Curio, and Müller (2007) used lateralized RPs to predict the actual or imaginary finger movements of users and translate them into commands in a Pac-Man game. Another widely exploited set of potentials are the *error potentials* (ErrPs), which are reactions of the brain to errors (Ferrez & del R. Millán, 2007). Förster et al. (2010) used ErrPs to train their hand gesture recognition system based on the errors occurring during interaction.

When we attend to a stimulus repeating with a certain frequency, the amplitude of the signal measured in the brain area processing the stimulation is enhanced at the frequency of the stimulation. This enhancement is known as the *steady-state* *evoked potential* and is another frequently used neuromechanism (Regan, 1977). By presenting multiple stimuli with distinct repetition frequencies, we can detect which of the stimuli a person was paying attention to. If each of these stimuli is associated with a message, then we can understand the person's intention. Martinez, Bakardjianm and Cichocki (2007) used four checkerboards, each flickering with a unique frequency and associated with a direction (up, down, left, and right) for navigating a car on the computer screen. As in this work, when the stimulation is a visual one, the resulting response is called a *steady-state visually evoked potential* (SSVEP) and is observed over the occipital (visual) cortex. In the literature there are also studies with auditory (Herdman et al., 2002) and vibratory (Muller-Putz, Scherer, Neuper, & Pfurtscheller, 2006) stimulation.

While using neuromechanisms, which are evoked through stimulation, attention should be paid to the features of the stimuli, the stimulation device, and the environment. The stimulation parameters might significantly affect not only the strength or the presence of the brain response but also the comfort and experience of the user. For SSVEP-based BCIs, Bieger and Molina (2010) wrote an excellent report on the influence of stimulation parameters (such as the environment; the stimulation device; and the flicker frequency, color, and shape of the stimulus) on recognition performance and user comfort. Also for P300 stimulation, effects of factors such as screen size (Li, Nam, Shadden & Johnson, 2011), and color of and distance between stimuli (Salvaris & Sepulveda, 2009) on recognition accuracy have been reported.

Apart from the aforementioned standard neuromechanisms, there are power changes at specific frequencies distributed across the scalp in correlation with emotions (Chanel, Kierkels, Soleymani, & Pun, 2009) and certain mental activities such as mental object rotation (Nikolaev & Anokhin, 1998) or problem solving (Fink et al., 2009). These correlates could, for instance, be used to detect a user's mental or emotional state for assisting the user. We would like to finally note that for some events there is more than one representative neuromechanism, such as ERD and RP signifying motor execution or imagery, so combined use of these can yield a better recognition capability (Fatourechi, Birch, & Ward, 2007).

Interacting via brain activity. Interpreting the brain activity based on neuromechanisms allows us to arrive at knowledge about a user's intention, mental processing, or emotional state. We differentiate between BCIs with respect to their ways of utilizing this knowledge in an application according to user's needs. In this work, we adopt the categorization of Zander, Kothe, Welke, and Roetting (2009) and identify three types of BCIs: active (aBCI), reactive (rBCI), and passive (pBCI) (see Table 2).

In aBCI, the user intends to interact with the BCI application and for this purpose directly generates brain activity. For instance, such brain activity could be generated by imagining movements to navigate in a VE. In rBCI, the user still intends to interact with the BCI application but the brain activity is generated indirectly, in reaction to external stimulation. The user voluntarily attends to a stimulus, but what causes the brain to react are the stimulus features, not the act of attending. For example in SSVEP navigation, where each direction is associated with a stimulus repeating at a unique frequency, the user looks at one of the stimuli to go in a certain direction. It is not the act of looking that generates the desired brain activity but the brain's reaction to the repetition frequency of the stimulus. In pBCI, the user's primary aim is not to interact with the BCI application, or possibly he does not have an aim at all. The BCI system watches the user passively in order to adapt the task or the environment for improving and enriching the HCI or the quality of life. This might be by monitoring the attention level, emotional state, or mental load of the user. pBCIs rely on brain signals generated during natural interaction of the user with his environment so they do not require any additional effort (such as attention to stimulation). Therefore, they can operate within aBCIs and rBCIs without demanding extra experimental requirements.

We would like to stress that the interaction methods are applicable to BCI applications but not to the BCI neuromechanisms, because a neuromechanism can be utilized in different ways in applications. For example a P300 speller (Serby, Yom-Tov, & Inbar, 2005) would be an rBCI because the

TABLE 2
Features and Application Domains of BCI Interaction Paradigms

Type of BCI	Interaction With BCI System	Generation of Brain Activity	Used for	Example Applications
Active	Intended	Consciously	Direct control	Motor imagery-based navigation
Reactive	Intended	In response to stimulation (i.e., unconsciously)	Direct control	SSVEP-based selection, P300 speller
Passive	Unintended	Through interaction	Supporting systems	User state detection, error handling via ErrPs

Note. BCI = brain–computer interface.

user interacts with the BCI system for spelling words. On the other hand, a P300 workload monitor (Allison & Polich, 2008) would be a pBCI as the user has a primary task to devote attention to other than responding to the workload monitor. Having made this distinction, we would like to draw reader's attention to yet another important detail. Depending on the context and the goal of the user, different interaction methods can be utilized to operate the very same BCI application. In the aforementioned BCI game, Alpha-World of Warcraft (Plass-Oude Bos et al., 2010), the player avatar changes between the elf and bear forms according to the relaxedness of the player. During the game players might intentionally try to regulate their relaxedness for better performance or they might simply enjoy seeing the game reflect their natural state. In the former scenario the game would be an aBCI, and in the latter a pBCI. Moreover, a blend of these two is highly probable during the game. Therefore BCI interaction methods are applicable to the applications but dependent on the user and the context.

3. BRAIN–COMPUTER INTERFACES IN MULTIMODAL INTERACTION

Having established basic knowledge about MMI and BCI, we now move on to real-time applications that involve both concepts. Following a pragmatist approach, we categorize the widely agreed benefits of using MMI in HCI and within each category survey related multimodal BCIs. As mentioned earlier, MMI has not been largely explored for BCI, so there are not many multimodal BCI applications. We discuss the potential contribution of multimodal BCIs in cases where no such applications exist for a category.

Before describing the use of BCIs in MMI, we want to mention a highly related concept called the hybrid BCI (hBCI), which was first coined by Pfurtscheller et al. (2010) and used mostly in the context of assistive technologies. They defined hybrid BCI as a system that uses two different brain signals (such as EEG and fMRI), or one brain signal associated with multiple neuromechanisms (such as ERD/ERS and ERP), or one brain signal and another input (such as EEG and eye gaze control system). The mechanism underlying hBCI corresponds to multimodal input and even more than that, as it concerns not only different modalities but also different neuromechanisms and signal acquisition methods. However, we want to make the distinction between our point of view of employing BCI in MMI and the ideology behind the hBCI research. First, hBCI is all about combining inputs, whereas in MMI we understand a two-way communication between the human and the computer. Therefore we care about output presentation as well as the coordination of input and output. Second, the motivation for current hBCI research is to compensate the weak points of modalities and improve the overall performance of a system. Taking this motivation also into account, we would like to go beyond that and investigate how the multimodal use of BCIs can improve the quality of HCI.

3.1. Improved Error Handling

Avoiding errors to improve robustness is essential in HCI. Current nonconventional HCI modalities are still not perfectly reliable, and each technology has its own restrictions for use. Functioning in noisy environments or even in the case of sensor failures are some of the many issues of concern. Consequently the technical weaknesses of HCI modalities often cause errors during interaction. Humans are not perfect in expressing themselves either. We make errors of action, such as misspeaking, and errors of intent, such as doing something that we actually did not mean to do. In some cases we do not make an error, but what we do is ambiguous. All the factors mentioned here can cause errors in HCI.

In traditional HCI, error prevention is mostly accomplished by asking the user to confirm an action before committing to it. Indeed, frequently, the confirmation dialog boxes save us from inconveniences such as deleting a file accidentally or sending an e-mail with an empty subject field. Nevertheless, some associated drawbacks are still present. Confirmation dialogs cost additional time, and when they become numerous they may introduce frustration to the user. Moreover, depending on the familiarity and frequency of the task, confirmation boxes sometimes become integrated into the actual command and lose their functionality. Therefore, in most systems, there is the possibility of reversing (undoing) a previously executed command to eliminate an unavoided error.

