Neurofeedback and Metacognition

Lars Hall & Petter Johansson
Lund University Cognitive Science,
Kungshuset, Lundagård, 222 22 Lund

Abstract. In this article we discuss the relationship between the concept of neurofeedback and research on metacognition. We analyze different aspects of neurofeedback, such as temporal and spatial resolution, range of possible cognitive modulation, long-time learning effects, ambulatory measurement, and advanced computer interfaces, and relate them to basic and applied research on metacognition. We argue that a refined and expanded practice of neurofeedback promises to be a helpful tool both to probe and investigate the basic nature of metacognitive processes, and as a powerful instrument to strategically shift the boundaries of our conscious control over a wide collection of cognitive phenomena.

Introduction

In the cognitive and clinical neurosciences the concept of executive functions (EF) has long served as an umbrella term for a loose collection of 'higher-order' abilities. It has been used to encompass actions of "planning, inhibiting responses, strategy development and use, flexible sequencing of action, maintenance of behavioral set, resistance to interference, and so forth" [1 p.3]. A more restricted classification comes from [2] who sorts five types of inhibitory control (orienting, preparing, filtering, suppressing and stopping) under the jurisdiction of a central 'supervisory system'. Finally, and perhaps the most common way to refer to this diverse set of processes, is as the notion of a 'central executive' [3].

No one doubts that the study of EF holds great potential for illuminating the related concept of *metacognition* as discussed within cognitive and educational psychology. For example, [4] have proposed the 'bridging' term *metacognitive regulation* to refer to various processes that coordinate cognition. In this category they include both bottom-up processes such as error detection and source monitoring in memory retrieval (cognitive monitoring), as well as top-down processes like conflict resolution, error correction, inhibitory control, planning, and resource allocation (cognitive control).

On the other hand, for someone with a chief interest in the development of computer-centered tools and environments, to scaffold and augment human metacognitive abilities, a deep chasm still exists between neuroscience research on EF, and the more applied focus of cognitive and educational psychologists. As [5] recently have argued, the cognitive neurosciences have generally failed to make good use of first-person, introspective data in their experiments and models. For EF-type theories that deal with processes frequently straddling the fence between the conscious and the subconscious, this has been particularly damaging, and lead to a proliferation of quasi-phenomenological concepts like 'monitoring' (of errors, or conflict), 'manipulation' (of attentional focus), 'maintenance' (of working memory contents) etc. Such concepts may well capture something important about the processes at hand, but from a first-person perspective they do not ring quite true [5]. Despite frequent use of terms like *strategic* or *controlled* in reference to different EF processes [6], it is often wholly unclear whether these are things that we *do*, or things that just *happen* in us. For a researcher with an interest in developing tools to support metacognitive processes this is a serious problem, as it makes it difficult to know which processes that are good candidates for support, and what kind of applications subjects are able to respond to.

In light of this, efforts to support metacognition with information and communication technology (ICT) have instead focused on features of metacognition that lend themselves more easily to external analysis. Since we possess detailed normative models of scientific reasoning (i.e. hypothesis testing, theory-evidence coordination, peer-review, etc. see [7]), we also have a good grasp of what an ideal metacognitive reasoning-process in this domain would look like. Applications that provide support for such activities as reflection and revision (i.e. things like structured argument-threads, or systems that demand formal giving and taking of reasons in simulation environments, etc.) quickly get students to start thinking about their own thinking, and learning about their own learning [8]. As reasoning to a large extent is a social process, there are also many examples of metacognitive learning platforms that do a good job of making visible the asking, judging, endorsing, revising, and clarifying that make up critical, communal discourse.

However, to move beyond the current state of affairs and engage *all* aspects of metacognition, we believe the mind/brain has to be brought back into the equation. In addition, we suspect this has to be done with a degree of specificity significantly greater than the previous generation of theories provided (i.e. future applications have to be targeted at a level well below such overarching categories as planning, monitoring, reflecting, etc.). As a modest contribution to this project, we will here discuss the concept of *neurofeedback* and its applications to metacognition: a concept that we believe promises to be helpful both as a basic device to probe and investigate the nature of EF and metacognitive activities, and as a powerful instrument to strategically shift the boundaries of our conscious control over a wide collection of cognitive phenomena.

1. Neurofeedback

*Neuro*feedback is a subclass of the general category of *bio*feedback, specifically concerned with the feedback and control over neural levels and patterns of activation (e.g. through the use of various properties of EEG, but more recently also by way of PET and fMRI). Over the course of the years, neurofeedback training has been applied in the most diverse domains of behavior, and has often been subject to exaggerated claims of beneficial effects [9, 10]. On the other hand it also represents a well-established and proven clinical practice for a few selected domains centered on different forms of attention and relaxation disorders (particularly attention deficit hyperactivity disorder, ADHD, see [11, 12]).

More recently, reliable effects have also been found in relation to normal cognitive performance. For example, in an experiment involving conditioned enhancement of a low beta band rhythm of the EEG (12-15Hz, also called sensorimotor rhythm; SMR) [13] found significant improvements of perceptual sensitivity and a reduction of commission errors in a continuous performance task (the well known TOVA, Test of Variables of Attention) in a sample of healthy volunteers. While the relationship between levels of SMR and its influence on cognitive performance is not yet fully understood, [13] hypothesize that learned voluntary control of SMR activity may facilitate information processing by decreasing interference from the motor system on other types of cognitive activity.

