Cognitive Rehabilitation with Realistic and Adaptive Computerized Cognitive Training: A Selective Interdisciplinary Review

Daniel Sjölie

UMINF-13.03
ISSN-0348-0542

Umeå University
Department of Computing Science
SE-901 87 Umeå, Sweden
Cognitive Rehabilitation with Realistic and Adaptive Computerized Cognitive Training: A Selective Interdisciplinary Review

Daniel Sjölie
Department of Computing Science,
Umeå University,
Umeå, Sweden
daniel@cs.umu.se

Abstract

Computerized cognitive training has the potential to be an important tool for both rehabilitation and prevention of cognitive impairment. However, the demonstrated effects of cognitive training have often been limited, and seemingly conflicting results have been reported. This paper reviews selected results from several disciplines that relate to the challenges and the potential of computerized cognitive training. Developments within cognitive neuroscience and human-computer interaction (HCI) are related to virtual rehabilitation, and to a need to develop training applications that are both adaptive and realistic.

Efficient rehabilitation and training depends on an ability to monitor the development of targeted functions. Monitoring cognitive functions is particularly challenging because of limited access to, and understanding of, these functions. Psychophysiological computing and brain-computer interfaces (BCIs) are reviewed as promising approaches for monitoring cognitive functions. Interpretation and understanding of such measurements should be supported by a basis in cognitive neuroscience and theories of brain function. This paper focuses on some general themes in such theories, with a strong connection to the importance of realism and interaction with reality. Such theoretical results are related back to practical applications through brain imaging studies concerning realistic interaction.

To summarize the interdisciplinary material reviewed here the concept of reality-based brain-computer interaction (RBBCI) is introduced. RBBCI is intended to capture important principles for the development of realistic and adaptive systems for cognitive rehabilitation, and to serve as a cornerstone concept for interdisciplinary development.
I. INTRODUCTION

Computerized cognitive training has been championed as a method with great potential for rehabilitation of cognitive deficits and for countering cognitive decline (Torkel Klingberg, 2010; Li et al., 2008). However, the desired general cognitive improvements, i.e., transfer effects, have been shown to be very difficult to achieve with common applications for cognitive training (Owen et al., 2010). This paper reviews a selection of previous work from several disciplines with a focus on how an improved understanding of cognition and brain function can be used to meet this challenge. Applications that respect users’ relation to everyday reality, before, during, and after training, is of key importance, as is the use of adaptive interaction in order to optimize training parameters dynamically. The importance of realistic interaction and the ecological validity of training are fundamental motivations for the use of virtual reality (VR) for rehabilitation and training (Pugnetti et al., 1995; Rizzo et al., 2001). The concept of reality-based brain-computer interaction (RBBCI) can be used to summarize these aspects and relate them to cognition and theories of brain function, thus facilitating the development of applications that operate in well-designed interaction with the brain.

The difficulty to gain general improvements from cognitive training is commonly described in terms of a distinction between near-transfer and far-transfer. Training almost always leads to improvements on the task that is actually trained, and often to improvements on very similar tasks (near-transfer). The challenge is to achieve transfer to tasks with no direct relation to the trained task (far-transfer) (Dahlin, Neely, Larsson, Backman, & Nyberg, 2008; Li et al., 2008). Much recent work on cognitive training has been focused on attaining general improvements related to, e.g., attention or working memory. Such general improvements should result in improved performance in everyday tasks, such as remembering what to shop. Unfortunately this corresponds to far-transfer, which is hard to achieve. The difficulty of achieving far-transfer motivates an increased focus on near-transfer, and it illustrates the need to train the right thing. This calls for an understanding of the differences between near and far transfer, and points to the need for interactive systems with high ecological validity. Reality-based interaction (RBI) in general, and virtual reality (VR) in particular, provides a foundation for ecologically valid HCI by building on the user’s skills and experiences from reality (Jacob et al., 2008; Rizzo, Schultheis, Kerns, & Mateer, 2004).

One challenge with ecologically valid cognitive training is how to adapt the intensity of the training. If the problem is to remember what to buy at the store a complete simulation of going to the store and buying a number of items might aid transfer to the real situation. However, the complexity of the task and the obscurity of the underlying functions can make it hard to monitor. Optimizing the level of training has been shown to be critically important (Ericsson & Charness, 1994; Ericsson, Prietula, & Cokely, 2007). This requires detection of what is challenging, and deliberate adjustment of the interaction environment in order to control the
challenging aspects. In general, this is a hard task in complex realistic training scenarios. The use of real-time brain measurements, as a form of passive, or adaptive, brain-computer interface (BCI) (Cutrell & Tan, 2007; Audrey Girouard, 2009), is a promising approach for solving this problem.