BCI can prevent errors in the same manner as just described but in a way that appears seamless to the user by utilizing ErrPs. Ferrez and del R. Millán (2005, 2007) identified four types of ErrPs. The first type, *response ErrP*, arises at the time a user realizes a self-made error. The second type, *feedback ErrP*, occurs also at the time a user realizes a self-made error, but in this case the user is not aware of his error until he is informed by some feedback. In the third type, *interaction ErrP*, the error is not caused by the user but by the system instead. Finally the fourth type, *observation ErrP*, arises when the user is not involved directly in the interaction but rather witnesses an error made by an operator during a choice reaction task, in which the operator is required to respond to stimuli as quickly as possible.

In a theoretical setup for preventing errors (see Figure 3), a BCI system would first interpret the user input but would not execute it immediately. The input would be evaluated or executed only if it was not followed by an ErrP. If the ErrP occurred after the input had already been executed (e.g., in the case of feedback ErrP), then the execution would be rolled back. In a way, ErrP would automatize the confirmation through the dialog box and the subsequent undoing in traditional HCI. Such a mechanism requires accurate single trial ErrP detection to be successful. Simulation studies have demonstrated the theoretical success of error handling with BCI. Ferrez and del R. Millán (2007) applied this methodology in a simulated, keyboard-based human–robot interaction experiment, by not executing a command if an interaction ErrP is detected afterward, and they reported significant improvement in BCI performance in

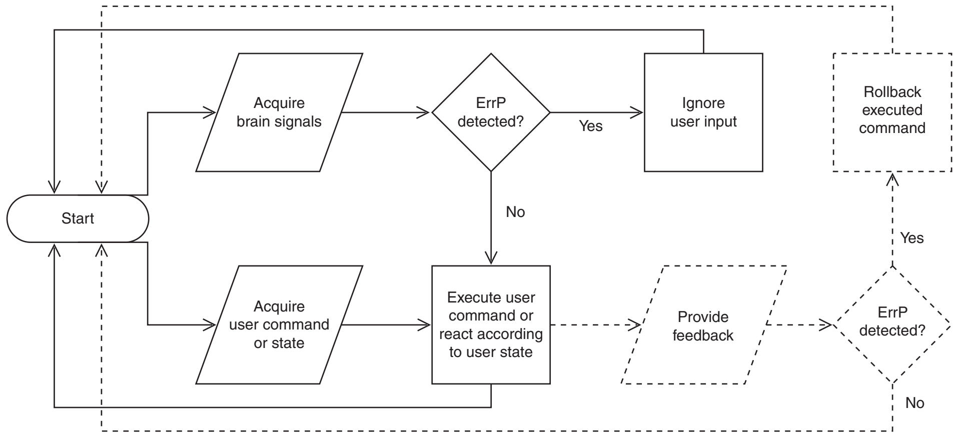

FIG. 3. Theoretical error prevention using brain–computer interfaces (BCIs). *Note.* Brain signals are acquired by the BCI and checked for the presence of error-related potential (ErrPs). If no ErrP is detected, the user command is executed or an action is taken according to the user state. Otherwise, user input is ignored. If the ErrP occurs after the feedback (dashed elements in the figure), the executed command is rolled back.

terms of bit rate. Similarly, Schalk, Wolpaw, McFarland, and Pfurtscheller (2000) used response ErrP in a simulated target selection experiment with ERD/ERS-based cursor movements and obtained gains in classification accuracy and bit rate.

3.2. Improved Task Performance

One of the features of MMI is redundancy (Mills & Alty, 1998) whereby the users are able to express themselves via more than one modality for the very same task and to receive feedback via more than one media for the very same response. While interacting simultaneously through multiple modalities that are imperfect in performance, weaknesses of one modality can be overcome by the others, thus the overall recognition performance can be improved. Leeb, Sagha, Chavarriaga, and del R. Millán (2011) combined EEG and electromyography (EMG; measuring the electrical activity of the muscles) signals during left/right hand movement tasks. They simulated muscle fatigue by degrading the EMG signals and showed that the recognition performance can be improved in comparison to EEG-only and EMG-only recognition. Another way to benefit from multimodal redundancy is to let the system/user switch modality during interaction. In this way, users can switch modality when they encounter an error or feel loss of control, or the system can automatically switch when an error or signal degradation is detected. Thereby, task performance can be improved. Kreilinger, Kaiser, Breitwieser, Neuper, and Müller-Putz (2011) designed an experiment where users could control a car game using a joystick and imaginary movements. For both types, four different quality measures were constantly applied to evaluate the quality of the signal. As soon as the quality of a control mode dropped below a certain threshold, the system would switch to the other mode. The authors concluded that their approach proved to be feasible for use, especially in assistive technologies where fatigue and other deteriorating factors are widely anticipated.

Brain responses and sensory processes can help in improving the performance of computers via BCIs. Analyzing the errors occurring during interaction can provide clues on the incongruity between a user and a system. We have already covered the principles of error detection with BCIs (through ErrPs) in the previous subsection. An adaptive system may train itself using the ErrPs and optimize its performance specifically for a user or in general. Use of errors to improve task performance has been demonstrated by Förster et al. (2010) on their hand gesture recognizer. In their experiment, the recognizer analyzed the users' ErrPs during interaction, and this way it became aware of its own performance. Thereupon, it was able to improve itself through reoccurring detection of ErrP signals.

Human sensory capabilities are superior to any computer technique in terms of speed and quality. It costs us just a blink of an eye to recognize an object, whereas a computer can recognize it only within a second through extensive training on a limited collection of objects (Tolba, El-Baz, & El-Harby, 2006). Moreover, we have the capability to recognize even the childhood photograph of a now-old person, whereas for a computer this is an extremely challenging task. Parra et al. (2008) developed a target detector that rapidly displays a collection of images (10–20 images per second) to the user and detects if an image matches the target image based on the P300 response from the user. For an image with a low number of matches in a collection, trained image analysts were able to identify a greater number of matching images in a shorter time using the target detector than by simply observing the thumbnails.

A user's task performance might be affected by his mental and emotional state. For assessing mental workload (Kramer & Weber, 2000) and emotional state (Chanel et al., 2009; Mandryk, Inkpen, & Calvert, 2006), physiological methods such as galvanic skin response (GSR; measuring conductivity of the skin), electrocardiography (ECG; measuring electrical heart activity), and electrooculography (EOG; measuring electrical eye activity) have been proposed. However, physiological sensors make use of much slower responses of the body when compared to the fast brain activity acquired by the EEG. Therefore, they require baseline periods and longer intertrial intervals, in which the subject establishes a true neutral state. In contrast, EEG studies do not require long intervals as EEG provides high temporal resolution. For these reasons it is advantageous to use EEG-based BCIs for mental/emotional state detection over the physiological methods.

Mental state is crucial for critical human tasks such as driving, operating an aircraft, or controlling the traffic. By monitoring the mental workload of the user, a system may adapt task difficulty or intensity for safe operation and optimal task performance. There are a number of examples of BCI applications for real-time mental state detection while the user is performing multiple tasks. In a typical scenario, the user performs a critical task and the BCI system monitors the user's mental load or drowsiness. The system warns the user or adjusts the settings in order to improve task performance or prevent undesired consequences. Among many indicators of mental state, ERD of parietal alpha rhythm, ERS of frontal theta rhythm, and amplitude of P300 response are the most prominent neuromechanisms used in mental workload (Holm, Lukander, Korpela, Sallinen, & Müller, 2009) and drowsiness (Oken, Salinsky, & Elsas, 2006) detection. Kohlmorgen et al. (2007) developed a system that monitors the mental workload of drivers in a real car on the highway in moderate traffic conditions. The drivers also had auxiliary tasks other than driving. The system mitigated the workload on the auxiliary tasks in high workload conditions, which improved the performance of drivers in terms of shorter response time. Lin et al. (2008) implemented a system that monitors the drowsiness of drivers in a highway VE and warns them when this condition occurs. They computed the system performance as the correlation between the detected drowsiness and actual driving errors and reported an average performance of 75%. However, the influence of the real-time warning system on decreasing the amount of errors was not evaluated.

Emotions also play a role in cognition. For example, frustration diminishing abilities with respect to attention, memory retention, learning, and thinking creatively (Klein, Moon, & Picard, 2002). Thus, by helping to regulate certain emotional states a better HCI performance may be achieved for certain tasks. Emotions are complex to understand, model, and thus measure. They are subjective and might often involve other brain functions, such as memory access. Thus, although a lot of research has been done on finding emotion correlations in the brain, there is much inconsistency as well. We do not delve into the details of these studies, and refer the reader to the comprehensive survey by Chanel et al. (2009). Use of BCIs as therapeutic tools for regulating the emotions through neurofeedback has already been proposed (Johnston, Boehm, Healy, Goebel, & Linden, 2010). But so far no real-time BCI system has been developed for investigating the effect of regulating emotions on task performance.