Positive results from neurofeedback training in healthy populations have also been reported for semantic working memory performance. In a study by [14] enhancement of SMR (conducted over a period of 4 weeks, with each participant receiving two training sessions a week) lead to clear performance gains in a cued-recall conceptual span task compared to a control group who received no training. As an explanation of the results [14] discuss the motor interference hypotheses described above, but they also suggest the possibility that regulation of activity in the 12-15Hz frequency range may directly influence encoding and retrieval operations in semantic memory.

Electrical responses recorded from the scalp can be used to measure events happening in the millisecond range. Thus, the temporal resolution of EEG makes it a suitable instrument for capturing the rapid dynamics of cognition over the whole cortical surface. It is within this tradition that most research on neurofeedback has been conducted – i.e. with a focus on generic processes of cerebral arousal and attention (e.g. [13, 12]). Because of the conduction distortions introduced by cerebrospinal fluid, the skull, and the scalp, EEG collected from any point on the scalp may include activity from a large volume of the brain. Such spatial smearing makes it difficult to relate EEG measurements to the underlying brain processes. However, the diffficulty inherent in source localization does not preclude the use of EEG for more focused neurofeedback applications. For example, feedback of slow-cortical potentials (SCP) recorded exclusively above left-hemisphere language cortices, has been used to induce speeded responses on a word comprehension task while no difference could be seen on a simple reaction time task [15]

Nevertheless, a more powerful solution for spatially constrained neurofeedback is to take advantage of the great precision inherent in modern equipment for functional magnetic resonance imaging (fMRI). One recent and interesting study using this technique involved subjects self-inducing an emotional reaction in correspondence to pictures of either happy, neutral or sad faces shown in the scanner [16, see also 17]. After each trial the subjects rated the intensity of their experienced emotion/mood. As a feedback manipulation the experimenters responded back with real-time images of activation level in the amygdala (a subcortical region known to play a vital role in emotional responses and processing) to guide the participants in their further regulatory attempts [16]. While the experiment was not tied to any particular behavioral or psychological goal, it clearly demonstrates the potential of fMRI as an instrument of neurofeedback (aside from greatly increased spatial resolution, and added bonus of fMRI neurofeedback compared to surface measurements like EEG, is the ability to measure activity in deep subcortical structures like the amygdala).

As real-time processing and feedback in an fMRI environment is a very computationally intensive operation, few such studies have been conducted to date (e.g. see discussion in [17]). However, this will soon change. With a slight sacrifice of analytic power [18] have demonstrated the feasibility of using (near) real-time feedback in an fMRI environment without the need for a customized network of data-transfer, and with software running on standard personal computers.

2. Neurofeedback and Metacognition

In the introduction we mentioned the problem of finding a meaningful relationship between introspection and the various EF processes that are conceived of as ‘controlled’ and/or ‘strategic’ in the cognitive and clinical neurosciences. As we see it, this is one of the most critical and difficult problems for applied cognitive and educational psychology to solve. We agree with [5] when they argue that this situation is in part caused by failure to properly incorporate and handle introspective evidence in basic research on EF and metacognition. Thus, to better align EF with what people *think* and *believe* they are doing in their minds, care has to be taken to bring the subject back into center stage of research.

On the other hand, it is a well known fact that introspection is a very blunt, and often downright misleading instrument [19] With research on EF proceeding at a breathtaking speed, it will sooner or later reach a level of explanation where the ‘subjective point of view’ will have to be dissolved into a host of mechanistic processes (as [20] argues, there is no ghost in the machine, in the end all homunculi have to be discharged). In this sense, we have to be prepared to let the subject completely vanish from sight.

Obviously, we cannot even begin to address this problem here. Instead, what we would like to suggest, is that neurofeedback, as an active research instrument, can cut across all facets of this discussion, and pinpoint the pragmatic question of whether different processes identified in the EF and metacognition literature *could* come under conscious control. Our argument is not that neurofeedback enjoys some kind of 'privileged' status in influencing and gaining control over various cognitive processes (i.e. as compared to verbal or behavioral influences, or environmental manipulations), but rather that it is a relatively straightforward and underutilized way of probing the limits and boundaries of our metacognition.

As we also mentioned in the introduction, this is an important thing to test because success or failure to manipulate metacognitive processes with neurofeedback gives a clear indication of which processes are ripe for intervention with more traditional means of support. The twin questions we want to ask are: Can we control it? Does it do us any good?

2.1 Probing the Boundaries of EF and Metacognition

[21] write: "with the accumulation of functional neuroimaging data, it has become obvious that the brain is not organized like a cognitive psychology textbook, with dedicated systems for perception, attention, working memory, episodic memory, and so forth" (p 141). But if the brain is not organized like a cognitive psychology textbook, what have we been handed as a replacement for the old categories?