The further development of such systems, in particular the implementation of adaptations and the interpretation of measurements, should be supported by an understanding of cognition and brain function. In particular, it is helpful to consider the role that internalization of real-world phenomena serves in the development of cognitive capabilities (Kaptelinin & Nardi, 2006; Leontiev, 1978). Several recent theories of brain function present plausible neural implementations that line up well with this perspective (K. Friston, 2010; K. J. Friston, 2005; George & Hawkins, 2009; Hawkins, 2005; Gallese & Lakoff, 2005; Gallese & Goldman, 1998). Brain activity is described as largely related to predictions, based on past interactions with reality, and familiarity with real objects and phenomena. The critical role of errors in relation to such predictions is highlighted, pointing to the importance of the “difficulty” of the required predictions. This in turn depends on the general complexity of the phenomenon, and on the user’s familiarity with it, suggesting familiarity and complexity as two particularly interesting factors for adaptations and interpretation. The fact that several of these theories can be related to computational neuroscience and methods for analysis of brain measurements makes them particularly interesting for interpretation of brain measurements for use in computer applications.

II. COGNITIVE TRAINING

The potential effectiveness of cognitive training has been under investigation for a long time. Today the basic cognitive and neural plasticity of the brain is well established, motivating the use of cognitive training for humans of all ages, in general (Dahlin et al., 2008; Erickson et al., 2007; Jones et al., 2006; Li et al., 2008). The limits on what is possible, however, and what is needed to optimize the efficiency of cognitive training, remain unclear.

One form of cognitive training that has attracted much attention is working memory (WM) training. Working memory capacity predicts performance in a wide range of cognitive tasks, and many neuropsychiatric conditions such as stroke or attention-deficit hyperactivity disorder (ADHD) coincide with impaired WM (Conway, Kane, & Engle, 2003; Torkel Klingberg, 2010). Several studies have shown that performance on specific WM tasks such as 2-back (comparing the last item in a sequence to the one presented 2 steps before) does improve with training, and that this effect does transfer to similar (near-transfer) tasks with associated changes in brain activity (Dahlin et al., 2008; Torkel Klingberg, 2010; Li et al., 2008; Olesen, Westerberg, & Klingberg, 2004; Owen et al., 2010). However, the magnitude and range of transfer, in particular the potential for far-transfer, remains disputed. For example, studies comparing transfer effects in old and young adults have presented seemingly conflicting
results. A study by Dahlin et al. concluded that while transfer to untrained tasks is possible for both young and old, the magnitude varies, and it is often harder to demonstrate transfer in old adults (Dahlin et al., 2008). In other studies transfer effects in young and old have been compared without any reliable differences (Li et al., 2008). Suggested reasons for such differences in results include variations in the amount and intensity of the training, as well as differences in the degree of overlap between trained tasks and the evaluated transfer task.

In a recent study by Owen et al. 11,430 participants trained on cognitive tasks online for several weeks but failed to show any general cognitive improvements outside of the tasks that were actually trained (Owen et al., 2010). How can this be explained given the previously demonstrated potential for cognitive plasticity and transfer? Can faith in the potential of cognitive training be maintained? The first thing to consider is that the primary goal of the study in question was to investigate potential general cognitive improvements. Even though the results include remarks about a lack of transfer even between relatively similar tasks, the potential for near-transfer to similar tasks is not developed. This motivates a closer look at near-transfer, and a focus on how to achieve the necessary overlap and similarity between tasks, and between the corresponding brain functions. The use of realistic interaction and VR technologies in computer applications for cognitive training is an important move towards increased overlap between training and desired improvements.

Another possible reason for the lack of transfer in the study by Owen et al. is that the amount or intensity of the training might simply have been insufficient. The amount of training is addressed in their paper by pointing out that the average number of training sessions in the study (25) should be enough to see a measurable effect if there was one, and by noting that differences in amount of training between participants did not correspond to similar differences in transfer effects. However, it is acknowledged that it cannot be ruled out that more training may give results. The large variation in amount of training between subjects, with a standard deviation of approximately 17 training sessions with a mean of 25, also leaves some room for further questions. These between-subject differences are not further analyzed in the paper and may conceal important differences, e.g., in terms of individual motivation or discipline and related changes in training intensity.

