Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence

A Brain-Computer Interface to a Plan-Based Narrative

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

Eyal Soreq, Gal Raz,
Limor Ikar, Ayelet Or-Borichov,
Udi Ben-Arie, Ilana Klovatch,
Talma Hendler
Functional Brain Center
Tel Aviv Sourasky Medical Center
Tel Aviv, Israel
hendlert@gmail.com

Abstract

Interactive Narrative is a form of digital entertainment heavily based on AI techniques to support narrative generation and user interaction, significant progress arriving with the adoption of planning techniques. However, there is a lack of unified models that integrate generation, user responses and interaction.

This paper addresses this by revisiting existing Interactive Narrative paradigms, granting explicit status to users' disposition towards story characters as part of narrative generation as well as adding support for new forms of interaction. We demonstrate this with a novel Brain-Computer Interface (BCI) design, incorporating empathy for a main character derived from brain signals within filmic conceptions of narrative which drives generation using planning techniques.

Results from an experimental study with a fully-implemented system demonstrate the effectiveness of a EEG neurofeedback-based approach, showing that subjects can successfully modulate empathic support of a character in a medical drama. MRI analysis also shows activations in associated regions of the brain during expression of support.

1 Introduction

Among all applications of AI to digital entertainment, Interactive Storytelling (IS) has emerged as one that requires a wide range of techniques for plot generation, user interaction, character behavior, and presentation management. A key challenge is that these techniques should be able to capture the various affective characteristics of the narrative experience. Planning has established itself as one of the most successful technologies for narrative generation [Porteous *et al.*, 2010; Riedl and Young, 2010] it has also supported discourse presentation [Jhala and Young, 2005] and some affective components of the narrative [Gratch, 2000] or user responses, for instance, suspense [Cheong and Young, 2006]. It would appear that departing from a plot-centric approach and considering a user's affective response to generated stories provides a stronger basis to develop Interactive Narrative technologies.

Recent research in media psychology has emphasized the central role of characters in the affective response of users and the overall entertainment experience. This suggests that direct interventions on the bond between user and character could not only provide a powerful interaction mechanism, but one that would be fully aligned to the user response. This bond between users and story characters has generally been characterized as empathy, despite the many different interpretations of the concept. Major filmic theories [Frijda, 1988; Tan, 1996; Vorderer *et al.*, 2004; Janicke and Raney, 2012] have embraced a similar conception of empathy.

In previous work [Gilroy *et al.*, 2012], we have investigated the use of peripheral physiological signals (GSR and facial EMG) as an input modality to an Interactive Narrative system. This approach was implemented in a prototype which was able to produce narratives of up to 8 minutes in duration. However, in addition to the limited correlation between physiological signals and affective dimensions, one major limitation of this work was the lack of conceptual integration between the affective computing model, the user response, and the filmic strategy adopted by the narrative generation process.

We were thus in search of a physiological mechanism, possibly neural in source, that could more directly relate to empathy, attachment or disposition, and that could be measured in real-time, so as to be usable as an input mechanism. Numerous studies that correlate affective responses to EEG signals in the alpha band (8–12Hz), have led to the development of a prefrontal asymmetry metric [Henriques and Davidson, 1991] that can be used to characterize modulation of affective response. Frontal asymmetry has since been extensively studied (e.g., [Allen *et al.*, 2001; Hammond and Baehr, 2009], and related to empathy [Tullett *et al.*, 2012] as well as positive thinking [Allen *et al.*, 2001]. This has led us to consider it as a measure of dis-

*corresponding author

position towards story characters, which could be the basis for user input, provided it could be captured in real-time as part of an evolving narrative. This is indeed suggested by the finding that frontal asymmetry can be controlled through neurofeedback (NF) of EEG [Rosenfeld, 2000]. Furthermore, an EEG-NF approach is well suited to an IS application, since its voluntary nature is adapted to user intervention, and feedback mechanisms can be embedded into the multimedia staging of the narrative itself.