Currently, great controversy exists over the detailed nature of EF (e.g. see [22]). An abundance of candidate mechanisms can be found in the literature (e.g. conflict-monitoring, error-detection, inhibition, filtering, suppressing, stopping, interference resolution, dual-task coordination, resource allocation, etc.), but as [23] concludes, most researchers work with a few pet tests and act as if their results are generally applicable. At the same time, a consensus is beginning to emerge on the vital importance of EF for real world behavior (not only in the educational domain, but in all aspects of life). To give two examples:

Response inhibition is a (putative) EF ability measured with the Go-NoGo test. This test is considered to be a straightforward assessment of the ability to suppress motor responses [2]. In Go-NoGo tests subjects are presented with series of 'go' signals for a predetermined response, and occasionally a 'no-go' signal that requires the subject to refrain from responding. Other variants of the response inhibition test include the 'stopping task' or the Continuous Performance Test (CPT). Clear developmental trends have been observed over a variety of response inhibition test, such that older children perform significantly better both in terms of speed of responding and lack of commission errors [2]. In relation to the Go-NoGo test, imaging studies have also shown correlations between some prefrontal areas and performance. Greater activation in the orbitofrontal cortex (OFC) is associated with better performance, while increased activation in the anterior cingulate cortex (ACC) is associated with more errors [2].

In terms of real world implications, response inhibition is thought to be related to the all important (especially in academic settings) ability to exercise self-control [1]. According to some researchers poor performance on response inhibition test signal an implulsive personality (these correlation can and should of course be questioned, more on ecological validity below).

Interference resolution is another candidate EF mechanism that is measured by a whole set of EF test (such as the familiar Stroop Test, or the Eriksen Flanker Test, or the Sternberg Item Recognition Test). Like the response inhibition tests described above such tests specifically engage prefrontal regions [24]. As an example, [25] used a version of the Sternberg Item Recognition Test designed to test both working memory load, and resistance to interference (i.e. both keeping information *in* and *out* of mind). The results showed that in terms of brain activation there was a high degree of overlap between the load and interference conditions (different areas in prefrontal cortex, ACC, and parietal cortex), and no new regions were activated exclusively by the interference task. [25] also found significant differences across individuals, and clear correlations between brain-activation and task performance. Activation in the right dorsolateral PFC was significantly correlated with successful task performance. Another more recent study expands upon this conclusion: [26] measured performance on a working-memory task, with critical lure items intermingled in the test (i.e. items designed to be almost like the targets that subjects were instructed to remember), and found that activity in lateral PFC strongly correlated with the ability to overcome the lures (i.e. resolve the interference). Even more interesting was that ability to overcome the critical lures also strongly correlated with performance on a standard IQ-test (Raven's Advanced Progressive Matrices). [26] conclude that conflict monitoring, and resolution of conflict is an important component of general intelligence, and that lateral PFC is a crucial site that underlies this ability.

Thus, it seems like both 'response inhibition' and 'interference resolution' are important abilities that any researcher interested in applied metacognition ought to consider when designing tools for metacognitive support. The problem is that these processes (important as they may be) are almost entirely devoid of phenomenology. *What is it like* to resolve interference in the Eriksen Flanker Test (to paraphrase Nagel's famous question)? When [26] discusses the specific cognitive functions that lie behind good cognitive performance in their experimental setup they come up with a long list of candidates:

> /…/ there are many possible cognitive functions that any given region might be supporting: the inhibition of incorrect responses… the detection, monitoring or reduction of conflict, increased mental checking under high-interference conditions, maintaining or updating task goals and subgoals, re-engaging task processes after a lure trial, other forms of switching between task components, and so on (p.320).

As far as we can see, it is unclear whether *any* of these processes can be consciously influenced or controlled by the subject. What is it that we do when we go about 'updating task subgoals', or 're-engage task processes after a lure trial'? However, if differential brain activity can be identified in different tasks (e.g. Go-NoGo, Stroop, Ravens Matrices, etc.), or differential profiles of activity can be discerned in sub-components of the tasks, then this activity would also be a natural target for neurofeedback manipulations. Initially it would be of interest to simply investigate whether the highly specific patterns of activity discussed in EF studies (like the ones described above) *can* be influenced by neurofeedback. As such, this would be an instrument to probe the degree of *modularity* or encapsulation of the processes at hand. If training by neurofeedback is capable of influencing such processes (which we deem highly likely, see discussion below) it could then be further used to reveal highly interesting and complex interrelationships between the different putative EF tests – i.e. if training to influence the neural substrates that underlie a particular test, also (for better or worse) influences outcomes on another part of the battery, then this would be an indication of important commonalities.

For example, with regards to the concepts of *attention* and *inhibition* [24] favor the idea that they represent a single mechanism. As they see it, selective attention and behavioral inhibition are two sides of the same coin: "Attention is the effect of biasing competition in favor of task-relevant information, and inhibition is the consequences this has for the irrelevant information" (p 186). [27] on the other hand, are of another opinion: "In order to accomplish a task that requires 'executive function' there is not one unitary process that is implemented /…/

attentional allocation alone cannot account for the execution of both switching and interference-resolution processes. There are also separate mechanisms that mediate the switching of attention and the inhibition of a prepotent response (p. 369)".