3.3. Improved User Experience

The most up-to-date definition of user experience according to the ISO standard (ISO 9241-210:2010, 2010) is, "A person's perceptions and responses that result from the use or anticipated use of a product, system or service." The HCI community is continuously trying to improve the definition of user experience (Law, Roto, Hassenzahl, Vermeeren, & Kort, 2009) and suggesting models for it (Nacke, Drachen, & Goebel, 2010). One of these models, upon which we base our discussion, is the model from Hassenzahl (2004) suggesting two sets of attributes of a product that affect the user experience: pragmatic and hedonic. Pragmatic attributes are those relating to a product's functionality (utility) and ways to access functionality (usability). Biological research has shown that human perception is multisensory (Murphy, 1996). We perceive things around us by integrating information provided by our multiple senses. This implies that HHI is multisensory, so MMI is analogous to HHI in this respect. By this analogy it should be natural and easier to interact with an MUI than a unimodal one, and therefore MMI should improve user experience. Several studies have practically confirmed that people prefer to use multiple-action modalities for certain tasks. Examples include using speech and gestures for manipulating virtual objects (Hauptmann & McAvinney, 1993) and maps (Oviatt et al., 2000). Hedonic attributes, on the other hand, are those that provide stimulation, communicate user identity, and provoke valued memories. A vintage product might bring back a memory so has a hedonic attribute. A user interface offering novel features also has such an attribute, as it might stimulate the user with the new possibilities it offers. In a study by Wechsung and Naumann (2009), a multimodal remote controller was found to improve user experience in comparison to a conventional remote TV controller, although the authors note that there are studies reporting opposite findings.

Until very recently, BCI research has concentrated only on optimizing the performance of the systems. Not much attention has been paid to the user's experience while interacting with the system. Nonetheless, a very few examples did consider the factors affecting user experience in BCI applications. Van de Laar, Reuderink, Plass-Oude Bos, and Heylen (2010) examined the differences in user experience between actual and imaginary movements during a BCI game. They reported that performing imagined movement is more of a challenge than actual movement. Groenegress, Holzner, Guger, and Slater (2010) assessed how mental workload imposed by a BCI affects sense of presence in a VE. They compared P300-based selection against gaze-based selection and found that the reported presence scores for the P300-based selection were much lower than those for the gaze-based selection. All the BCI systems considered in aforementioned studies are unimodal, and so far there has been no study investigating the added value of BCI to user experience when used in MMI. Using BCI may add to user experience because of the novelty or might detract from it due to the effort it may demand. More applications and research are required to answer questions about user experience in the multimodal use of BCIs.

3.4. Broader User Spectrum

While using a redundantly multimodal interface, users' choice of the modality depends on various factors such as their skills, impairments, or preferences. From the input point of view, a user with a broken hand bone would prefer speech over the mouse, and a user who can draw well would prefer pen input over speech. Likewise, for the output, a hearing impaired individual would appreciate visual output while a visually impaired user an acoustic one. So redundant design leads to a broader range of users.

Sometimes there are very limited options to choose from, or there are no options at all as in the case of locked-in syndrome (LiS). LiS is a neurological condition consisting of tetraplegia and paralysis of all cranial nerves except vertical eye movements (Bauer, Gerstenbrand, & Rumpl, 1979) and is a characteristic of multiple sclerosis or amyotrophic lateral sclerosis (also known as Lou Gehrig's disease). The clinical BCI applications have long been addressing this case. For example, a patient with LiS can use a prosthetic hand (Guger, Harkam, Hertnaes, & Pfurtscheller, 1999) or control a cursor on the computer screen (Wolpaw, McFarland, Vaughan, & Schalk, 2003) by imagining left- and right-hand movements. For a comprehensive overview of invasive and noninvasive BCIs to restore the interaction of the LiS patients, the reader is referred to Kübler, Nijboer, and Birbaumer (2007).

Gaze and blink trackers based on camera (Betke, 1998) or EOG (Barea, Boquete, Mazo, & Lopez, 2002) input can also be used for LiS patients because their vision is intact. The true value of the BCI becomes clear in the case of people with complete (or total) LiS (CLiS), which is the condition of virtually total immobility including all eye movements combined with preserved consciousness (Bauer et al., 1979). A nonvisual BCI is the sole means of restoring the interaction capability of a CLiS user. Despite successful demonstrations of auditory BCIs for the locked-in (Furdea et al., 2009), so far no patient in CLiS has been able to reliably use such a device (Kübler & Birbaumer, 2008). Nevertheless, research has shown that it is possible to detect the consciousness in CLiS with an auditory BCI (Schnakers et al., 2009), so BCIs are still promising devices for this group of people.

4. TECHNIQUES, CHALLENGES AND GUIDELINES FOR EMPLOYING BCI IN MMI

So far we have covered the objectives of MMI and the principles of BCI together with the recent applications. We now continue with the techniques for employing BCI in MMI and provide some guidelines for the associated challenges. We base our discussion on the driving principles of MMI as mentioned by Dumas, Lalanne, and Oviatt (2009) as well as Jaimes and Sebe (2007). These principles are fusion of input modalities, fission of output modalities, integration architectures, and data collection procedures.

4.1. Multimodal Information Integration (Fusion)

The integration of information from different input modalities is a typical process in MUIs. In this process, also known as fusion, the goal is to extract a common meaning from multiple inputs and make a decision accordingly. The multimodal fusion can be executed in three main levels (see Figure 4), namely, the

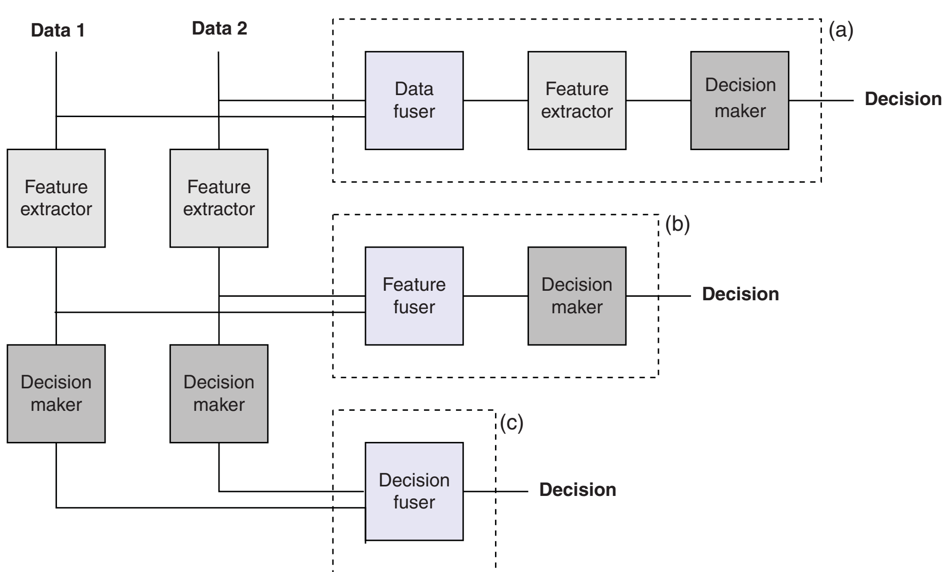

FIG. 4. Three main levels of multimodal signal integration (fusion) illustrated on bimodal data: (a) data-level fusion, (b) feature-level fusion and (c) decision-level fusion (color figure available online).

data, feature, and decision levels (Sharma, Pavlovic, & Huang, 1998). Next we define each of these levels and describe relevant multimodal BCI applications.

Data-level fusion. Data-level fusion is performed while integrating multiple observations of the same kind, such as audio data coming from camera and microphone. The data are not processed before the fusion, so a high level of information detail is preserved. On the other hand, the absence of preprocessing makes this level of fusion susceptible to noise and sensor failure as well. In HCI, data-level fusion is not commonly applied as the modalities usually deliver different types of data, such as speech and gestures yielding audio and visual data, respectively.

Within BCI research, data-level fusion is mostly used for subtracting electrical eye activity (EOG) from the EEG in the domain of artefact correction. EOG contamination is difficult to avoid, especially in applications for nondisabled users, and simple artifact rejection causes data loss. Therefore, artifact reduction and correction is an essential topic in BCI research. We recommend that reader would look at the survey by Fatourechi, Bashashati, Ward, and Birch (2007) for a very good review of EEG artifacts and artifact removal methods.