In this paper, we revisit Interactive Narrative from an empathy perspective using a Brain-Computer Interface (BCI) for user interaction. Our objective was to allow users to support their favorite character through the power of their thoughts, by mentally expressing their support and empathy (Figure 1). This is an active process rather than the analysis of passive emotional response; it is however compatible with recent neuroscience findings about filmic response. This experiment thus realizes a unified approach between an affective filmic theory (Tan's character empathy [Tan, 1996]), a character-based narrative generation technique [Porteous *et al.*, 2010], and a BCI mechanism compatible with empathy (pre-frontal alpha wave asymmetry as proposed by [Henriques and Davidson, 1991]). We have created a baseline Interactive Narrative based on a Medical drama, which features a junior female doctor facing a variety of challenges in her work, both personal and professional. The evolution of generated stories depends on user intervention—if it is successful the doctor manages to overcome these challenges whilst in the absence of successful user intervention the story spontaneously evolves towards the character's demise.

2 Previous and Related Work

BCIs have, from their inception, been used in conjunction with interactive media (i.e., video games), but mostly from the perspective of an interface technology either in an entertainment perspective [Nijholt *et al.*, 2009] or for therapeutic applications [Pope and Palsson, 2004], with little exploration of the relationship to the actual media content itself. A critical analysis of the performance of existing BCIs [Lotte, 2012] has led to emphasis on user training and the increasing relevance of NF as an implementation paradigm. This was demonstrated in a commercially available game environment in AlphaWoW [Plass-Oude Bos *et al.*, 2010], which used a version of *World of Warcraft*. However, this only went as far as replacing a single control variable (switching between two character forms) with an NF interface based around relaxation. The OpenViBE framework [Renard *et al.*, 2010] has been created for interpreting EEG signals in both medical and entertainment settings, including an NF game-like therapeutic system built on the Unity engine for ADHD therapy.

Recent work in Neuroimaging has provided evidence for specific activation pathways corresponding to a range of empathic responses when viewing films of high emotional content [Raz *et al.*, 2012], while Tikka et al. [2012] have proposed a similar approach using BCI, while not reporting an implementation of the system. Meanwhile Hand and Varan [2009] showed that there was increased empathy towards characters in interactive drama and narrative-based advertisements compared to passive versions, and suggest a strong narrative structure is required to allow this empathy to develop.

To our knowledge, the implemented interactive application which we introduce in this paper is the first to combine the specific concepts of empathetic disposition, interactive narrative generation and frontal asymmetry NF. Frontal asymmetry in particular has been identified previously as an element of a model of intrinsic affect evident while playing games [Reuderink *et al.*, 2013] and used as part of longer-term NF ther-