Of course, the most straightforward application of neurofeedback and metacognition would be to look for performance gains in relevant environments (for the academic context, different forms of tests, or competent problem solving, or transfer of learning, or other aspects of self-regulated learning). Here it must be noted that the prospects of neurofeedback as an applied tool for metacognition is not all too intimately bound up with the fate of EF theories. Applications based on neurofeedback can proceed in a relatively non-theoretical, ad-hoc, task based fashion, using comparisons between individual differences in the variables of interest as a basis for intervention. For instance, in one recent study by [28] an EEG-profile database was created from 59 subjects during two different auditory memory tasks, and memory performance was correlated with the EEG variables. Next, a clinical group was tested on the same two tasks, and *deviations from the values that predicted success* in the other group was measured. After this stage, a neurofeedback training regime, aimed at modulating those critical variables, was established for the clinical group. While the improvement rates in this particular case were very good (in the range of 68% to 181% better for the clinical group, see [28]), it is the general methodology that is of primary importance. Obviously, the establishment of EEG variables to discriminate between good and bad performance on a particular task is not limited to comparisons between normal and clinical populations. As long as a range of individual differences can be identified, and reliable neural indicators of performance can be found, the same procedure can be used to train individuals from good to even better (of course, the variables that correlated with success may have nothing to do with the *causes* of success, so nothing is guaranteed). In any case, with this methodology it is possible to create very specific and potentially powerful applications based on neurofeedback, without detailed knowledge of the underlying cognitive architecture.

2.2.1 Ecological Validity

Apart from the promise of immediate or long-term individual benefits, neurofeedback can also be used as an explicit test case for the ecological validity of laboratory based measures of EF and metacognition.

As a rough example of what such a process might look like, take the simple and understandable notion of *goals* and *planning*. According to EF-theorists, maybe the most encompassing and most outstanding executive function is the active maintenance of behavioral goals, and goal relevant contextual information [24]. In line with the studies of inhibition and interference resolution described above, it is the job of the PFC to actively maintain current goals and coordinate behavior in accordance with those goals.

But once we are outside of the laboratory, what does it mean to hold a goal 'active' in the prefrontal cortex? Are 'active' goals impervious to distractions and temptations? Or is it more likely that indices of prefrontal activity that successfully predict efficient work on a particular EF task, lose much of their explanatory power in the face of real world complexities? Strong indications that this is the case already exist (correlations between EF-tests and ratings by patients and observers about performance in natural settings is disturbingly low, the shared variance between such measurements often being below 10%, [1]). With neurofeedback as an additional tool this can be tested in a much more efficient manner.

For example, a common outcome of frontal lesions is a marked impairment in the capacity for goal representation and goal persistence – what variously has been called "Dysexecutive Syndrome" [29] or "Strategy Application Disorder" [30]. In the words of [31]:

> Careful testing has revealed that the behavior of humans and monkeys with prefrontal damage can be described as *stimulus-bound*. Their behavior is captured by salient sensory cues that reflexively elicit strongly associated actions. They are unable to override these impulses to engage in behaviors that depend on knowledge of a goal and the means to achieve it (p 61, our italics).

Despite this, it is not *impossible* to pursue long-term goals even with severe frontal lobe damage. [29] have shown, in rehabilitations studies with 'dysexecutive' patients, that impaired performance on the very demanding Six Elements Task [30], in which subjects are asked to allocate limited time among six semi-realistic assignments within a common theme, can be restored almost to the level of control subjects by the simple use of a wireless auditory pager system that at random intervals alerts the patients to think about their goals and what they are doing.

If we allow ourselves to use this simple alertness intervention in the Six Elements Task as a proxy for real life situations (where the pager might be substituted for a family member, or a conscientiously organized day planner), it is not difficult to see that the representational and coordinative power of the PFC can easily be eclipsed by the intelligence inherent in well designed cultural artifacts and environments [32]. However, as of the present day we simply do not know the relative importance of structure coming from cultural artifacts and environments, compared to structure coming from a well honed 'lateral PFC' (constantly "maintaining or updating task goals and subgoals", as [26] suggest). With the use of neurofeedback to manipulate activity in PFC-regions in an orderly manner it would be possible to make much more rigorous tests of the relationship and interplay between levels of brain activity and the use of cultural artifacts. As we see it, this would represent a substantial step forward for research on EF and metacognition, and a strong reason to further pursue the possibilities inherent in neurofeedback as an applied research tool.

2.2. What can we do with Neurofeedback?

A natural question to ask at this stage is whether we have rushed ahead of ourselves, and identified domains of cognitive activity that are too detailed and specific to be influenced by neurofeedback. The actual empirical examples of neurofeedback modulation we have presented have all had a rather coarse-grained character (testing attention, working memory, word comprehension, etc, e.g. see [13, 14, 15]), that do not live up to the 'new standard' of specificity advertised in the introduction. The question is, does this spring from a particular limit inherent in the process of neurofeedback? Exactly what kind of modulations of cognitive (or metacognitive) activity can we expect from different forms of neurofeedback? How wide is the range of processes that can be influenced? While only preliminary answers can be given, we believe it is reasonable to expect that with refined instruments of neurofeedback, a great share of cognitive and emotional activity previously outside of our scope of influence may come under conscious control (it is another question whether improved levels of control always is a good thing).