Feature-level fusion. Feature-level fusion is used for modalities, which are not necessarily of the same type but are tightly coupled or synchronized, such as speech and lip movements. As the data are processed in this case, information loss is inevitable. On the other hand, relying on extracted features rather than the raw data increases robustness. In this level of fusion, the size of feature sets can be very large so dimensionality reduction is a commonly applied technique.

A typical application of this level of fusion in BCI is for mental and emotional state detection. Just as speech and lip movements are indicators of the act of speaking, neurophysiological changes in the body are indicators of certain states. As an example, ECG measures the electrical activity of the heart. Heart rate, interbeat interval, heart rate variability (HRV), and respiratory sinus arrhythmia are some features which can be extracted from the ECG data. These features are representatives for certain states. For example, heart rate can differentiate between positive and negative emotions, and HRV is associated with stress (Mandryk et al., 2006). A state can correlate with multiple features extracted from different sensors. For instance, stress can be traced in respiration rate, HRV, and some other features extracted from GSR, EEG, and EMG data. Thus by combining these features, it is possible to assess a person's stress level.

In general practice, the feature sets obtained from different modalities are concatenated to form a single vector and fed into a classifier. A recent example is the work by Chanel, Kronegg, Grandjean, and Pun (2006) for arousal assessment using EEG and peripheral sensors (a temperature sensor, GSR to measure the skin conductance, plethysmograph to measure blood pressure, and respiration belt to evaluate abdominal and thoracic movements). Six EEG features and 18 peripheral features were extracted, concatenated, and fed into naive Bayes and Fisher's discriminant classifiers. The results showed that fusion provides more robust results than using only EEG or physiological data.

Feature-level fusion is also used for combining multiple brain measurement methods. As mentioned in the Acquiring (Measuring) Brain Activity (Imaging Modalities) section, the measurement methods have their strong and weak properties. For instance EEG and MEG are weak in spatial resolution but strong in temporal resolution, whereas the situation is the opposite for fMRI. Therefore, integrating the methods with complementary assets can overcome the drawbacks of each. Merzagora et al. (2009) extracted two EEG and three fNIRS features during a working memory task carried out by six people and combined in a single feature vector. They showed that integrating the modalities improved the spatiotemporal resolution. Another approach is to use the features extracted from one measurement method as constraints (prior knowledge) in source reconstruction of another one. Friston (2009) provided some examples for the fusion of EEG and fMRI in this fashion.

Decision-level fusion. The third type of fusion is the decision level (or classifier) fusion, which is used for integrating modalities that are not necessarily tightly coupled, such as pen and speech input. In this case, data from each input are modeled independently and the unimodal recognition results are combined. Therefore low-level information is not preserved. However, because data are preprocessed, tolerance to sensor failure and noise is higher.

In this level of fusion, modalities can be combined to perform a single task or separate tasks contributing to a single higher level task. In the former case, the combination scheme is the main concern. There are various alternatives depending on the format of the results and availability of training data. If scores are available from the modalities for every alternative decision and the scores are in the same scale (or can be normalized), then they can directly be combined. This can be done by simply averaging the scores, but there are more advanced methods as well (Aslam & Montague, 2001). If information is available on the reliability of modalities—for example, via some training data—then the scores can be weighted in proportion to the reliability and then combined (also known as linear combination). On the other hand, if there are no scores but only a ranking of possible decisions (i.e., an n-best list) or the scores coming from the modalities are not comparable, then ranks can be combined. Again, depending on the availability of training data, the contribution of each modality can be weighted. Nuray and Can (2006) provided a nice overview of rank-based fusion methods, from an information science point of view. This type of decision-level fusion has been used for combining neurophysiological data such as EEG and fNIRS in a working memory task (Merzagora et al., 2009) or neurological and physiological data, such as EEG and pupil data for visual detection analysis (Qian et al., 2009).

The latter type of decision-level fusion just mentioned (i.e., where there are multiple tasks but a single higher level task) is sometimes called semantic-level fusion. It is used in combining BCI with HCI (i.e., non-neurophysiological) modalities. Mühl et al. (2010) employed decision-level fusion in their multimodal BCI game, which uses keyboard input and EEG signals. They combined key presses with the player's parietal alpha activity, associated with relaxed alertness, to position the player avatar in the game world. The main concern in this type of decision-level fusion is not the combination scheme but rather the synchronization of modalities. We mention this issue in the next subsection.

Decision-level fusion can be used to combine any sorts of modalities, even with simple classification algorithms. Nevertheless, the underlying assumption of conditional independence between the modalities must be taken into account. Especially with tightly coupled modalities, the data provided by the modalities might contain mutual information. So modeling the data independently would cause loss of information of mutual correlation. In this case one could opt in for model-level fusion in which data from each modality are modeled separately (using a model such as the Hidden Markov Model) and the resulting models are fused together using a probabilistic fusion model that is optimal according to the maximum entropy principle and a maximum mutual information criterion (Zeng, Pantic, & Huang, 2009).

4.2. Integration Architectures

In multimodal fusion, synchronization of the modalities is an indispensable step. Lags may occur due to technology (as in the case of automatic speech recognition or BCI), multimodal system architecture, or user differences in multimodal integration patterns (Dumas et al., 2009). For this reason, multiagent architectures are desirable for distributing the processes and coordinating the modalities. Ideally, all data should be time-stamped at each individual modality and should be verified before the fusion.

Theoretically BCI modalities can provide a temporal resolution of less than a second (van Gerven et al., 2009). But because BCI is not as robust as other HCI modalities, often a task is repeatedly performed or a mental/emotional state is kept stable within a time window and the cumulated data are analyzed. For example Savran et al. (2006) used blocks 12.5 s long to fuse EEG and functional NIRS (fNIRS) data for emotion detection. In each block, they showed five stimulating images to the subject to accumulate data. Mühl et al. (2010) used 3 s of overlapping EEG windows to combine with the keyboard input in their multimodal game. The windowing technique introduces a trade-off between speed and accuracy. The longer the window is (i.e., the more repetitions there are), the more the data accumulates, thus the higher is the chance of correct recognition. On the other hand, as the window gets longer, the observed speed of the interface decreases up to several seconds.

The trade-off between speed and latency might make it difficult to achieve fast-paced BCI applications with high performance. Zander, Gaertner, Kothe, and Vilimek (2011) developed a multimodal speller where users make selections by dwelling (i.e., staring for some time) or imagining hand movements on the choices. They compared short and long dwell selections against the BCI selection in terms of accuracy and speed. The results showed that BCI control (with 83.3%) was significantly more accurate than short dwelling (with 67.4%) and as accurate as long dwelling (with 84.7%). However, BCI was the slowest method in task completion time (easy task: 5.90 s; hard task: 8.84 s) compared to short (3.98 s; 5.38 s) and long (4.79 s; 7.37 s) dwelling. This study shows that the proportion of the amount of time spent on collecting data to the amount of collected data is a matter of consideration.

4.3. Data Collection Procedures

There are experimental challenges associated with combining different measurement methods. The multimodal data can be collected either in separate sessions or simultaneously. Separate recordings are preferred if an experiment can be repeated more than once with a high degree of reliability of the data, whereas simultaneous measurements are essential in cases such as when a subject's state might influence the results as in monitoring spontaneous activity or sleep state changes (Halchenko, Hansen, & Pearlmutter, 2005). During separate acquisition habituation effects, variations in the stimulation paradigm, or any other difference between sessions, might lead to differential activity of the brain. On the other hand, when recording simultaneously, reciprocal electromagnetic perturbations between the measurement devices pose a constraint. For example, during fMRI-EEG recordings, the magnetic resonance field strength and positioning of the EEG recording equipment induce voltages that add linearly to the EEG signal and obscure the biological signal of interest (Daunizeau, Laufs, & Friston, 2010). Halchenko et al. (2005) listed some artifact removal techniques and protocols for simultaneous fMRI-EEG recordings.

There are also physical constraints in simultaneous multimodal measurements. The number of sites on the head where sensors can be placed is limited, and once one type of sensor is placed, another one cannot simply be applied on the same location. For this reason, different types of sensors can record at different locations exploiting different correlates of the very same activity. Alternatively, integrated probes can be used as done by Cooper et al. (2009) for simultaneous EEG and NIRS recordings.

