Towards Emotional Regulation through Neurofeedback

Marc Cavazza, Fred Charles, Gabor Aranyi, Julie Porteous, Stephen W. Gilroy
School of Computing
Teesside University
Middlesbrough, United Kingdom
m.o.cavazza@tees.ac.uk

Gal Raz¹,², Nimrod Jakob Keynan¹,², Avihay Cohen¹, Gilan Jackont¹,², Yael Jacob¹,², Eyal Soreq¹,², Ilana Klovatch¹, Talma Hendler¹,²
¹ Functional Brain Center, Tel Aviv Sourasky Medical Center, Tel Aviv, Israel
² Tel Aviv University, Tel Aviv, Israel
hendlert@gmail.com

ABSTRACT
This paper discusses the potential of Brain-Computer Interfaces based on neurofeedback methods to support emotional control and pursue the goal of emotional control as a mechanism for human augmentation in specific contexts. We illustrate this discussion through two proof-of-concept, fully-implemented experiments: one controlling disposition towards virtual characters using pre-frontal alpha asymmetry, and the other aimed at controlling arousal through activity of the amygdala. In the first instance, these systems are intended to explore augmentation technologies that would be incorporated into various media-based systems rather than permanently affect user behaviour.

ACM Classification Keywords
H.5.2 [Information Interfaces and Presentation]: User Interfaces

General Terms
Human Factors; Design; Measurement.

Author Keywords
Brain-Computer Interfaces, Neurofeedback, Virtual Reality, Affective Computing

INTRODUCTION AND OBJECTIVES
The role of emotions in human cognitive abilities and performance has been an object of constant debate. Prominent authors in Cognitive Science such as Damasio [7] and, more recently, Thagard [41] have advocated a constructive role for emotions as part of cognitive abilities. However, as noted by Gross [13], behavioural responses prompted by emotions may not be those most appropriate to the fulfilment of our objectives. Ochsner and Gross [32] have proposed a hypothetical continuum between the various sorts of cognitive control of emotions, ranging from selective inattention to emotional stimuli, to reappraisal methods. They identified various neural structures involved for each of these mechanisms, highlighting the involvement of the pre-frontal cortex in higher-forms of emotional control, as well as the amygdala for its role in pre-attentive emotion appraisal.

In line with a large set of evidence from animal studies [11], Beauregard et al. [3] have identified neural correlates for the conscious regulation of emotion as a combination of the pre-frontal cortex (PFC) and limbic structures, including the amygdala. They have shown that attempts at voluntarily inhibiting emotional responses resulted in activation peaks in the prefrontal cortex (BA 10), while amygdala was involved in the generation of the emotional response itself, in particular arousal.

In this paper, we advocate emotional control as an important component of human augmentation and suggest that neurofeedback (NF) techniques, incorporated into multimedia systems, could form the basis for a new technology supporting situated emotional control. After reviewing current discussions in the field of neuroscience on the neural basis of emotional control, we present two case studies of experimental, yet fully implemented, systems exploring emotional regulation, essentially from a perspective of response-focussed modulation. We conclude by discussing potential directions to extend these to real-world augmentation challenges.

PREVIOUS AND RELATED WORK
Aside from fundamental Neuroscience work on emotion control, few researchers have addressed how computational systems could be developed to support metacognition and the control of executive functions. Hall and Johansson [15] have discussed extensively the relationship between metacognition and NF from the perspective of developing computer tools to augment human performance (in particular metacognition). One of their arguments is that, for NF to progress, subjects should gain a better awareness of executive functions, including many higher-level control functions. However, they also suggest that NF itself may be part of the solution, as novel experiments tend to bring hitherto unknown control abilities to the subjects’ consciousness, as if there was a bootstrapping effect in experimenting with NF.

Affective computing has been recognised as an important topic for human augmentation [37] with previous work
addressing emotional feedback through facial expressions [46], or companion robots [42]. Emotion recognition for human augmentation can make use of physiological computing [47] or EEG techniques adapted to a use outside laboratory conditions [29], or for specific applications such as augmented reading [33]. Brain-Computer Interfaces (BCI) are poised to play an important role in human augmentation, from traditional application in assistive technologies to specific techniques for video augmentation [22] or control of avatars [23]. In addition, there has been extensive research in biofeedback, NF and BCI in digital entertainment and serious games (see for instance [27] for a recent review).