An initial reason to be optimistic about the level of control we may achieve comes from the wider study of the role of external prompts and scaffolding in influencing our self-regulatory efforts. A common observation here is that it often is a world of difference between what we can do on our own, with our bare brains (so to speak), and what we can do when embedded in a matrix of strong external prompting/feedback, and timely strategic input. As an example, take the very interesting set of studies conducted by [33] and [34]. They have shown that *hypnotic suggestion* about either the degree of unpleasantness of painful stimuli, or the degree of intensity of the pain, can induce a double dissociation between these dimensions.

Effective suggestions that reduce the rated *unpleasantness* (the affective-motivational dimension) also reduce activity in the anterior cingulate cortex (ACC, a region that has regularly been implicated in the representation of pain valence), but not in primary somatosensory cortex (S1) [33]. Suggestions that target the *intensity* of the pain (the sensory-discriminative dimension) instead reduce activation in S1, but leave ACC unaffected [34].This is a remarkable and unintuitive result. As [35] have shown, we may have a natural ability to temporarily inhibit pain (again, dorsolateral PFC seems to be the culprit), but the finer details of pain perception are very difficult to influence by way of strategic metacognitive processing.

On the other hand, a functional, anatomical 'fragmentation' of a previously unitary subjective experience is a perfect target for neurofeedback applications, as feedback of the localized activation would give subjects a unique opportunity to work to achieve control over that particular dimension (although we do not want to suggest that training by neurofeedback can automatically reproduce all the causative agents involved in hypnotic induction).

As the relevant studies have not yet been conducted with human subjects (i.e. with high-resolution fMRI and very specific sites of influence) it is interesting to note that in animal studies of neurofeedback it was established, already in the early seventies, that monkeys with electrodes implanted in cortex could learn to self-regulate the firing-rates of *individual neurons* (as much as 50-500% higher) in response to visual or auditory feedback combined with food reward (e.g. see overview in [36]). On top of being a fascinating example of fine-grained cortical modulation, this latest example also suggests that a good way of approaching the question of the range of possible modulations, is to look at *how* the modulation is achieved. In this case, it is believed that the learning in the monkeys was heavily influenced by operant, reward learning (as evidenced by the fact that the learning reversed when the reward was removed, [36]).

2.2.1 How do we do it?

The question of exactly *how* and *why* we come to be able to control specific neural states by feedback training has not yet been answered properly by the neurofeedback research community [37]. In some cases it turns out to be very difficult to say how the regulation is accomplished. [38] studied EEG dynamics in a highly trained paralyzed subject operating a brain-computer interface (BCI) (where power shifts at 12Hz over his left- and right- central scalp controlled a cursor that could move to target boxes in different places on a computer screen). In this case it was shown that the specific regulation of the EEG required by the BCI was accompanied by other large scale, simultaneous modulations of the EEG at several frequencies and in multiple cortical areas. Exactly what 'strategy' this activity represented could not be determined (more than the contention that it was not a muscle or motor-type event, see [38]). This can be compared to [39] who also studied the mental strategies of a paralyzed patient using a BCI. In this case the patient was able to write a long and detailed letter describing very specific strategies for control (using the BCI, it took him 6 months to write the letter at an average pace of two characters per minute) that matched the actual EEG patterns remarkably well [39].

But it is not only in the special case of paralyzation that we may have great difficulty telling what 'strategic manipulations' that lie behind the self-regulation. It is a common case in neurofeedback experiments to find that the ability to control some specific parameter develops far ahead of knowledge of how this is accomplished (e.g. [40], this is also reminiscent of metacognitive strategy development in the educational domain, as studied by [41]). As such, much of the general debate of the efficacy of neurofeedback has been centered on the relative importance of specific cognitive strategies vs. (more or less) implicit operant conditioning. On the one hand we have examples like [42] who found that children claimed to use identical ‘mental strategies’ to regulate their slow cortical potentials even though (unbeknownst to the participants and the conducting therapist) the relation between the feedback and the amplitude of SCP was inverted in mid session (i.e. the feedback allowed them to regulate the EEG into opposite directions while still being under the impression that they did the same thing). On the other hand it has frequently been shown that patients in BCI programs supposed to be based exclusively on operant training, claim to use the same mental strategies as those that have been clearly validated in other programs (i.e. most often some version of motor imagery of specific actions, see [37]).

If neurofeedback is to become a useful and reliable tool for practical metacognitive applications, this is clearly an important issue to solve. But there are also other types of learning effects that deserve attention, and which point to important and unexplored implications of neurofeedback training. In a recent study conducted with epilepsy patients trying to induce shifts (both positive and negative) in their SCP it was determined that, in addition to actual control, reliable knowledge of *the degree of success* on each trial developed throughout the training sessions [40]. Thus, in a non-metaphorical sense, it seems like the subjects learned to *perceive their own brain states* (the authors of the experiment dismisses the option that this knowledge is based on some global feeling or arousal, or something similar). If the interpretation of [40] is correct, this is an effect worthy of further exploration. It is a well-known fact that experience by training can affect many ‘objective’ psychophysical measures of perceptual performance, including thresholds for stimulus detection. [43] for example, discusses the finding that the inclusion a few suprathreshold trials in a perceptual learning procedure can lead to a sudden drop in the threshold for detection. She speculates that the “stronger signals available in these suprathreshold trials ‘show’ the subject where the relevant representations can be found in the nervous system” (p. 107). Thus, the neurofeedback study by [40] can bee seen as an example of a particular training and feedback regime allowing subjects to learn new access relations *within* the brain.