When the experiments are conducted in a laboratory (posed, controlled) setting, the optimal conditions for the operation of the recording device are satisfied. Isolation against electromagnetic inference and audio noise would be ensured, and subjects would be instructed to avoid movements, blinks, or speaking, which could cause physiological noise (Gunes, Piccardi, & Pantic, 2008). However in real-world (spontaneous, natural) contexts, all the aforementioned noises occur naturally and should either be handled or, better, included in the analyses. Solovey et al. (2009) identified some considerations and provided guidelines for using fNIRS in realistic HCI settings. They examined whether typical human behavior (e.g., head and facial movements) or computer interaction (e.g., use of keyboard and mouse) interfere with fNIRS measurements. They stated that, provided the interference is corrected according to the guidelines they proposed, fNIRS can be used in realistic experiment environments. Lotte et al. (2009) investigated the feasibility of using an EEG system based on P300 signals with a moving subject. They found that it was possible to detect the P300 signal while the subject was sitting, standing, or walking. Nam, Li, and Johnson (2010) investigated the usability of a P300 speller by assessing how background noise affected user performance and BCI usage preference. They reported that participants had better performance in the noisy condition than in the quiet condition. Gürkök, Poel, and Zwiers (2010) classified EEG signals of imaginary movements in presence of speech. They showed that with their method the presence of speech during motor imagery did not affect the classification accuracy significantly.

4.4. Multimodal Information Presentation (Fission)

In MMI, fission is the process of conveying the output of a system using multiple modalities, such as the audio, visual, and haptic channels. As an input modality, BCI does not contribute to multimodal information presentation (fission). However multimodal fission may play a role in the functioning of BCIs.

Multimodal stimulation can enhance brain activity or response due to superadditivity, in which the multisensory response exceeds the sum of those evoked by the modality-specific stimulus components individually (Stanford & Stein, 2007). In a multimodal feedback P300 study by Brouwer, van Erp, Aloise, and Cincotti (2010), participants attended to the vibrations and/or flashes of a target presented in a stream of standards. The authors reported that classification accuracy was highest in the bimodal condition and concluded that bimodal stimuli could enhance classification results within a BCI context compared to unimodal presentations. In another study (Belitski, Farquhar, & Desain, 2011) participants used a P300 speller with visual, auditory, and audiovisual stimuli. The bimodal version outperformed the others in terms of single trial classification. Based on the findings, the authors suggested that when both sensory modalities are available, multimodal stimulation improves the performance over unimodal stimulation. They also concluded that multimodal stimulation is robust to the loss of a single sensory modality due to, for example, a disease or distraction.

On the other hand, attentional resource competition between the modalities may impede the user's ability to fully use a combined feedback. In a study by Hinterberger et al. (2004), subjects learned to regulate their brain potentials with auditory, visual, and combined visual-auditory feedback from the BCI system. The results of the study revealed that the combined feedback modality showed the smallest learning effect, indicating that combining modalities may impede learning.

With some BCI interaction paradigms, output modality selection may become critical. For rBCI, in which a stimulus is necessary to evoke the desired brain activity, the user is supposed to attend only to the stimulus presented. In this case, using mutually exclusive modalities for the feedback and the task-related stimulus might mitigate the interference between the two. Lalor et al. (2005) developed an EEG-based 3D, immersive BCI game that uses SSVEPs (thus visual stimulation) and audiovisual feedback during training. Their subjects reported that audio feedback aided in the successful sustained fixation on a particular stimulus and the inhibition of responses to distractions. Cincotti et al. (2004) used a tactile feedback modality and compared it to the visual feedback while subjects were required to perform a visually guided navigation task using imaginary movement of their hands. A significantly higher rate of mistakes was made when visual attention was divided between the control and task monitors.

For aBCI and pBCI, this is less of a concern as they do not depend on stimulation. As an example of the aBCI, Tangermann et al. (2008) demonstrated the multimedia feedback in a pinball machine controlled by imaginary movements. The users were able to control the machine, which provided rich and complex feedback as well as acoustic and visual distracters. Typical examples of pBCI are the neurofeedback applications in which the user is provided with visual and auditory feedback about his mental or emotional state (Jackson & Mappus, 2010).

5. CONCLUSIONS

Until recently, BCI and HCI researchers have not attempted to look at their own research fields from one another's perspectives. BCI research disregarded human factors such as usability, aesthetics, and social interaction but concentrated more on system performance. Likewise, HCI research did not consider BCI as a regular modality and did not include it in HCI topics such as ubiquitous computing, multimodality, and interactive systems. With the penetration of portable, wireless, easy-to-use BCI headsets into homes and the development of everyday applications, nondisabled users are becoming a target group in BCI research. This calls for investing time to explore the HCI aspects in BCI. On the other hand, the unique features offered by the BCIs can overcome the limitations of conventional modalities, so HCI research is attracted to the benefits of BCI as a modality.

For a long time, MMI was associated only with speech and gesture (or pen-based) interfaces. Despite some demonstrations as MUIs, the BCI has not been commonly employed in MMI. In this article we defined the research domains of MMI and BCI and indicated the mutual benefits between them. We showed that BCIs can improve error handling, task performance, and user experience and can broaden the user spectrum by multimodal implementations. Then we applied the principles of MMI to recent BCI research and applications in order to identify the possible challenges. We provided some guidelines to the issues such as noise and interference during multimodal fusion, distraction during multimodal output, physical constraints in multimodal deployment, and experimental constraints with respect to the user, the task, and the environment.

Perhaps there are two main requirements that HCI demands to BCI to accept it as an alternative modality: ease of operation and reliability. BCI technology is not too far from satisfying these requirements at a certain level. With consumer BCI devices, which are easy to use for the users as well as for the researchers, operation of BCIs no longer requires neuroscientists. Although much promising work has been going on to improve the reliability of the BCI, better signal processing and classification methods will always be expected. Nevertheless, as pointed out by Oviatt (1999), the main advantage of MMI is not enhanced efficiency but decreased error rate, flexibility to choose between alternating input modes, and a wider range of users. Throughout this text we have considered the contribution of BCIs in providing these and other advantages when used in MMI. The individual inadequacies experienced with BCI should not hinder concurrent exploration of the practical possibilities of integrating BCI in MMI. We do use previously developed algorithms for BCI specific tasks such as artifact removal or response detection in our applications and attempt to improve them. However, with respect to MMI methods for BCI, we are not sufficiently aware of each other. This is mainly because there is not much work being done on integrating MMI and BCI yet. Despite being far from mature, there are methods for multimodal input fusion. However the issues including but absolutely not limited to output fission methods and effects, synchronization with other modalities, and privacy and security are still open. MMI and BCI researchers should start regarding each other's work as specialized tracks of attention. We need to develop more multimodal BCI applications and report the challenges encountered. As the number of multimodal BCIs increase and their contribution to enriching HCI is appreciated, BCIs will be more widely accepted as regular HCI modalities.

REFERENCES

Allison, B. Z., & Polich, J. (2008). Workload assessment of computer gaming using a single-stimulus event-related potential paradigm. *Biological Psychology*, *77*, 277–283.

Aslam, J. A., & Montague, M. (2001). Models for metasearch. In *Proceedings of the 24th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 276–284). New York, NY: ACM.

Barea, R., Boquete, L., Mazo, M., & Lopez, E. (2002). System for assisted mobility using eye movements based on electrooculography. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *10*, 209–218.

Bauer, G., Gerstenbrand, F., & Rumpl, E. (1979). Varieties of the locked-in syndrome. *Journal of Neurology*, *221*(2), 77–91.

Belitski, A., Farquhar, J., & Desain, P. (2011). P300 audio-visual speller (Article 025022). *Journal of Neural Engineering*, *8*(2).

Berger, H. (1929). Über das elektrenkephalogramm des menschen. *Archiv für Psychiatrie und Nervenkrankheiten*, *87*(1), 527–570.

Betke, M. (1998). *Camera-based interfaces and assistive software for people with severe motion impairments* (Tech. Rep. BU-CS-2008-025). Boston, MA: Department of Computer Science, Boston University.

Bieger, J., & Molina, G. G. (2010). *Light stimulation properties to influence brain activity: A brain-computer interface application* (Tech. Rep. TN-2010-00315). Eindhoven, the Netherlands: Philips Research.

Brouwer, A.-M., van Erp, J. B. F., Aloise, F., & Cincotti, F. (2010). Tactile, visual and bimodal P300s: Could bimodal P300s boost BCI performance? *SRX Neuroscience*, *2010*, 967027.

Campbell, A. T., Choudhury, T., Hu, S., Lu, H., Mukerjee, M. K., Rabbi, M., & Raizada, R. D. (2010). NeuroPhone: Brain–mobile phone interface using a wireless EEG headset. In *MobiHeld '10: Proceedings of the second ACM SIGCOMM workshop on networking, systems, and applications on mobile handhelds* (pp. 3–8). New York, NY: ACM.