A powerful argument for the critical importance of both the amount and the intensity of training can be found in research into the nature of expertise (Ericsson & Charness, 1994; Ericsson et al., 2007). In short, it has been shown that what is needed to become truly skilled is a large amount of training at a deliberately directed and adapted level of intensity and difficulty. Humans are not born to become chess masters or elite musicians but “experts are always made, not born” (Ericsson et al., 2007). Deliberate practice must be directed to a level where the training in question includes elements that one is not already skilled with, while at the same time building on elements that one is familiar with. In essence, one needs to make some errors in order to have something to correct and improve, but too many errors will
hamper learning. An example of how deliberate adaptation of the training can be successfully employed together with cognitive training is the use of WM training to reduce symptoms of ADHD in children. The methods behind the training program offered commercially by Cogmed ("Cogmed Working Memory Training," 2011) for this purpose has been evaluated by Klingberg et al. in several studies (T. Klingberg, Forssberg, & Westerberg, 2002; Torkel Klingberg et al., 2005). This program includes both continuous adaptation of the difficulty based on training performance, and regular contacts with a human “coach” to further guide and motivate the training.

A similar focus on the importance of training and learning at the right level can be found in the idea of the zone of proximal development (ZPD) (Cole, 1985; Kaptelinin & Nardi, 2006). This concept is of particular interest in the current context because it was pioneered by Lev Vygotsky, who is also recognized as one of the founding fathers of activity theory. Activity theory is a theoretical framework that has gained some popularity within the HCI community in the last decades, in large part because of its focus on the context of interaction and human activities in the real world (Kaptelinin & Nardi, 2006; Kuutti, 1996; Bonnie A. Nardi, 1996). The zone of proximal development was conceived as a way to assess development and learning, in particular in children. Measures of development that were based on current performance failed to predict how a child would develop in the future. Instead, Vygotsky suggests that a measure of the difference between what a child can accomplish on its own and what it can accomplish with the aid of an adult is a far better guide to the current developmental potential of the child. It is this difference that has become known as the zone of proximal development. Similar ideas can be used to guide adaptations of computer aided cognitive training by varying the support given and thus probe the developmental potential of the user in the particular context.

It should be noted that it was important for Vygotsky that development was facilitated by social interactions. Aid from other humans was crucial for the development of cognitive functions. One of the many advantages that humans have over computers in this context is their expertise at judging the cognitive state of other humans. In order to reach a point where computers can step in and aid human cognitive development in a similar manner it is essential to extend the communication between human and computer beyond the mouse and keyboard.

III. HUMAN-COMPUTER INTERACTION

The field of human-computer interaction (HCI) continues to expand as computers become part of both everyday society and personal life. Several areas of HCI research are of great interest for the development of realistic and adaptive computerized cognitive training.
A. Adaptive Psychophysiological Computing

One recent development within HCI is an increasing interest in using measurements from the brain or body, i.e., physiological measurements, as extra input channels for computer applications (Fairclough, 2009; Tan & Nijholt, 2010a). In particular, these measurements can be related to the psychological state (thus the term psycho-physiological computing) of the user, and used to adapt the behavior of the application to psychological states such as frustration, overload, or excitement (Daly & Wolpaw, 2008; Picard, 2000; Tan & Nijholt, 2010a; Zander, Kothe, Jatzev, & Gaertner, 2010). This provides a basis for computer applications that may start to take on the role of a human coach that guides training based on the state of the user. Common physiological measurements for this use include heart rate (using electrocardiography, ECG), skin conductance (primarily as galvanic skin response, GSR, or skin conductance reactions, SCR), respiratory rate (breathing), skin temperature and electroencephalography (EEG). Some examples of how these measurements can be used are given below.

Wilson and Russel used EEG and ECG together with electrooculography (EOG, used to measure electrical activity resulting from eye movements) to adapt the amount of support given to an operator in a complex aviation task (Wilson & Russell, 2007). Features based on these measurements were fed into an artificial neural network (ANN) that was trained to classify the measurements as corresponding to easy or hard conditions in the task. It was shown that assistance given on the basis of the classified operator state significantly increased performance. A similar approach to classification was used by Koenig et al. to estimate the mental engagement of patients and nondisabled users during robot-assisted gait training (Koenig et al., 2011). Their study measured ECG, breathing, GSR and skin temperature, and they demonstrated a high accuracy of classification when compared to subjective measures. Skin temperature and skin conductance were pointed to as the most reliable psychophysiological responders across both patients and nondisabled. This result partly matches an earlier result by Novak et al. (2010). This study investigated the same measurements, in a similar motor rehabilitation task, and they found that skin conductance was the most reliable measurement, while skin temperature provided better results for the control group than for stroke patients.