Not the least is this an important result for anyone interested in realistic applications for neurofeedback in the domain of metacognition, for it indicates the possibility that neurofeedback training regimes do not have to be sustained indefinitely, as subjects eventually may learn to perform and assess the outcome of the regulation without feedback correction (apparently in contrast to the monkeys with implanted electrodes described by [36]). This, of course, means that training mediated by neurofeedback could be useful even in environments where continuous and concurrent measurements are not feasible or practical.

2.2.2 Ambulatory Measurement

As a final point we would like to suggest that it is very important not only to look at the question of how control is accomplished (whether by conditioning or use of specific strategies), but also in what context the training is situated, and how elaborate the feedback environment is. A factor that could point to a substantial role for neurofeedback in practical metacognitive applications (in educational, work, or sports settings, etc.) is the relatively (and surprisingly) *constrained* use of neurofeedback that has been established practice so far.

As we have seen, neurofeedback (indeed the whole field of biofeedback) has been largely preoccupied with the narrow achievement of *control* over physiological and brain states. We believe a broader take than this is called for. In fact, as we see it, the great boon of neurofeedback as a practical metacognitive tool will not manifest itself unless we look beyond training to control brain states, to the prospect of proactively using the sensed neural-context in wider systems of regulation. The possibilities of automated context-awareness as a new *interface* for ourselves go far beyond simple feedback-control. Only detailed experimentation can determine what function (metacognitive, modulating, communicative, explanatory, rewarding, facilitating, distractive, evidential, etc.) that neurofeedback might play in any given computer-mediated system designed to support metacognition.

One example of this wider use of neurofeedback is systems designed to provide ambulatory measurements (i.e. at the site of performance, with an element of mobility). [44, see also 45] have explored this concept in industrial environments where human operators risk facing work overload. Human vigilance research has shown that for operators engaged in attention-intensive tasks, retaining a constant level of alertness is next to impossible. In this context, [44] constructed a neural network based system for real-time monitoring of operators that extracts critical indices of alertness in the EEG. The point of the system was not to give the operator feedback from the EEG in order for her/him to introspectively control the EEG activity. Instead, the systems was designed to act as a *context-aware* application [46] that can provide the operator with timely cues regarding their cognitive status, and automatically adjust the information intensity in the task environment based on the measured level of vigilance. In analogy with more traditional computer mediated tools designed to scaffold metacognitive processes the system detects and monitors important parameters during performance, and suggests solutions based on the state of the user and the priorities of the situation (to go one step beyond this is to have the system in turn *learn* what kind of effects its feedback have on the user, and consequently fine-tune operations based on this assessment - i.e. a concept of *intelligent* biofeedback, as proposed by [47]).

Naturally there are problems inherent in ambulatory sensing of cognitive states. Take EEG, and the study of event related potentials (ERPs, activity phase-locked to different tasks or stimuli). To study ERPs, data is usually averaged prior to analysis in order to increase the signal to noise ratio – e.g. to filter out the weak source signals from the strong background EEG-activity, as well as to avoid contamination from motion or muscle artifacts (such as eye-blinks, or gross body movement). This method works well in the laboratory, but it also ignores the fact that responses may vary widely across trials, or between subjects, effectively hiding temporal and spatial variability that may be linked to changes in subject performance, or to fluctuations in attention, arousal, task strategy, or other important cognitive variables [48]. Averaging is simply not an option in real-world settings. But as was illustrated in the example [44] described above, ERP activity *can* be tracked continuously in real time. Independent Component Analysis (ICA), for example, is one class of algorithms that can be applied to the EEG to separate the activity into distinct components that represent all variables of interest, as well as noise and artifacts [48]. In fact, given recent developments in the field, many researchers now believe that we are on our way towards robust, fully automated tools for removal of artifacts in electrophysiological recordings [49, 50].

Still, as a vision for the immediate future, we believe it is a little far fetched to assume that robust and precise EEG recordings will be portable and efficient enough to serve as a universal tool for metacognitive support (and for the foreseeable future, hulking basement dwellers like MRI-scanners, will certainly not be something we casually don for an afternoon stroll). While we believe this to be a very promising avenue of research, in the near future neurofeedback will most probably make a bigger impact as a research tool.

3. Summary and Conclusion

We have discussed different strands in the relationship between the concept of neurofeedback and research on EF and metacognition. Neurofeedback has traditionally been mostly concerned with the achievement of *control* over brain states. We believe this to be an important, core function, and foresee an expanded role for neurofeedback in the near future, both as a helpful tool to probe and investigate the basic nature of metacognitive processes, and as a powerful instrument to strategically shift the boundaries of our conscious control over a wide collection of cognitive phenomena. However, we also believe a broader take on neurofeedback is warranted. As we see it, the great boon of neurofeedback as a practical metacognitive tool will not manifest itself unless we look beyond training to control brain states, to the prospect of proactively using the sensed neural-context in wider systems of regulation. The possibilities of automated context-awareness as a new *interface* for ourselves go far beyond simple feedback-control.