Cassell, J., Sullivan, J., Prevost, S., & Churchill, E. (Eds.). (2000). *Embodied conversational agents*. Cambridge, MA: MIT Press.

Chanel, G., Kierkels, J. J., Soleymani, M., & Pun, T. (2009). Short-term emotion assessment in a recall paradigm. *International Journal of Human–Computer Studies*, *67*, 607–627.

Chanel, G., Kronegg, J., Grandjean, D., & Pun, T. (2006). Emotion assessment: Arousal evaluation using EEG's and peripheral physiological signals. In B. Gunsel, A. K. Jain, A. M. Tekalp, & B. Sankur (Eds.), *Multimedia content representation, classification and security* (pp. 530–537). Berlin, Germany: Springer-Verlag.

Cincotti, F., Kauhanen, L., Aloise, F., Palomäki, T., Caporusso, N., Jylänki, P., . . . del R. Millán, J. (2004). Vibrotactile feedback for brain–computer interface operation. *Computational Intelligence and Neuroscience*, *2007*, 048937.

Cooper, R. J., Everdell, N. L., Enfield, L. C., Gibson, A. P., Worley, A., & Hebden, J. C. (2009). Design and evaluation of a probe for simultaneous EEG and near-infrared imaging of cortical activation. *Physics in Medicine and Biology*, *54*, 2093–2102.

Daunizeau, J., Laufs, H., & Friston, K. J. (2010). EEG–fMRI information fusion: Biophysics and data analysis. In C. Mulert & L. Lemieux (Eds.), *EEG–fMRI: Physiological basis, technique, and applications* (pp. 511–526). Berlin, Germany: Springer-Verlag.

Deuschl, G., & Eisen, A. (Eds.). (1999). *Recommendations for the practice of clinical neurophysiology* (2nd ed.). Amsterdam, the Netherlands: Elsevier.

Dumas, B., Lalanne, D., & Oviatt, S. (2009). Multimodal interfaces: A survey of principles, models and frameworks. In D. Lalanne & J. Kohlas (Eds.), *Human machine interaction: Research results of the MMI program* (pp. 3–26). Berlin, Germany: Springer-Verlag.

Edlinger, G., Holzner, C., Groenegress, C., Guger, C., & Slater, M. (2009). Goal-oriented control with brain–computer interface. In D. Schmorrow, I. Estabrooke, & M. Grootjen (Eds.), *Foundations of augmented cognition. Neuroergonomics and operational neuroscience* (pp. 732–740). Berlin, Germany: Springer-Verlag.

Fabiani, M., Gratton, G., & Federmeier, K. D. (2007). Event-related brain potentials: Methods, theory, and applications. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (3rd ed., pp. 85–119). Cambridge, UK: Cambridge University Press.

Farwell, L. A., & Donchin, E. (1988). Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography and Clinical Neurophysiology*, *70*, 510–523.

Fatourechi, M., Bashashati, A., Ward, R. K., & Birch, G. E. (2007). EMG and EOG artifacts in brain computer interface systems: A survey. *Clinical Neurophysiology*, *118*, 480–494.

Fatourechi, M., Birch, G. E., & Ward, R. K. (2007). A self-paced brain interface system that uses movement related potentials and changes in the power of brain rhythms. *Journal of Computational Neuroscience*, *23*(1), 21–37.

Ferrez, P. W., & del R. Millán, J. (2005). You are wrong!: Automatic detection of interaction errors from brain waves. In *IJCAI'05: Proceedings of the 19th International Joint Conference on Artificial intelligence* (pp. 1413–1418). San Francisco, CA: Morgan Kaufmann.

Ferrez, P. W., & del R. Millán, J. (2007). Error-related EEG potentials in brain-computer interfaces. In G. Dornhege, J. del R. Millán, T. Hinterberger,

D. J. McFarland, & K.-R. Müller (Eds.), *Toward brain-computer interfacing* (pp. 291–301). Cambridge, MA: MIT Press.

Fink, A., Grabner, R. H., Benedek, M., Reishofer, G., Hauswirth, V., Fally, M., . . . Neubauer, A. C. (2009). The creative brain: Investigation of brain activity during creative problem solving by means of EEG and fMRI. *Human Brain Mapping*, *30*, 734–748.

Förster, K., Biasiucci, A., Chavarriaga, R., del R. Millán, J., Roggen, D., & Tröster, G. (2010). On the use of brain decoded signals for online user adaptive gesture recognition systems. In P. Floréen, A. Krüger, & M. Spasojevic (Eds.), *Pervasive computing* (pp. 427–444). Berlin, Germany: Springer.

Friston, K. J. (2009). Modalities, modes, and models in functional neuroimaging. *Science*, *326*(5951), 399–403.

Furdea, A., Halder, S., Krusienski, D., Bross, D., Nijboer, F., Birbaumer, N., & Kübler, A. (2009). An auditory oddball (P300) spelling system for brain–computer interfaces. *Psychophysiology*, *46*, 617–625.

Gevensleben, H., Holl, B., Albrecht, B., Vogel, C., Schlamp, D., Kratz, O., . . . Heinrich, H. (2009). Is neurofeedback an efficacious treatment for ADHD? A randomised controlled clinical trial. *Journal of Child Psychology and Psychiatry*, *50*, 780–789.

Gloor, P. (1969). *Hans Berger on the electroencephalogram of man. The fourteen original reports on the human encephalogram*. Amsterdam, the Netherlands: Elsevier.

Groenegress, C., Holzner, C., Guger, C., & Slater, M. (2010). Effects of P300 based BCI use on reported presence in a virtual environment. *Presence: Teleoperators and Virtual Environments*, *19*(1), 1–11.

Guger, C., Harkam, W., Hertnaes, C., & Pfurtscheller, G. (1999). Prosthetic control by an EEG-based brain–computer interface (BCI). In *Proceedings of the 5th European Conference for Advancement of Assistive Technology (AAATE '99)*. Amsterdam, the Netherlands: IOS Press.

Gunes, H., Piccardi, M., & Pantic, M. (2008). From the lab to the real world: Affect recognition using multiple cues and modalities. In J. Or (Ed.), *Affective computing: Focus on emotion expression, synthesis, and recognition* (pp. 185–218). Vienna, Austria: InTech Education and Publishing.

Gürkök, H., Poel, M., & Zwiers, J. (2010). Classifying motor imagery in presence of speech. In *IJCNN '10: Proceedings of the International Joint Conference on Neural Networks* (pp. 1235–1242). Piscataway, NJ: IEEE.

Halchenko, Y. O., Hanson, S. J., & Pearlmutter, B. A. (2005). Multimodal integration: fMRI, MRI, EEG, MEG. In L. Landini, V. Positano, & M. Santarelli (Eds.), *Advanced image processing in magnetic resonance imaging* (pp. 223–265). Boca Raton, FL: CRC Press.

Hassenzahl, M. (2004). The thing and I: Understanding the relationship between user and product. In M. A. Blythe, K. Overbeeke, A. F. Monk, & P. C. Wright (Eds.), *Funology: From usability to enjoyment* pp. 31–42. Dordrecht, the Netherlands: Kluwer Academic.

Hauptmann, A. G., & McAvinney, P. (1993). Gestures with speech for graphic manipulation. *International Journal of Man–Machine Studies*, *38*, 231–249.

Herdman, A. T., Lins, O., van Roon, P., Stapells, D. R., Scherg, M., & Picton, T. W. (2002). Intracerebral sources of human auditory steady-state responses. *Brain Topography*, *15*(2), 69–86.

Hinterberger, T., Neumann, N., Pham, M., Kübler, A., Grether, A., Hofmayer, N., . . . Birbaumer, N. (2004). A multimodal brain-based feedback and communication system. *Experimental Brain Research*, *154*, 521–526.

Holm, A., Lukander, K., Korpela, J., Sallinen, M., & Müller, K. M. I. (2009). Estimating brain load from the EEG. *Scientific World Journal*, *9*, 639–651.

ISO 9241-210:2010. (2010). *Ergonomics of human-system interaction – Part 210: Human-centred design for interactive systems*. Geneva, Switzerland: ISO.

Jackson, M. M., & Mappus, R. (2010). Applications for brain–computer interfaces. In A. Nijholt & D. Tan (Eds.), *Brain–computer interfaces: Applying our minds to human–computer interaction* (pp. 89–103). London, UK: Springer-Verlag London.

Jaimes, A., & Sebe, N. (2007). Multimodal human–computer interaction: A survey. *Computer Vision and Image Understanding*, *108*, 116–134.