All of the studies above make some use of subjective measurements. Psychophysiological measurements are often combined with subjective measures, if for no other reason than to validate the psychophysiological measurements and classification results. Popular alternatives are the NASA Task Load Index (NASA TLX) (Hart & Staveland, 1988), the Self-Assessment Manikin (SAM) (Bradley & Lang, 1994), and subjective workload assessment technique (SWAT) (Reid & Nygren, 1988). NASA TLX and SWAT are typically administered after completing a task since these are relatively comprehensive questionnaires. The SAM is a faster alternative. The original three dimensions of SAM, arousal, valence and control, are
often reduced to two questions, dropping control (Koenig et al., 2011). Although the SAM explicitly targets emotions these can often be related to cognitive states such as mental workload. E.g., if you are unhappy and aroused you are stressed, if you are happy and aroused you are engaged, if you are unhappy and not aroused you are bored, etc.

The alternative to relying on classification and subjective measures is to directly use some form of index to adapt an application, and take the evaluation to be whether or not it actually “works”, and, e.g., improves performance. One of the purest examples of this approach can be seen in a recent study by Walter (2010). In this study subliminal feedback was used to change the interface or environment of the user directly depending on measured SCR, based on the strong connection between skin conductance and arousal (Critchley, Elliott, Mathias, & Dolan, 2000). Factors such as background color, text size and ambient room lighting were subliminally adapted in small steps, unnoticed by the user. These adaptations slowly moved the application towards a configuration producing increased arousal. The interaction environment resulting from a series of such adaptations was shown to lead to significantly improved performance in the investigated tasks. The potential for subliminal adaptations in VR environments is indeed great and the level of performance is of prime interest for cognitive training applications.

The use of adaptive systems and psychophysiological measurements in combination with VR is not new. Recent examples are a system for cognitive telerehabilitation, capable of automatically adjusting the difficulty level of tasks and even change tasks depending on performance (Tost et al., 2009), and a system for motor rehabilitation incorporating psychophysiological signals to adapt to emotions (Matjaz Mihelj, Novak, & Munih, 2009). The recent development of such systems demonstrates that the use of measurements from brain and body to adapt virtual environments (VEs) for efficient interaction, training and rehabilitation is becoming increasingly feasible and common. The use of brain measurements in such systems is, however, still uncommon.

B. Brain-Computer Interfaces

Brain measurements provide the potential for a direct connection to psychological and cognitive state that may be very valuable for applications targeting cognitive training. The integration of brain measurements into computer applications has traditionally been in the form of brain-computer interfaces (BCIs) enabling the user to consciously control an application. See Tan & Nijholt (2010a) for an introduction to BCI from a HCI perspective and Lécuyer et al. (2008) for comments on BCI in combination with VR and videogames. The use of similar BCI methods for the passive adaptation of an application has been suggested (Cutrell & Tan, 2007; Audrey Girouard, 2009; Zander et al., 2010) and this is a growing research area. Millán et al. recently reviewed the potential for combining BCIs with assistive technologies and touch upon the use of BCI for adaptive applications (Millán et al., 2010). The
use of a BCI as an additional channel is described as a hybrid BCIs, but is not related further to previous work on passive and adaptive BCIs.

The techniques most commonly used for BCIs are EEG and functional near-infrared spectroscopy (fNIRS). More details about these techniques are given in section V.B. EEG is the most common and most thoroughly investigated method. Grimes et al. demonstrated that the accuracy of EEG-based classification of mental workload scales gracefully with the amount of training data, the time window, etc (Grimes, Tan, Hudson, Shenoy, & Rao, 2008). By using a method for automatically selecting features for training based on information gain between conditions, in relation to information gain between continuous blocks, they show that EEG can be used to automatically classify mental workload with high accuracy. Solovey et al. have investigated the use of fNIRS to detect different states of multitasking (Solovey et al., 2011). Based on previous results using functional magnetic resonance imaging (fMRI) they demonstrated that the same patterns could indeed be identified using real-time fNIRS. This provides a basis for adaptive BCI applications that change in response to the multitasking state of the user.