References

[1] Barkley, R. A. (2001). The executive functions and self-regualtion: An evolutionary neuropsychological perspective. *Neuropsychology Review, 11*(1), 1-29.

[2] van der Molen, M. W. (2000). Developmental changes in inhibitory processing: Evidence from psychophysiological measures. *Biological Psychology, 54*, 207-239.

[3] Baddeley, A. D. (1996). Exploring the central executive. *The Quarterly Journal of Experimental Psychology*, 49:A(1). 5-28.

[4] Fernandez-Duque, D., Baird, J. A., & Posner, M. I. (2000). Executive attention and metacognitive regulation. *Consciousness and Cognition, 9*, 288-307.

[5] Jack, A. I., & Roepstorff, A. (2002). Introspection and cognitive brain mapping: From stimulus-response to script-report. *Trends in Cognitive Sciences, 6*, 333-339.

[6] Badre, D., & Wagner, A. D. (2002). Semantic retrieval, mnemonic control, and prefrontal cortex. *Behavioral and Cognitive Neuroscience Reviews, 1*(3), 206-218.

[7] Kuhn, D. (1996). Is good thinking scientific thinking? In D. R. Olson & N. Torrance (Eds.), *Modes of Thought;: Explorations in Culture an Cognition* . New York: Cambridge University Press.

[8] Vye, N. J., Schwartz, D. L., Bransford, J. D., L., Z., & CTGV. (1998). SMART environments that support monitoring, reflection, and revision. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in Educational Theory and Practice* (pp. 305-346). New Jersey: Lawrence Erlbaum.

[9] Ossebaard, H. C. (2000). Stress reduction by technology? An experimental study into the effects of brainmachines on burnout and state anxiety. *Applied Psychophysiology and Biofeedback, 25*, 93-101.

[10] Fell, J., Elfadil, H., Klaver, P., Röschke, J., Elger, C. E., & Fernandez, G. (2002). Covariation of spectral and nonlinear EEG measures with alpha biofeedback. *International Journal of Neuroscience, 112*, 1047-1057.

[11] Thompson, L., & Thompson, M. (1998). Neurofeedback combined with training in metacognitive strategies: Effectiveness in students with ADD. *Applied Psychophysiology & Biofeedback, 23*(4), 243-263.

[12] Lubar, J. F. (1997). Neocortical dynamics: Implications for understanding the role of neurofeedback and related techniques for the enhancement of attention. *Journal of Applied Psychology & Biofeedback, 22*(2).

[13] Egner, T., & Gruzelier, J. H. (2001). Learned self-regulation of EEG frequency components affects attention and event-related brain potentials in humans. *NeuroReport, 12*(18), 4155-4159.

[14] Vernon, D., Egner, T., Cooper, N., Compton, T., Neilands, C., Sheri, A., & Gruzelier, J. (2003). The effect of training distinct neurofeedback protocols and aspects of cognitive performance. *International Journal of Psychophysiology, 47*, 75-85.

[15] Pulvermüller, F., Mohr, B., Schleichert, H., & Veit, R. (2000). Operant conditioning of left-hemispheric slow cortical potentials and its effect on word processing. *Biological Psychology, 53*, 177-215.

[16] Posse, S., Fitzgerald, D., Habel, U., Rosenberg, D., Moore, G. J., & Schneider, F. (2001). Real-time fMRI of single trial amygdala activation during sad mood induction with feedback. *Proceeding of International Society of Magnetic Resonance in Medicine, 9*, 1-2.

[17] Posse, S., Fitzgerald, S., Kunxiu, G., Habel, U., Rosenberg, D., Moore, G. J., & Schneider, F. (2003). Real-time fMRI of temporolimbic regions detects amygdala activation during single-trial self-induced sadness. *NeuroImage, 18*, 760-768.

[18] Yoo, S.-S., & Jolesz, F. A. (2002). Functional MRI for neurofeedback: Feasibility study on a hand motor task. *Brain Imaging, 113*(11), 1377-1381.

[19] Nisbett, R. E., & Wilson, T. D (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*. 84: 231-259.

[20] Dennett, D. (1991). *Consciousness Explained*. Boston: Little, Brown & Company.

[21] Cabeza, R., & Nyberg, L. (2002). Seeing the forest through the trees: The cross-function approach to imaging cognition. In A. Zani (Ed.), *The Cognitive Electrophysiology of Mind and Brain* : Academic Press Inc.

[22] Wood, J. N., & Grafman, J. (2003). Human prefrontal cortex: Processing and representational perspectives. *Nature Reviews: Neuroscience, 4*, 139-147.

[23] Logan, G. D. (2003). Executive control of thought and action: In search of the wild homonculus. *Current Directions in Psychological Science, 12*(2), 45-48.

[24] Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24*, 167-202.