NEUROFEEDBACK FOR EMOTIONAL REGULATION

Although NF has a relatively long history, it has recently been the subject of renewed interest following progress in neuroscience, which has provided a more rigorous basis to NF experiments and applications [14], including clinical ones [31]. Furthermore, NF has been suggested as one of the methods of choice for experimenting with emotional control [21].

NF can be defined as the process by which a subject learns to gain control over specific aspects of neural activity, putatively through operant conditioning or volitional control [14]. There are several components to NF implementation: the identification of a target brain region; the choice of measurement for this region’s activity (EEG, fMRI [45]); the presentation of the feedback signal to the user (including ‘rewarding’ desirable patterns of brain activity); the experimental paradigm (specifying the number and duration of NF trials, success thresholds); and the user instructions. Although not strictly speaking an element of the experimental setup (in which case it would pertain to the “user instructions”), users’ spontaneous mental strategies also play an important role in the success of NF experiments [24, 30].

When using NF to explore emotion control, one is inevitably faced with the much-debated issue of identifying the neural basis of target emotions. Several recent reviews have concluded that mental categories for universal emotions are not respected by the brain [26], and that previous studies failed to find consistent and specific correlates for basic emotions [16]. Cacioppo [6] has suggested that the prominence of feelings was hindering the analysis of finer-grained mechanisms for emotion generation. From a similar perspective, Lindquist et al. [26] have advocated the need for what they called “mid-level scientific categories” that better relate basic emotional processes to neural structures.

We will illustrate this strategy through two separate examples, which consider emotions of different complexity and different hierarchical control levels. Both have inspired our proof-of-concept experiments, described in further sections. We have decided to identify specific processes, for which previous literature has established a clear role in emotional processes, and for which NF experiments have been previously reported (although not necessarily for emotional control). One is amygdala’s activation (and down-regulation); the other is prefrontal cortex (PFC) asymmetry. The amygdala has been considered as an important region for emotion, and has most often been associated with fear: yet, in their recent review of neural reference spaces, Lindquist et al. [26] synthesise the role of the amygdala as concerned with “high arousal emotions”. Beauregard et al.’s [3] experiments have shown a decrease in amygdala activity when subjects attempted to control their emotional responses; Hare et al. [17] have suggested the involvement of the amygdala in evaluating emotional significance. Blair [5] has proposed a role for the amygdala in morality, in conjunction with the ventromedial prefrontal cortex (vmPFC).

Thus it should be possible to use amygdala activation as part of NF experiments, in which the subject is controlling emotional responses, in particular arousal. Tracking the activation of an anatomical structure can be implemented using real-time fMRI [48], but this is unlikely to result in technologies easily available outside specialised laboratories. A novel possibility that has recently emerged is to use an innovative EEG ‘fingerprinting’ technique [28], whose activity pattern reflects that of the target structure, in this case the amygdala. It then becomes feasible to use an almost standard approach for EEG-based NF, including in terms of the timing of the feedback loop, which will benefit from EEG’s high sampling frequency.

Our second example is constituted by PFC asymmetry as investigated by Davidson [8]. Davidson has identified alpha wave asymmetry in that region as a marker of approach/avoidance [39], which again can be seen as one of the mid-level categories advocated by Lindquist et al. [26]. PFC asymmetry has attracted much interest, and been further associated to a variety of emotional and clinical phenomena including mood disorders [19] and empathy [25]. In his interim review of progress in the field of PFC asymmetry, Davidson [9] identifies amygdala as one of the main targets of PFC-mediated control, which also supports Blair’s [5] findings on the neural basis of morality. The connection between PFC asymmetry and empathy has been the object of many conjectures. Decety and Moriguchi [10] seem to imply that approach, in a social context, could be a phylogenetic precursor of empathy. Light et al. [25] have demonstrated the implication of PFC asymmetry in various sorts of empathy: empathic concern, empathic happiness and empathic cheerfulness (the latter having inspired our own experiments, see below).
Tullett et al. [43] have identified a link between withdrawal and empathy when considering empathic concern, which is known to be associated with negative feelings, but this is not inconsistent with a link between approach and empathy for empathic cheerfulness. Finally, alpha PFC asymmetry has been used as part of NF paradigms in clinical settings, including for the treatment of depression or mood disorders [2], although with variable results [34]. Importantly, such research has established the suitability of alpha PFC asymmetry as a target for NF, although several experimental difficulties (e.g., reference electrodes) have been identified [9].