The studies above are only a few selected examples from areas of research that are currently developing quickly. However, such applications are still rare and the space of possible adaptations is relatively unexplored. Further development should benefit from considerations of recent theoretical developments in HCI and cognitive neuroscience.

C. Reality-Based Interaction

As the interest in more complex and realistic interaction methods grows within HCI the overlap between VR research and HCI research increases. The framework of reality-based interaction (RBI) has recently been introduced as an attempt to capture the underlying advantages (and disadvantages) of designing interaction with computers to be similar to interaction with physical reality (Jacob et al., 2008). The framework of RBI relates realistic interaction to human awareness of and skill with body, environment and social situation as well as naïve human understanding of physics. These themes are becoming increasingly common in emerging HCI applications, e.g., in the form of tangible interfaces building on our naïve understanding of physics, friction and gravity (Jacob et al., 2008). The connection to VR is especially apparent for the theme of environmental skill and awareness, exemplified by the human proficiency with spatial navigation and our ability to keep track of objects and events at different spatial locations. This can be further related to the sense of presence. Presence has traditionally been described as the sense of “being there” in a virtual environment (Slater, 2002) but recent elaborations of this description are easier to relate to brain function. Such accounts include a greater emphasis on presence as hypothesis selection (Sanchez-Vives & Slater, 2005) and as “the ability to act there” (Jäncke, 2009). The conception of presence as the ability to act in a given environment has been directly related to brain function and neural
correlates of presence as it relates to the use of existing motor representations to interact in a virtual reality (Jäncke, 2009). This is also described as the ability to relate the virtual space to “real motor space” and to build on simulations of real motor responses. Such simulations are based on past interactions with reality and together with the description of presence as hypothesis selection this maps very well into recent theories of brain function that may be helpful to consider when developing applications to take advantage of brain measurements.

IV. THEORIES OF BRAIN FUNCTION AND INTERACTION

The description of presence as the selection of a hypothesis corresponds directly to the establishment of expectations, based on predictions about how interaction with the assumed reality should and/or could proceed, according to the hypothesis in question. In these terms, breaks in presence, suggested as a fundamental complement to presence (Slater, 2002), correspond to errors in relation to predictions that are based on beliefs about the environment. The importance of such predictions and the associated prediction errors is a central theme in a number of recent theories of cognition and brain function. It should be noted here that these predictions should not be equated with conscious predictions. It may just as well be an unconscious expectation about what to feel when touching a door handle, etc. Also, errors in prediction may be considered as matters of degree rather than match or no match. If predictions are present on many levels of the brain a complete match can rarely (if ever) be expected.

One emerging theme in recent theories of cognition is the importance of mental simulations. Within the general framework of grounded cognition mental simulation is presented as a fundamental aspect of cognition, together with the grounding of higher cognition in the modal systems of the brain (Barsalou, 2008; Barsalou, Simmons, Barbey, & Wilson, 2003; Gallese & Lakoff, 2005). The modal systems in the brain correspond to the low level interfaces between the brain and “the rest of the world”, i.e., the lowest levels in a hierarchy. This includes the common perceptual senses but also action, body information and (according to some accounts) introspection (Barsalou, 2008). The grounding of higher-level cognition in such lower-level modalities corresponds to a hierarchical description of cognition that matches the hierarchical structure of the brain, in particular the cortex (K. Friston, 2005).

Mental simulation builds on an extension of results concerning mental imagery: a subject that has gathered a lot of research over the last decades, e.g., concerning mental rotation (Cohen et al., 1996; Shepard & Metzler, 1971) and motor imagery (Jeannerod, 1995). The concept of mental simulations explicitly includes the unconscious and flexible reactivation of memories, employed to recognize the current context and to simulate, or predict, possible actions and expected results. This idea, that predicting future events based on previous experience is a critical aspect of how the brain works, has gathered increasing support in recent years. The renowned memory-researcher Daniel Schacter recently argued for such a perspective as “helpful” for understanding the brain (Schacter, Addis, & Buckner, 2007, p. 660) and it is
prominent in recent theories of cognition and the brain by Hawkins (George & Hawkins, 2009; Hawkins, 2005), Friston (K. Friston, 2005, 2010), and others (Bar, 2007; James M Kilner, Friston, & Frith, 2007; R. P. N Rao & Ballard, 1999). These theories further develop the importance of hierarchies and memories (past experience), and the critical importance of joint representation of action and perception for predictions. “What can I expect to perceive/experience if I act thus? How do I need to act in order to experience that?” This corresponds to representations of interaction and, in the common case, interaction with reality.