[25] Bunge, S. A., Ochsner, K. N., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. E. (2001). Prefrontal regions involved in keeping information in and out of mind. *Brain, 124*, 20074-2086.

[26] Gray, J. R., Chabris, C. F., & Braver, T. S. (2003). Neural mechanisms of general fluid intelligence. *Nature Neuroscience, 6*(3), 316-322.

[27] Sylvester, C.-Y., Wager, T. D., Lacey, S. C., Hernandez, L., Nichols, T. E., Smith, E. S., & Jonides, J. (2003). Switching attention and resolve interference: fMRI measures of executive functions. *Neuropsychologia, 41*, 357-370.

[28] Thornton, K. (2000). Improvement/rehabilitation of memory functioning with neurotherapy/QEEG biofeedback. *Journal of Head Trauma Rehabilitation, 15*(6), 1285-1296.

[29] Manly, T., Hawkins, K., Evans, J., Woldt, K., & Robertson, I. H. (2002). Rehabilitation of executive function: Facilitation of effective goal management on complex tasks using periodic auditory alerts. *Neuropsychologia, 40*, 271-281.

[30] Burgess, P. W. (2000). Strategy application disorder: The role of the frontal lobes in human multitasking. *Psychological Research, 63*, 279-288.

[31] Miller, E. K. (2000). The prefrontal cortex and cognitive control. *Nature Reviews Neuroscience*, 1:59-65.

[32] Hutchins, E. (1995). *Cognition in the Wild*. Cambridge, Mass.: MIT Press.

[33] Rainville, P., Carrier, B., Hofbauer, R. K., Bushnell, M. C., & Duncan, G. H. (1999a). Dissociation of sensory and affective dimensions of pain using hypnotic modulation. *Pain, 82*, 159-171.

[34] Hofbauer, R. K., Rainville, P., Duncan, G. H., & Bushnell, M. C. (2001). Cortical representation of the sensory dimensions of pain. *Journal of Neurophysiology, 86*, 402-411.

[35] Lorenz, J., Minoshima, S., & Casey, K. L. (2003). Keeping pain out of mind: The role of the dorsolateral prefrontal cortex in pain modulation. *Brain, 126*, 1079-1091.

[36] Nicoledis, M. A. (2003). Brain-machine interfaces to restore motor function and probe neural circuits. *Nature Reviews Neuroscience, 4*, 417-422.

[37] Curran, E. A., & Stokes, M. J. (2003). Learning to control brain activity: A review of the production and control of EEG components for driving brain-computer interface (BCI) systems. *Brain and Cognition, 51*, 326-336.

[38] Delorme, A., & Makeig, S. (2002). EEG changes accompanying learned regulation of 12-Hz EEG activity. *IEEE Transactions on Rehabilitation Engineering*.

[39] Neumann, N., Kübler, A., Kaiser, J., Hinterberger, T., & Birbaumer, N. (2003). Conscious perception of brain states: Mental strategies for brain-computer interaction. *Neurophysiologia, 41*, 1028-1036.

[40] Kotchoubey, B., Kübler, A., Strehl, U., Flor, H., & Birbaumer, N. (2002). Can humans perceive their brain states? *Consciousness and Cognition, 11*, 98-113.

[41] Siegler, R. S., & Stern, E. (1998). Conscious and unconscious strategy discoveries: A microgenetic analysis. *Journal of Experimental Psychology General, 127*, 377-397.

[42] Siniatchkin, M., Kropp, P., & Gerber, W.-D. (2000). Neurofeedback - the significance of reinforcement and the search for an appropriate strategy for the success of self-regulation. *Applied Psychophysiology and Biofeedback, 25*(3), 167-175.

[43] Kanwisher, N. (2001). Neural events and perceptual awareness. *Cognition, 79*, 89-113.

[44] Jung, T.-P., Makeig, S., Stillwell, D., & Harm, D. (2001b). *Accessing cognitive state from physiological data*. Paper presented at the Bioastronautics Investigators' Workshop

[45] Gevins, A., Smith, M. E., Leong, L., McEvoy, L., Whitfield, S., Du, G., & Rush, G. (1998). Monitoring working memory load during computer-based tasks with EEG pattern recognition methods. *Human Factors, 40*, 79-91.

[46] Weiser, M. (1991). The computer for the 21 st century, Scientific American, 265 (3): 94-104.

[47] Bersak, D., et al. (2001). *Intelligent biofeedback using an immersive competetive environment.* Paper presented at Ubicomp 2001.

[48] Jung, T.-P., Makeig, S., McKeown, M. J., Bell, A. J., Lee, T.-W., & Sejnowksi, T. J. (2001). *Imaging* brain dynamics using independent component analysis. *Proceedings of the IEEE*. 89:7. 1107-1121

[49] Delorme, A., Makeig, S., & Sejnowski, T. (2001). *Automatic artifact rejection for EEG data using high-order statistics and independent component analysis.* Paper presented at the Third International Workshop on Independent Component Analysis and Signal Separation, San Diego.

[50] Browne, M., & Cutmore, T. R. H. (2002). Low-probability event-detection and separation via statistical wavelet thresholding: An application to psychophysiological denoising. *Clinical Neurophysiology, 113*, 1403-1411.