CASE STUDY I: CONTROLLING FILMIC EMPATHY

Movies are a form of entertainment conspicuously designed to elicit various emotional responses, which are an integral part of the spectator’s experience, and one of the actual motivations for film viewing. However, emotions induced by film are, by definition, beyond the control of the spectator. Although competing theories exist about filmic emotion [38], one of the most popular posits that spectators’ dispossession towards characters plays an important role in their overall emotional responses. Such model could be described as empathic [40]. This echoes neuroscience research in empathy which has used filmic emotional responses to identify neural networks specific to various sorts of empathy [36].

Our objective was to design a system in which the user could gain control over her disposition towards the main character of a story, and convert her feelings into actual support, in a way that would impact on the unfolding story.

We used interactive narrative techniques [35] to implement the baseline story, which appears to the subject as a 3D animation movie, generated in real-time, whose total duration is 7-8 minutes on average. Interactive narratives are based on Artificial Intelligence techniques (such as planning), which dynamically generate a consistent plot from a set of baseline actions, each action instance being visualised in real-time using 3D computer graphics. The use of interactive narrative makes it possible to dynamically generate dramatic situations eliciting emotional response (Figure 1), rather than using pre-defined ones: this brings experimental flexibility when researching emotional responses. In addition, successful user intervention can be translated into a modification of the story course, which provides an additional motivation as well as an immediate measure of success.

Our baseline story is a medical drama whose main feature character is a female junior doctor facing multiple challenges in her work, both in terms of medical emergencies and interpersonal relationships with colleagues. The medical drama genre was originally chosen for its popularity and its ability to elicit a whole range of emotional reactions. Its interactive narrative implementation will face her with a string of difficult situations (overworking, harassment, and death of her patients): once a critical stage is reached, the user will be given the opportunity to support her using a single NF session of limited duration (30 seconds). In the absence of successful user support, the character’s situation will continue to deteriorate; on the contrary, should the user succeed in harnessing empathy, the story will evolve towards a much more favourable ending. This is an attempt to unify different aspects of filmic emotion around the concept of empathy: the generation of threatening situations for a character using interactive narrative techniques, the emotional response of users to such situations, and the neural correlates of this response turned into a Brain-Computer Interface based on empathic NF (a virtual transposition of empathic cheerfulness using alpha PFC asymmetry).

Figure 1 describes the system architecture. The user watches the animated story generated in real-time on a monitor, equipped with an EEG cap. Once the story has reached a critical step (which is dynamically generated hence varies across sessions), the system triggers a short NF
session creating a bespoke scene within the story itself, then resumes story generation. The NF mechanism is based on alpha-wave prefrontal asymmetry, using Davidson’s $A_2$ asymmetry score (F4-F3/F3+F4) calculated in real-time during the NF session. This metric has been correlated to empathy in previous studies (see below for a discussion). In addition, PFC alpha-asymmetry has been the object of extensive NF studies in clinical settings and, although pursuing different objectives, these studies have established the suitability of PFC alpha-asymmetry for NF. Feedback is considered successful if the $A_2$ score exceeds the $A_2$ baseline (obtained for each subject during a calibration stage at rest) at a statistically significant level during the embedded NF session (this is to account for the difficulty in actually sustaining NF over prolonged periods of time) (Figure 2).

We undertook early experiments using fMRI as a baseline validation and EEG as BCI input [12]. EEG data was acquired using a 32-electrode MRI-compatible BrainAmp MR system (Brain Products, Munich, Germany). Data was recorded at a sampling rate of 5000Hz and collected on a PC running RecView software (Brain Products) for gradient and cardioballistic artefact removal. Twelve right-handed male subjects took part in this experiment. Of these, six were successful with NF despite minimal training, which is well in line with success scores reported for NF, since it is generally considered that up to one third of subjects fail at NF even with significant training [24]. In addition, fMRI scans taken during the experiment confirmed pre-frontal BOLD asymmetry in all successful subjects [12], albeit of variable magnitude, in those areas consistent with previous work, in particular the frontopolar and ventromedial PFC (BA 10).