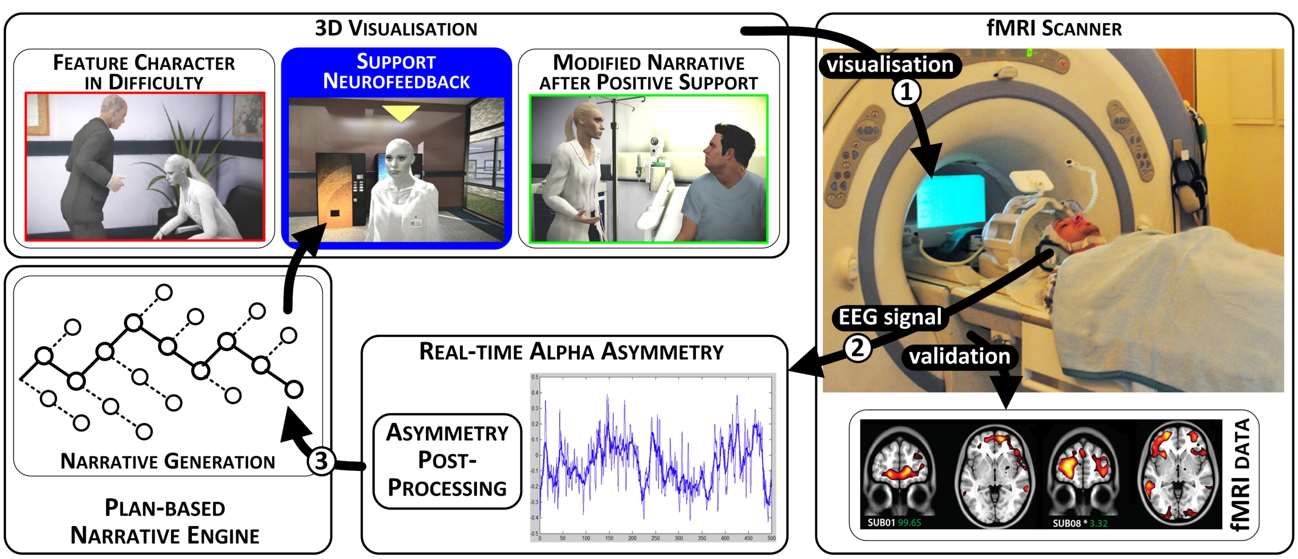

Figure 1: System Overview and Experimental Setting: (1) the user watches the narrative generated in real-time from inside an MRI scanner; (2) BCI input is interpreted in terms of empathy towards the feature character; (3) successful neurofeedback influences the narrative, improving the character's situation.

apy [Hammond and Baehr, 2009], but we take the next step in creating a BCI placed in an entertainment setting aligned to an intentional interpretation, utilizing strong narrative structure and characters that viewers can relate to.

3 System Overview

We have developed a fully implemented system in which a BCI based on neurofeedback of frontal asymmetry is integrated in to the control of an Interactive Narrative that we configured for experimental validation via the use of MRI. The overall architecture of the system in use is shown in Figure 1.

The interactive storytelling system is an extended version of our previous implementation based around a medical drama [Gilroy *et al.*, 2012]. This interactive narrative offers multiple opportunities for varied levels of tension and conflict across the unfolding story. Using a principled planning approach, the narrative engine generates sequences of individual narrative actions which are staged within the Unreal® 3D game engine (Unreal Development Kit). Narrative actions are visualized in real-time by selecting appropriate animations and camera placements to produce compelling dramatic scenes which match the intended affective tone. To enable experimental validation, the resulting visualization presented to the viewer was projected onto a screen positioned in front of the viewer while they lie in an MRI scanner.

The feature character in our medical drama is a junior female doctor who faces adversity as the narrative unfolds. The realistic nature of the visuals, combined with the use of filmic conventions in shot selection and camera placements, facilitates the induction of the appropriate empathetic feelings for the feature character. Beyond the generation of compelling graphics, specific periods of user interaction through NF are explicitly indicated to the viewer by the 3D visualization engine, in which they are prompted to "support" her (as illustrated in the scene labeled "Support Neurofeedback" in Figure 1).

Interaction consists of an NF process, in which the success of attempted user support measured by the BCI is visualized to the viewer in real-time (described in Section 5), and the subsequent affect on the narrative depends on the final level of user support at the end of a period of NF. Support is calculated from EEG measurements of frontal asymmetry, which in experiments we compared to MRI data measurements captured in experiments for offline validation of the approach. Support output is mapped to fluent values in the planning domain, while the NF periods themselves are staged using the same visualization mechanism as the main narrative. Neurofeedback is initiated for limited periods if, while iterating through the staging of the current sequence of narrative actions, the progression of the feature character has deteriorated below a pre-defined threshold. The updated value of user support is added to the current state of the narrative domain and provides the potential for narrative re-planning.

This closes the loop of affective response and empathetic thought by making sure the story progresses antagonistically for the main character unless the viewer intervenes, giving the viewer the agency to act upon their concern. If their support is not enough to alter her course through the narrative, she is still shown in adversity, and the viewer is given further opportunity to support her again.