Several of the theories introduced above make specific claims about what is expected to give rise to increased brain activity (Friston, 2005; George & Hawkins, 2009). The critical role of prediction error in these models is prominent. According to this view, the brain works by predicting what comes next and if the prediction is correct no further reaction is needed. It is only when the prediction fails that one must reevaluate the situation and consider alternative interpretations of our current environment. Objects and phenomena that are easy to predict give rise to few (and/or small) prediction errors and result in less brain activity compared to phenomena that are more unpredictable. Predictions are made in relation to a context. The brain consists of a complex hierarchy where the context is defined at a higher level and predictions are checked at a lower level. When prediction errors occur they are passed upwards, and assumptions about the context are refined as needed. Such theories can be used for model-based analysis and interpretation of brain measurements. Some readers may also find the connection to information content and compression interesting. Higher level predictions correspond to the larger and/or more general trends in “the data”, i.e., gathered experience, and lower level prediction errors correspond to the new information and the details in the current situation.

A recurring theme in the theories above is the importance of how hierarchical representations of observed phenomena develops (Friston, 2010; Friston, 2005; George & Hawkins, 2009; Hawkins, 2005). Learning and skills development is explained as necessarily based on what you already know, utilizing existing levels in the hierarchy to compress representations and to enable better predictions of events. The idea that critical aspects of cognition rely on our ability to internalize, represent and in some sense simulate events and processes that are observed in the environment has been defended by many researchers in cognition (Barsalou, 2008; Gallese & Lakoff, 2005; Hutchins, 1996; Kaptelinin & Nardi, 2006). Research on mirror neurons has brought forward some particularly clear examples of how mental simulations may be implemented in the brain. The primary result is that parts of the brain are activated both when executing actions and when observing someone else doing the same action (Fadiga, Fogassi, Pavesi, & Rizzolatti, 1995; Gallese & Goldman, 1998). This has been extended to include the imaging of future actions on the same neural basis (Gallese & Goldman, 1998; Kilner et al., 2007) which in turn in tightly related to remembering similar actions from the past (Schacter et al., 2007), providing a clear connection to the development
of mental simulations. The suggested importance of such internalized phenomena is in large part related to the idea of cognitive tools. I.e., the idea that cultural constructs such as language and mathematics support human cognition as they are internalized and become available as building blocks in mental simulations that facilitate interaction with reality (Kaptelinin & Nardi, 2006; Leontiev, 1978). In light of such ideas cognitive training can be imagined as an attempt to expose the user to the cognitive tools that he or she needs to internalize. E.g., mental techniques (not necessarily conscious) for keeping track of a shopping list in a virtual scenario. A critical question when developing systems for cognitive training then becomes how to adapt the created environments to support the desired internalization. New or changed simulations must fit into the existing structure (be familiar enough) and they must give rise to prediction errors (in order to elicit corrections and change).

V. BRAIN MEASUREMENTS AND REALISTIC INTERACTION

A. Whole brain imaging

A number of studies have been conducted on brain activity in realistic interaction environments employing VR technologies. Brain activity related to navigation has been studied several times using virtual 3d-environments and functional magnetic-resonance imaging (fMRI). FMRI measures brain activity by detecting changes in blood flow that are assumed to depend on the metabolic demands of neural activity. See Mráz et al. (2003) for a general discussion of the combination of fMRI and VR. In a study by Aguirre et al. (1996) fMRI was used to investigate topographical learning and recall while navigating a simple 3d-maze. Spiers and Maguire (2006, 2007) have extended upon this basic combination of a 3d-environment and fMRI in several important ways. They made sure that the virtual environment (VE) was realistic and full of life by taking advantage of an existing commercial game where the user was able to drive a car in the middle of the busy London traffic. In order to determine the thought process corresponding to a task executed in the VE a verbal report protocol (Ericsson & Simon, 1980) was used in conjunction with a video recording of the subject’s viewpoint in the VE. The video was shown to the subject directly following completion of the VE task, and the subject was asked to describe verbally what he/she had thought at the times shown on video.