Since the role allocated to the user is to comfort or cheer up the feature character, it is fair to consider that the most relevant sort of empathy is empathic cheerfulness [25]. It is generally defined as the production of positive emotions in response to the recognition of a negative situation affecting another subject. This sort of empathy is strongly correlated with prefrontal asymmetry [25]. This should be translated into the NF instructions issued to subjects. There is no agreement on how to design such instructions, in particular taking into account the highly variable strategies reported by users. In this specific case, we could identify a number of constraints on the instructions: they should be specific to the type of empathy considered without being too prescriptive about the mental strategy; they should minimise false positives for the type of NF signal considered, and they should not introduce artefacts. Our instructions to the subjects were “to support the feature character by expressing positive thoughts”. It should be noted that these are compatible with current knowledge on alpha PFC asymmetry, as well as neutral to the motivation/valence controversy [18]. We also described that such support will be called for when the character’s situation within the story would become critical. Finally, the users were explained the NF success criterion which consists in restoring the normal appearance of the character from her fading to grey (Figure 2).

During debriefing sessions, successful subjects expressed various strategies they used during NF: these included mentally “cheering up” the character, thinking about positive and pleasant personal situations, or thinking about real-word friends or family. It is interesting to note that these spontaneous strategies are in line with those previously reported in the NF literature: cheering was mentioned in [24] and real-life events and family in [30].

The full implementation, together with its fMRI validation, is described in greater detail in [12].

**CASE STUDY II: AMYGDALA REGULATION**

This second test case investigated subjects’ ability to control the activity of their amygdala, which, unlike the pre-frontal cortex, is a control and integration structure that is involved in emotional responses in a more primitive fashion. As previously discussed, while for long the amygdala has been associated to fear, this locationist assumption has been largely abandoned in favour of more subtle associations [26]. It is currently accepted that amygdala activation correlates more specifically with arousal, and its functional role is to assess novelty in a situation.

We have developed another NF setting based on the simulation of a realistic situation rather than abstract feedback signals such as gauges, meters or sounds. The rationale is that potential human augmentation applications would be set precisely in similar, albeit real rather than virtual, environments, supporting emotional control in social situations, but also our belief that more sophisticated feedback visualisations give new insights into the NF process. There has been limited research published on this issue (one of the few exceptions being [20]), but we have anecdotal evidence that realistic visualisations are able to elicit a better range of emotions, again through some of their more narrative properties. This is important as well when the initial situation is meant to induce an emotional response. In addition, realistic environments offer more opportunities to adapt the response time of the feedback loop which can be used to mitigate oscillations; this will be briefly discussed in a further section.
This experiment (Figure 3) is also based on a NF paradigm, and comprises a target emotional response (amygdala activity / arousal), an input probe (amygdala EEG Fingerprint), and mapping to a metric that governs the behavior of the visual environment. Once again, appropriate instructions are issued to the user so she can follow the NF paradigm. The overall experiment investigates the control of arousal based on down- or up-regulation of the amygdala. The sensor is a bespoke EEG fingerprint (EFP) of the amygdala activity obtained from fMRI data, consisting in a non-standard electrode placement as described in [28]. In terms of signal acquisition, it brings flexibility, ease of use and specificity to the target region. It allows a response time, which is that of the EEG signal, making it possible to design appropriate feedback loops not limited by signal acquisition. In addition, this amygdala EFP offers more specificity, better correlation and faster response times than peripheral signals of arousal, such as Galvanic Skin Resistance (GSR).

The virtual environment reproduces the waiting room of a public building (hospital), where a representative sample of the population is waiting to be called at the admissions desk. There is an imbalance between the number of characters waiting and only one admission staff at the counter, which in the real-world would quite logically trigger some form of unrest in the waiting population. Unrest manifests itself through the fact that some patients stand up from their seats to reach the counter and enquire or complain. The “global unrest” of the waiting room as a whole is the matching variable, and it is given as the proportion of patients that stand at the counter (all patients are calmly sat down at the start of the simulation). A matching soundtrack, part of which has been actually recorded inside a hospital as part of a documentary (used with permission) complements the system output. In the baseline state, all 20 virtual patients take seats in the waiting room, while at the peak of unrest, half of them congregate at the front desk also expressing their frustration through body language. This “level of unrest” is thus set to constitute the feedback signal and is quantified using the ratio between the number of characters at their seats and characters at the front desk, with matching sound levels. The system is implemented using the Unreal Development Kit (UDK) game engine, which allows to control walking animations for individual characters. Changes to the distribution of sat and standing characters are made as a function of the level of unrest by randomly selecting the next character to change state and triggering the corresponding animation.