4 Character-Based Narrative Generation

In order to generate narratives that can integrate user support a mechanism was required for specification of story control information, so that generated stories would be initially

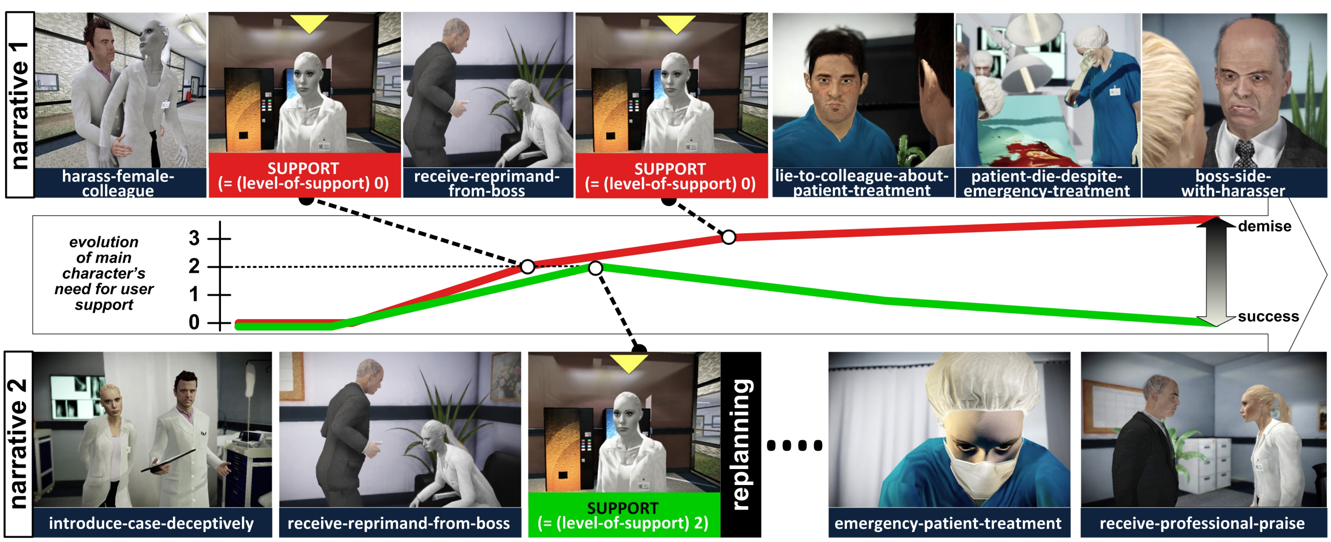

Figure 2: Excerpt from the Interactive Narrative. The feature character faces a difficult situation *(receive-reprimand-from-boss)*, which creates an opportunity for the user to support her; BCI NF input is mapped into the planning domain *(= (level-of-support) 2)*; successful support modifies the course of the narrative, leading the character to take the right decisions—the junior doctor is eventually praised for saving her patient *(receive-professional-praise)*.

```
(sometime-before
    (patient-outcome Jones) (humiliated DrDixon))
(sometime-before
    (doctor-outcome DrDixon) (harassed-by DrDixon DrAdams))
(at-end (doctor-outcome DrDixon))
(at-end (patient-outcome Jones))
```

Figure 3: Sample PDDL3 modelling of narrative landmarks.

skewed towards the demise of the feature character. In addition the planning representation needed to be capable of capturing the valence of actions that support or degrade the feature character.

Narrative control was achieved via the use of landmarks, as introduced in [Porteous *et al.*, 2010], which provide a general mechanism to ensure the inclusion of dramatic content within a narrative in a way that enables control over story order and progression. Within the system, this landmarks-based approach provides a means to control the evolution of different story variants to show the early demise of the feature character, followed by further deterioration towards a negative ending should user support be unsuccessful. If the user successfully supports the character, the same mechanism is used to evolve the story towards a positive ending for the feature character.

The landmarks control is used in combination with representational aspects in order to dynamically generate invocation of NF periods for user support, with an initial attempt made once the user's situation has passed a threshold value and a second call generated should the feature character's situation deteriorate further (the consequence of failed user support as the story evolves towards its original negative ending). The dynamic triggering of these user support requests is discussed further in section 4.1 and illustrated in Figure 2.