When these studies look at the whole brain they generally show distributed networks of activity. The details of this network vary between studies, possibly because of differences in the task setup, but the hippocampus in particular plays a central role according to most reports. Unfortunately, the hippocampus is located deep within the brain and brain activity in the hippocampus cannot be measured with common BCI techniques such as EEG or fNIRS. Another common finding that may be more directly relevant for BCI applications is brain activity in frontal regions. In the Maguire and Spiers study brain activity in frontal regions was related to route planning or different forms of expectation violation (Spiers & Maguire, 2006).
This corresponds to cognitive states where your current hypothesis about your context is in flux. I.e., predictions at lower levels can no longer explain the current experience in terms of higher level goals and desires, and prediction errors propagate upwards through the hierarchy until they reach the frontal areas of the brain where they need to be resolved. In other words, at moments when there is no clear plan, goal, or belief, the mind is more open to external expressions and prediction errors that correspond to detailed information about the current situation can reach and influence higher-level cognition. The dorsolateral prefrontal cortex (DLPFC) in particular has been identified as a key region in networks related to presence (Jäncke, 2009) and spatial working memory (Constantinidis & Wang, 2004). The relation to working memory constitutes a direct connection to more general theories of brain function, such as the ones discussed above. The relation between spatial working memory and mental imagery constitutes a particularly clear example (Postma & Barsalou, 2009), building on such classical results as the speed of mental rotation (Shepard & Metzler, 1971).

We have previously conducted a study concerning the effect of certain aspects of a realistic and dynamic interaction environment on brain activity measurements using fMRI (Sjölie et al., 2010). Brain activity was measured while performing a mental rotation task in a virtual environment with varying degrees of motion and interactivity. Our results show that interactivity leads to increased activity in frontal and medial areas while the effect of automatic, easily predicted, motion is restricted to posterior, primarily visual, areas (Fig. 1). These increases are primarily within areas already activated for the general mental rotation task and not (simply) explained by the sensorimotor activity associated with interactivity. Within the perspective given by the theories described above the increased activity from interactivity can be understood as related to an increased unpredictability in the environment.
The more frontal nature of the effect of interactivity also fits well with the hierarchical structure of predictions and prediction errors in these theories. Environments that are more dynamic and harder to predict lead to more prediction errors being fed upwards, and to increased activity in higher-level, more frontal, regions.

B. Understanding Practical Measurements

While fMRI measurements have great advantages for investigating brain function there are many disadvantages when attempting to develop systems for practical use and wide distribution. Functional near-infrared spectroscopy (fNIRS) is one method that is starting to show some promise (Solovey et al., 2009). One advantage with fNIRS is that the measurements are of the same kind as fMRI measurements, i.e., they are based on blood flow and neural metabolism, making comparisons to results from the vast fMRI literature relatively straightforward. However, the dominant method of brain measurements for BCI is electroencephalography (EEG) and recent developments speak to the future potential of this method. The emergence of commercially available and affordable EEG-headsets, such as the Emotiv Epoc (Emotiv Corporate, 2011), lowers the threshold for new researchers and developers to integrate BCI features into their systems. This development plays right into the increasing interest for psychophysiological computing among HCI researchers and the desire to tap into the human mind to extend the HCI toolkit (Tan & Nijholt, 2010b). An increasing interest in combining fMRI and EEG and in developing the relation between these different forms of brain measurement also speaks to the feasibility of relating theories of brain function to brain measurements and applications utilizing EEG (Mulert & Lemieux, 2010).

An understanding of EEG in the context of previous fMRI results and theories of brain function is perhaps best supported by the development of neuronal models that aim to explain both fMRI and EEG measurements as resulting from the same underlying neuronal activity. Such models are still in their infancy and details should be expected to change in the future but the overall picture presented has a solid basis in empirical evidence and illustrates how EEG can fit in with other results and theories (Kilner & Friston, 2010; Kilner, Mattout, Henson, & Friston, 2005). In short, the model describes neuronal populations as dynamic systems with neurons that fire at different rates and consume energy in relation to this. FMRI measurements are related to metabolism and thus to the energy consumption while EEG measurements are related to a combination of neuronal spiking frequencies giving a spectral profile where the power varies with the frequency. The key idea is that both the energy consumption and the spectral profile are affected as the dynamics of the systems speeds up or slows down. Thus an activation measured with fMRI should correspond to a speedup of the neuronal dynamics and thus to a shift in the spectral profile of corresponding EEG measurements towards higher frequencies. In both cases the underlying neuronal phenomenon is that the neurons fire more
frequently within a given population. In terms of standard EEG measurements this corresponds to a reduced power in the lower frequency bands such as the alpha band (8-12 Hz) and an increase in higher frequency bands such as the gamma band (30-70 Hz). This is largely in agreement with results on the relation between EEG, fMRI and working memory load (Michels et al., 2010) but previous experience of using EEG to classify working memory load in real time in HCI research has demonstrated large individual differences that complicate matters (Grimes et al., 2008).