Subjects’ instructions consist precisely in trying to either appease the situation or, conversely, in fostering unrest in the waiting room: however, once again the instructions do not prescribe specific mental strategies. The level of unrest has a metaphorical connection with arousal that is easy to grasp: nevertheless there is still a need to provide a principled mapping between the amygdala probe signal and the level of unrest to implement NF. The basis for this mapping is to consider the ratio between characters sat down and protesting at the counter as a two-state Boltzmann distribution [1], whose evolution is driven by a virtual “temperature” whose value is derived from the amygdala activation measured by the EFP (Figure 4).

\[
\frac{N_i}{N_f} = \frac{e^{-\frac{E_i}{k_B T}}}{e^{-\frac{E_f}{k_B T}}}
\]

Where the equivalent “energy” levels are determined empirically as a function of desired system behavior during calibration.

---

1 We have tentatively compared this to direct NF from an amygdala ROI using fMRI.
Valid data were obtained from five volunteers (4 male, 1 female) who participated in the preliminary experiment. EEG was obtained using BrainAmp-MR EEG amplifier (Brain Products, Munich, Germany) and the BrainCap electrode cap with Ag/AgCl electrodes providing 30 EEG channels, and additional ECG and EOG channel (Falk Minow Services, Herrsching-Breitbrunn, Germany). EEG was sampled at 5 kHz and recorded using the brain vision recorder software (Brain Products).

The EEG fingerprint of the amygdala was computed as described in [28]. It was sampled continuously during a 5 minutes baseline session. The participants were instructed to sit still while no external stimuli were introduced to them. The EFP computed during this session was used for standardising data collected during forthcoming NF sessions. During these NF sessions, the standard score of the EFP relative to the baseline distribution served as a target for the NF manipulation.

Each of the NF sessions included four baseline epochs of passive viewing of the waiting room environment (30 sec each) and four blocks of active NF (60 sec each); except for one session of one subject that included only two NF blocks. In the first session, the participants were instructed to “appease the waiting room using their brain activity” (Figure 5). As with our previous case study, no preferred behavioural strategy was specified for NF, to avoid influencing subjects. In the last session, the participants were asked to foster unrest in the waiting room. One of the participants had double relaxation and agitation sessions.

The analysis of these preliminary data suggests that participants acquired ability to down-regulate their amygdala, while no evidence was found for a successful learning of up-regulation of this region. A successful relaxation block was defined as a block during which amygdala EFP values were significantly lower than those recorded in the previous baseline block (two-tailed Student’s t-test; p < 0.05, corrected for multiply comparisons using False Discovery Rate [4]). Success was found in 43.3% of the relaxation blocks, in comparison to only 10.71% of the agitation blocks during which EFP was lower than the previous baseline block ($\chi^2 = 7.71, p = 0.005$ in a chi square test of proportion difference). No difference was found between the conditions in the proportion of blocks during which the EFP was higher relative to the previous baseline (3.5% and 3.3% in the agitation and relaxation blocks, respectively).

CONCLUSION AND PERSPECTIVES

We have presented two examples of EEG-based NF systems for emotion regulation, embedded in realistic virtual reality applications. These have explored intervention at various levels of integration, amygdala and PFC, on different types of emotional responses. Despite these being proof-of-concept demonstrators, results obtained are encouraging, in particular the relatively high success rate considering minimal training received by subjects.

The limitations of the approaches presented here derive from well-known issues in NF. The first one corresponds to inter subject variability and the potential limitations to NF known as “BCI illiteracy” [44], the difficulty for some subjects (up to 30% depending on studies) to succeed at NF at all.

On the technical side, we have reported experiments still carried out in a laboratory setting. The development of augmentation mechanisms will require technologies that can operate in real-world conditions, to be incorporated, for instance, in ubiquitous computing systems; some research is starting to address this issue [29].

NF is receiving increasing interest as a BCI technique, and is poised to become one of the techniques of choice for emotional control. Despite encouraging results obtained in a variety of early (clinical) applications, many problems
require additional attention. It remains unclear whether NF should be considered a skill and the extent to which it can be developed from pre-existing abilities. In the specific field of emotional control, personality traits that determine baseline features of control structures may also play a role. NF techniques are also in need of significant development: our understanding of its mathematical properties (control aspects), and of mental strategies that can inform subjects instructions, all merit further investigation.

REFERENCES