Johnston, S., Boehm, S., Healy, D., Goebel, R., & Linden, D. (2010). Neurofeedback: A promising tool for the self-regulation of emotion networks. *NeuroImage*, *49*(1), 1066–1072.

Klein, J., Moon, Y., & Picard, R. W. (2002). This computer responds to user frustration: Theory, design, and results. *Interacting with Computers*, *14*, 119–140.

Kohlmorgen, J., Dornhege, G., Braun, M., Blankertz, B., Müller, K.-R., Curio, G., . . . Kincses, W. (2007). Improving human performance in a real operating environment through real-time mental workload detection. In G. Dornhege, J. del R. Millán, T. Hinterberger, D. J. McFarland, & K.-R. Müller (Eds.), *Toward brain–computer interfacing* (pp. 409–422). Cambridge, MA: MIT Press.

Kramer, A. F., & Weber, T. (2000). Applications of psychophysiology to human factors. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (2nd ed., pp. 794–814). New York, NY: Cambridge University Press.

Kreilinger, A., Kaiser, V., Breitwieser, C., Neuper, C., & Müller-Putz, G. R. (2011). Fusion of manual control and BCI using long term and short term quality measures. *International Journal of Bioelectromagnetism*, *13*, 110–111.

Krepki, R., Blankertz, B., Curio, G., & Müller, K.-R. (2007). The Berlin brain–computer interface (BBCI)—Towards a new communication channel for online control in gaming applications. *Multimedia Tools and Applications*, *33*(1), 73–90.

Kübler, A., & Birbaumer, N. (2008). Brain–computer interfaces and communication in paralysis: Extinction of goal directed thinking in completely paralysed patients? *Clinical Neurophysiology*, *119*, 2658–2666.

Kübler, A., & Müller, K.-R. (2007). An introduction to brain–computer interfacing. In G. Dornhege, J. del R. Millán, T. Hinterberger, D. J. McFarland, & K.-R. Müller (Eds.), *Toward brain–computer interfacing* (pp. 1–25). Cambridge, MA: MIT Press.

Kübler, A., Nijboer, F., & Birbaumer, N. (2007). Brain–computer interfaces for communication and motor control - perspectives on clinical applications. In G. Dornhege, J. del R. Millán, T. Hinterberger, D. J. McFarland, & K.-R. Müller (Eds.), *Toward brain–computer interfacing* (pp. 373–391). Cambridge, MA: MIT Press.

Lalor, E. C., Kelly, S. P., Finucane, C., Burke, R., Smith, R., Reilly, R. B., & McDarby, G. (2005). Steady-state VEP-based brain–computer interface control in an immersive 3D gaming environment. *EURASIP Journal on Applied Signal Processing*, *2005*(19), 3156–3164.

Larson, J. A., Raman, T., & Raggett, D. (2003). W3C multimodal interaction framework (Tech. Rep.). Retrieved from http://www.w3.org/TR/mmi-framework W3C

Law, E. L.-C., Roto, V., Hassenzahl, M., Vermeeren, A. P., & Kort, J. (2009). Understanding, scoping and defining user experience: a survey approach. In *CHI '09: Proceedings of the 27th international conference on Human factors in computing systems* (pp. 719–728). New York, NY: ACM.

Lebedev, M. A., & Nicolelis, M. A. (2006). Brain–machine interfaces: past, present and future. *Trends in Neurosciences*, *29*, 536–546.

Leeb, R., Friedman, D., Müller-Putz, G. R., Scherer, R., Slater, M., & Pfurtscheller, G. (2007). Self-paced (asynchronous) BCI control of a wheelchair in virtual environments: a case study with a tetraplegic. *Computational Intelligence and Neuroscience*, *2007*, 1–12.

Leeb, R., Sagha, H., Chavarriaga, R., & del R. Millán, J. (2011). A hybrid brain–computer interface based on the fusion of electroencephalographic and electromyographic activities (Article 025011). *Journal of Neural Engineering*, *8*(2).

Li, Y., Nam, C. S., Shadden, B. B., & Johnson, S. L. (2011). A P300-based brain–computer interface: Effects of interface type and screen size. *International Journal of Human–Computer Interaction*, *27*(1), 52–68.

Lin, C.-T., Chen, Y.-C., Huang, T.-Y., Chiu, T.-T., Ko, L.-W., Liang, S.-F., . . . Duann, J.-R. (2008). Development of wireless brain–computer interface with embedded multitask scheduling and its application on real-time driver's drowsiness detection and warning. *IEEE Transactions on Biomedical Engineering*, *55*, 1582–1591.

Lotte, F., Fujisawa, J., Touyama, H., Ito, R., Hirose, M., & Lécuyer, A. (2009). Towards ambulatory brain–computer interfaces: a pilot study with P300 signals. In *ACE '09: Proceedings of the International Conference on*

Advances in Computer Entertainment Technology (pp. 336–339). New York, NY: ACM.

Luck, S. J. (2005). *An introduction to event-related potentials and their neural origins*. Cambridge, MA: MIT Press.

Mandryk, R. L., Inkpen, K. M., & Calvert, T. W. (2006). Using psychophysiological techniques to measure user experience with entertainment technologies. *Behaviour and Information Technology*, *25*, 141–158.

Martinez, P., Bakardjian, H., & Cichocki, A. (2007). Fully online multicommand brain–computer interface with visual neurofeedback using SSVEP paradigm. *Computational Intelligence and Neuroscience*, *2007*, 13–13.

Merzagora, A. C., Izzetoglu, M., Polikar, R., Weisser, V., Onaral, B., & Schultheis, M. T. (2009). Functional near-infrared spectroscopy and electroencephalography: A multimodal imaging approach. In R. Goebel, J. Siekmann, & W.Wahlster (Eds.), *Foundations of augmented cognition: Neuroergonomics and operational neuroscience* (pp. 417–426). Berlin, Germany: Springer-Verlag.

Mills, K. M., & Alty, J. L. (1998). Investigating the role of redundancy in multimodal input systems. In I. Wachsmuth, & M. Fröhlich (Eds.), *Gesture and sign language in human–computer interaction* (pp. 159171). Berlin, Germany: Springer-Verlag.

Mühl, C., Gürkök, H., Plass-Oude Bos, D., Thurlings, M. E., Scherffig, L., Duvinage, M., . . . Heylen, D. (2010). Bacteria hunt: Evaluating multi-paradigm BCI interaction. *Journal on Multimodal User Interfaces*, *4*(1), 11–25.

Muller-Putz, G., Scherer, R., Neuper, C., & Pfurtscheller, G. (2006). Steady state somatosensory evoked potentials: suitable brain signals for braincomputer interfaces? *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *14*(1), 30–37.

Murphy, R. R. (1996). Biological and cognitive foundations of intelligent sensor fusion. *IEEE Transactions on Systems, Man, and Cybernetics*, *26*(1), 42–51.

Nacke, L., Drachen, A., & Goebel, S. (2010). Methods for evaluating gameplay experience in a serious gaming context. *International Journal of Computer Science in Sport*, *9*(2).

Nam, C. S., Li, Y., & Johnson, S. (2010). Evaluation of P300-based brain–computer interface in real-world contexts. *International Journal of Human–Computer Interaction*, *26*, 621–637.

Nikolaev, A. R., & Anokhin, A. P. (1998). EEG frequency ranges during perception and mental rotation of two- and three-dimensional objects. *Neuroscience and Behavioral Physiology*, *28*, 670–677.

Nuray, R., & Can, F. (2006). Automatic ranking of information retrieval systems using data fusion. *Information Processing and Management*, *42*, 595–614.

Oken, B., Salinsky, M., & Elsas, S. (2006). Vigilance, alertness, or sustained attention: physiological basis and measurement. *Clinical Neurophysiology*, *117*, 1885–1901.

Oviatt, S. (1999). Ten myths of multimodal interaction. *Communications of the ACM*, *42*(11), 74–81.

Oviatt, S. (2008). Multimodal interfaces. In J. A. Jacko, & A. Sears (Eds.), *The human–computer interaction handbook: fundamentals, evolving technologies and emerging applications* (2nd ed., pp. 413–428). New York, NY: Erlbaum.

Oviatt, S., Cohen, P., Wu, L., Duncan, L., Suhm, B., Bers, J., . . . Ferro, D. (2000). Designing the user interface for multimodal speech and pen-based gesture applications: State-of-the-art systems and future research directions. *Human–Computer Interaction*, *15*, 263–322.