VI. SUMMARY AND PRACTICAL APPLICATIONS

A. Reality-Based Brain-Computer Interaction

Given the interdisciplinary spread of the works reviewed in this paper, the connections between them may not be obvious. The development of systems for realistic and adaptive cognitive training may be supported with a summarizing concept. Reality-based brain-computer interaction (RBBCI) describes a system where the computer interacts directly with the brain, by building on the principles and techniques described above. The input to the brain consists of computer-generated phenomena in a virtual reality. The output from the brain consists of brain measurements that can be related to properties of these phenomena in an informed manner. When computer applications deliberately modify these phenomena expecting certain responses in brain measurements there is an interaction loop involving computer and brain. Note that the conscious user is not included in this description. RBBCI is thus closely related to the concepts of passive and/or adaptive BCI (Cutrell & Tan, 2007; Audrey Girouard, 2009; Zander et al., 2010), described above. The adaptation in a RBBCI system consists of changing aspects of the presented reality that can be related to cognitive processes and changes in brain measurements, based on theories of cognition and brain function.

To develop an RBBCI application it is necessary to integrate the use of VR techniques and adaptive BCIs, with an understanding of how brain activity is affected by VR in general and by possible adaptations of VR in particular. The primary motivations for considering RBBCI as a unified concept are:

- Brain measurements and the course of events in the computer-generated reality should be considered together as tightly interrelated through the user’s perception of reality.
  - Because cognition and the corresponding brain activity is linked to the user’s current reality in a fundamental way with explicit theoretical formulations and practical implications.

- Applications targeting cognitive rehabilitation or training need a combination of realistic interaction and real-time adaptations to cognitive state.
Because cognitive training based on specific tasks does not give reliable improvements in general cognitive functioning, and because training at the right level is always essential.

- Interdisciplinary development of VR applications can benefit from a cornerstone concept such as RBBCI that can be closely tied to all related areas.
  - Because an ability to clearly communicate and share ideas among researchers, developers and practitioners in interdisciplinary projects is both important and inherently difficult.

The essence of RBBCI is that brain function is intimately related to reality, and that the use of VR technology makes it possible to manipulate and synchronize the perceived reality and the associated brain function of the user. Thus, the computer interacts with the brain through the presented reality and by interpreting brain measurements resulting from aspects of and changes in this reality.

B. Informed Adaptations

One clear benefit with the theoretical grounding reviewed in this paper is as a basis for developing adaptations in brain aware applications. The theories presented above suggest a way to understand how aspects of a complex realistic task affect brain measurements. In particular, familiarity and basic complexity or predictability emerges as particularly promising parameters for adjustment. Optimal training should depend on both these parameters being at the right level. Increased familiarity and predictability should contribute to reduced brain activity and more unfamiliar, complex or unpredictable stimuli should contribute to increased brain activity. An unfamiliar or unpredictable environment, and an increased amount of prediction errors, can also be expected to lead to brain activity further up in the hierarchy, which generally corresponds to more frontally in the brain.

It is also important to realize that not all realism is created equal. The work reviewed in this paper provides some further guidance concerning which aspects of realism are important for applications that target cognitive functions. Reality and realism is tightly coupled to experience and familiarity. In order to know what is truly realistic for a specific user in this sense, the culture and everyday environment of the user is of critical importance.

VII. CONCLUSIONS

Computerized cognitive training has potential for rehabilitation of neurological disorders and for fighting off or compensating for cognitive decline. This paper presents a selective interdisciplinary review of some key areas for the further development of such applications. Developments within HCI and cognitive neuroscience, including reality-based interaction, adaptive BCIs and the neurological basis for effects of cognitive training, provide a basis for
the implementation of systems that take advantage of an increasing understanding of the brain. To further aid the integration of these interdisciplinary results the concept of reality-based brain-computer interaction (RBBCI) is suggested as a supporting cornerstone concept with connections to many related disciplines. The implicit argument is that the development of efficient computerized cognitive training requires that all of these results are considered together, e.g., in order to understand how one may change a realistic interaction environment to challenge the brain in the right way by balancing familiarity and unpredictability.

ACKNOWLEDGMENTS

Thanks to Lars-Erik Janlert and Johan Eriksson for comments and feedback on the paper. Supported by Swedish Brain Power, Umeå University and the Kempe foundations, JCK-2613.

REFERENCES