In terms of representation, landmarks are facts representing narrative situations of interest and for the medical drama genre they might include such things as tense clinical situations, strained relationships between characters, deceptions and confrontations. The landmarks and partial orders over them, are specified as part of the domain model using PDDL3.0 modal operators and then used in a decomposition based planning approach to control the shape of a narrative trajectory: the landmarks are linearized and used to decompose the process of narrative generation into a sequence of sub-narratives, with the complete output narrative produced by conjunction of the sub-sequences. To ensure variation between generated (and re-generated) narratives, planning problem instances are created at run-time using non-deterministic selection of initial state facts and landmarks from sets of candidates. This ensures the generation of narratives containing suitable dramatic content and promotes story diversity. As an illustration, Figure 3 shows landmark facts and orders used in generation of the narrative example in Figure 2.

To implement the character support paradigm, generated stories must continue adversely in the absence of user intervention. This is achieved through representational means, whereby actions are categorized according to their potential to create difficult situations for the feature character, (i.e., in terms of the facts they achieve). For example, the action *receive-reprimand-from-boss*, shown in Figure 4, is adverse to the feature character, and thus can appear in phases of narrative that show their demise. In contrast actions such as *receive-professional-praise* (Figure 4) improve the character's situation. For our domain model, amongst a baseline set of 150 narrative actions, 5% are clearly adverse to the feature character, 20% are supportive and the remainder are neutral but acquire their significance in context.

```
(:init
    (= (in-difficulty) 0) ...
    (= (level-of-support) 0) (= (supported) 2) (...) ...)
```

Adverse	(:action receive-reprimand-from-boss :parameters (?d - doctor ?b - boss) :precondition (and (missed-work-deadline ?d) ...) :effect (and (increase (in-difficulty) 1) (humiliated ?d) ...))
Supported	(:action receive-professional-praise :parameters (?d1 ?d2 - doctor ?p - patient) :precondition (and (>= (level-of-support) (supported)) (emergency-treatment ?d1 ?p) (patient-ok ?p) ... :effect (and (doctor-outcome ?d1) ...))

Figure 4: Example Narrative Actions. Adverse actions such as *receive-reprimand-from-boss*, modify *in-difficulty* (initially 0). Actions such as *receive-professional-praise*, can feature in a narrative only if *level-of-support* matches a per-action threshold comparison to the *supported* fluent. The *level-of-support* (initially 0) is updated following successful NF.

With this representational approach, the combinatoric nature of narrative generation is preserved, since the configuration of states considered at run-time comes from the entire set of actions, rather than just those specifically tagged as adverse or supportive. Consequently, a domain model of this size is typically able to generate hundreds of different baseline stories with the number increasing with successful user intervention.

4.1 Triggering User Support

Unlike some systems that allow anytime user intervention, within this paradigm interaction is enabled a limited number of times during the unfolding narrative, due to the demanding nature of NF, which requires the user to concentrate in a manner that is difficult to successfully maintain for long periods of time. To implement dynamic generation of support opportunities, the planning domain model has been augmented to include the following numeric fluents: *in-difficulty*, which characterizes the feature character's situation; and *level-of-support*, which records the measure of user support as detected through NF.

The fluent, *in-difficulty*, is incremented by actions which have adverse consequences for the feature character (such as being *humiliated* by *receive-reprimand-from-boss* as shown in Figure 4). Consequently, facts such as these are used as

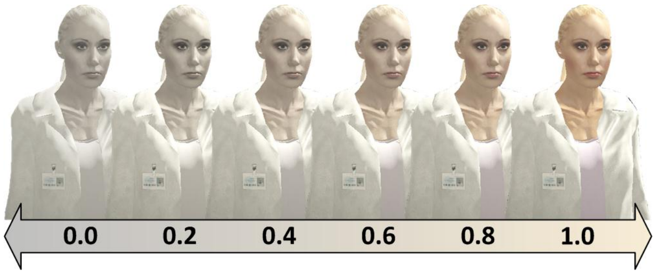

Figure 5: The character's appearance is used as a neurofeedback channel (color saturation is mapped to the alpha asymmetry score, A_2).

landmarks which are constrained to occur early in the partial order, thus ensuring that narrative generation is initially biased towards the demise of the feature character (e.g. all linearizations of the landmarks in Figure 4 place *humiliated* and *harassed-by* at the beginning of the total order).