Parra, L., Christoforou, C., Gerson, A., Dyrholm, M., Luo, A., Wagner, M., . . . Sajda, P. (2008). Spatiotemporal linear decoding of brain state. *IEEE Signal Processing Magazine*, *25*(1), 107–115.

Pfurtscheller, G., Allison, B. Z., Brunner, C., Bauernfeind, G., Solis Escalante, T., Scherer, R., . . . Birbaumer, N. (2010). The hybrid BCI. *Frontiers in Neuroprosthetics*, *4*, 30.

Pfurtscheller, G., Brunner, C., Schlögl, A., & Lopes da Silva, F. H. (2006). Mu rhythm (de)synchronization and EEG single-trial classification of different motor imagery tasks. *NeuroImage*, *31*(1), 153–159.

Pfurtscheller, G., & Lopes da Silva, F. H. (1999). Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clinical Neurophysiology*, *110*, 1842–1857.

Plass-Oude Bos, D., Reuderink, B., van de Laar, B., Gürkök, H., Mühl, C., Poel, M., . . . Heylen, D. (2010). Brain-computer interfacing and games. In A. Nijholt & D. Tan (Eds.), *Brain–computer interfaces: Applying our minds to human–computer interaction* (pp. 149–178). London, UK: Springer-Verlag London.

Qian, M., Aguilar, M., Zachery, K., Privitera, C., Klein, S., Carney, T., & Nolte, L. (2009). Decision-level fusion of EEG and pupil features for single-trial visual detection analysis. *IEEE Transactions on Biomedical Engineering*, *56*, 1929–1937.

Regan, D. (1977). Steady-state evoked potentials. *Journal of the Optical Society of America*, *67*, 1475–1489.

Salvaris, M., & Sepulveda, F. (2009). Visual modifications on the P300 speller BCI paradigm (Article 046011). *Journal of Neural Engineering*, *6*(4).

Savran, A., Ciftci, K., Chanel, G., Mota, J. C., Viet, L. H., Sankur, B., . . . Rombaut, M. (2006). Emotion detection in the loop from brain signals and facial images. In *Proceedings of the 2nd International Summer Workshop on Multimodal Interfaces* (pp. 69–80). Zagreb, Croatia: Faculty of Electrical Engineering and Computing, University of Zagreb.

Schalk, G., Wolpaw, J. R., McFarland, D. J., & Pfurtscheller, G. (2000). EEG-based communication: presence of an error potential. *Clinical Neurophysiology*, *111*, 2138–2144.

Scherer, R., Schloegl, A., Lee, F., Bischof, H., Janša, J., & Pfurtscheller, G. (2007). The self-paced Graz brain–computer interface: Methods and applications (Article 79826). *Computational Intelligence and Neuroscience*, *2007*.

Schnakers, C., Perrin, F., Schabus, M., Hustinx, R., Majerus, S., Moonen, G., . . . Laureys, S. (2009). Detecting consciousness in a total locked-in syndrome: An active eventrelated paradigm. *Neurocase*, *15*, 271–277.

Sellers, E. W., & Donchin, E. (2006). A P300-based brain-computer interface: Initial tests by ALS patients. *Clinical Neurophysiology*, *117*, 538–548.

Serby, H., Yom-Tov, E., & Inbar, G. (2005). An improved P300-based brain-computer interface. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *13*(1), 89–98.

Sharma, R., Pavlovic, V., & Huang, T. (1998). Toward multimodal human–computer interface. *Proceedings of the IEEE*, *86*, 853–869.

Shibasaki, H., & Hallett, M. (2006). What is the Bereitschaftspotential? *Clinical Neurophysiology*, *117*, 2341–2356.

Solovey, E. T., Girouard, A., Chauncey, K., Hirshfield, L. M., Sassaroli, A., Zheng, F., . . . Jacob, R. J. (2009). Using fNIRS brain sensing in realistic HCI settings: Experiments and guidelines. In *Proceedings of the 22nd Annual ACM Symposium on User Interface Software and Technology* (pp. 157–166). New York, NY: ACM.

Stanford, T. R., & Stein, B. E. (2007). Superadditivity in multisensory integration: Putting the computation in context (Article 787792). *NeuroReport*, *18*(8).

Tangermann, M., Krauledat, M., Grzeska, K., Sagebaum, M., Blankertz, B., Vidaurre, C., & Müller, K.-R. (2008). Playing pinball with non-invasive BCI. In *Advances in neural information processing systems 21* (pp. 1641–1648). Cambridge, MA: MIT Press.

Tolba, A. S., El-Baz, A. H., & El-Harby, A. A. (2006). Face recognition: A literature review. *International Journal of Signal Processing*, *2*, 88–103.

Turk, M., & Robertson, G. (2000). Perceptual user interfaces [Introduction]. *Communications of the ACM*, *43*(3), 32–34.

van Gerven, M., Farquhar, J., Schaefer, R., Vlek, R., Geuze, J., Nijholt, A., . . . Desain, P. (2009). The brain–computer interface cycle (Article 041001). *Journal of Neural Engineering*, *6*.

van de Laar, B., Reuderink, B., Plass-Oude Bos, D., & Heylen, D. (2010). Evaluating user experience of actual and imagined movement in BCI gaming. *International Journal of Gaming and Computer–Mediated Simulations*, *2*(4), 33–47.

Wang, Y.-T., Wang, Y., & Jung, T.-P. (2011). A cell–phone–based brain–computer interface for communication in daily life (Article 025018). *Journal of Neural Engineering*, *8*(2).

Wechsung, I., & Naumann, A. (2009). Evaluating a multimodal remote control: The interplay between user experience and usability. In *2009 International Workshop on Quality of Multimedia Experience* (pp. 19–22). Piscataway, NJ: IEEE.

Wolpaw, J., McFarland, D., Vaughan, T., & Schalk, G. (2003). The Wadsworth Center brain–computer interface (BCI) research and development program. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *11*(2), 1–4.

Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., & Vaughan, T. M. (2002). Brain–computer interfaces for communication and control. *Clinical Neurophysiology*, *113*, 767–791.

Zander, T. O., Gaertner, M., Kothe, C., & Vilimek, R. (2011). Combining eye gaze input with a brain–computer interface for touchless human–computer interaction. *International Journal of Human–Computer Interaction*, *27*(1), 38–51.

Zander, T. O., Kothe, C., Welke, S., & Roetting, M. (2009). Utilizing secondary input from passive brain–computer interfaces for enhancing human–machine interaction. In R. Goebel, J. Siekmann, & W.Wahlster (Eds.), *Foundations of augmented cognition: Neuroergonomics and operational neuroscience* (pp. 759–771). Berlin, Germany: Springer-Verlag.

Zeng, Z., Pantic, M., & Huang, T. S. (2009). Emotion recognition based on multimodal information. In J. Tao, & T. Tan (Eds.), *Affective information processing* (pp. 241–265). London, UK: Springer-Verlag.

ABOUT THE AUTHORS

Hayrettin Gürkök is a research assistant at the University of Twente, the Netherlands. He received his B.S. and M.S. in computer engineering from Bilkent University, Turkey. His research interests include user experience design and evaluation, natural and multimodal interaction, and entertainment computing.

Anton Nijholt received his Ph.D. from the Vrije Universiteit, Amsterdam. He held positions at universities in Europe and Canada and was scientific advisor of Philips Research Europe. Presently he is professor at the University of Twente. His research interests are multimodal and multi-party interaction, brain-computer interfacing and entertainment computing.

APPENDIX A

Nomenclature

aBCI	Active BCI
BCI	Brain–computer interface
CLiS	Complete (or total) locked-in syndrome
ECG	Electrocardiography (also known as EKG)
ECoG	Electrocorticography
EEG	Electroencephalography
EMG	Electromyography
EOG	Electrooculography
ERD/ERS	Event-related desynchronisation/ synchronisation
ERP	Event-related potential
ErrP	Error-related potential
fMRI	Functional magnetic resonance imaging
fNIRS	Functional near-infrared spectroscopy
GSR	Galvanic skin response
hBCI	Hybrid BCI
HCI	Human–computer interaction
HHI	Human–human interaction
HRV	Heart rate variability
LiS	Locked-in syndrome
MEA	Microelectrode (or multielectrode) array
MEG	Magnetoencephalography
MMI	Multimodal interaction
MUI	Multimodal user interface
NIRS	Near infrared spectroscopy
pBCI	Passive BCI
rBCI	Reactive BCI
RP	Readiness potential (also known as the Bereitschaftspotential)
SSVEP	Steady-state visual evoked potentials
VE	Virtual environment