The process of determining these user support opportunities is embedded in the overall planning loop, by assessing the numerical value of the *in-difficulty* fluent, which is increased by adverse events. Once it reaches a threshold value, a period of NF is started, as described in Section 5.

As shown in Figure 2, an initial request is made when the threshold value of *in-difficulty* is reached, and a second chance to support the character is offered should a further increase be detected in the value of this fluent.

The fluent *level-of-support*, is directly updated by a dummy operator whose effects are the results of the user support returned at the end of the NF period. The response of the narrative generator depends on whether the user was successful at supporting the character reflected in the numerical value of *level-of-support*. If the user's support can be interpreted as successful, the remainder of the narrative is immediately regenerated by re-planning using a planning problem instance revised to include supportive landmarks, and the current state of the narrative world which now includes the updated *level-of-support*. This will redress the course of action to favor the feature character. Should the user fail to provide sufficient support through NF, the original narrative resumes its execution leading to a negative ending for the feature character. These differing outcomes are shown in Figure 2.

5 Frontal Asymmetry Neurofeedback

As our BCI design relies on NF of frontal asymmetry measured through EEG, we have adapted the asymmetry score A_2, derived from work conducted by Henriques and Davidson [1991] and further refined and implemented in an NF depression paradigm by Hammond and Baehre [2009]. In this, successful positive thinking is characterized by greater relative *left* frontal activity in the brain, which can be derived from alpha rhythm signals.

As α-rhythm (8–12Hz), reflects cortical hypoactivity, increase in left frontal activity corresponds to a positive A_2 score (which we measure as the ratio of $F_4(R)$ & $F_3(L)$ electrodes with a reference electrode at position FCz, using the 10–20 electrode placement standard). The EEG-NF mechanism involves the viewer modulating this activity using mental thoughts of empathetic support, attempting to achieve the highest ratio of *left* vs. right cortical activity they can (i.e., a positive A_2 score approaching 1).

Neurofeedback visuals are expressed as the saturation of the graphical representation of the feature character, normalized from 0.0 (de-saturated) to 1.0 (rich saturation). This is illustrated in Figure 5. When NF is invoked, the main character's situation has been adversely affected, as detected via the *in-difficulty* fluent. In the normal visualization of the current action the character's appearance is "faded out" to a completely de-saturated appearance or "grey", as a prompt to the impending "support" NF section (as shown in the scene labelled "Feature Character In Difficulty" in Figure 1). To preserve visual consistency, and to effectively visualize the "fading" within the narrative (representing the character's distress reaching critical levels), this "prompt" scene is a narrative action that features the main character in a prominent manner.

Neurofeedback itself takes place over a 30-second window, during which a unique visualization is shown, with the main character in mid-shot, (as shown in the scene labelled "Support Neurofeedback" in Figure 1). During NF, if the character's appearance remains de-saturated, this indicates the viewer has not successfully communicated their positive thought (i.e., has a below threshold asymmetry score). Saturation is increased as the asymmetry score increases, mapped through a sigmoid function, to avoid over-saturation. When the character is fully saturated with color and the viewer is able to maintain this (by successful modulation of a higher asymmetry score), this is recognized as successful support and generates the corresponding modification of the *level-of-support* fluent in the planning domain (e.g., the increase to *(= (level-of-support) 2)* as shown in Figure 2).

6 Experimental Study

We used simultaneous acquisition of EEG/fMRI where the information on activated loci gathered from the fMRI serves as a validation to brain areas (cortical and sub-cortical) involved in the EEG alpha-asymmetry NF during periods of attempted user support. We hypothesized that successful EEG-NF guided by the probe of alpha asymmetry during the support window will activate mainly prefrontal cortex areas that were previously accounted to be related to regulation processes. The fMRI validation overcomes the low spatial resolution of the EEG, thus providing a superior anatomical account of the NF success especially with regard to suspected regulation of deep cortical areas such as the prefrontal cortex (PFC).

Participants

Fifteen healthy volunteers (3 female, 3 left-handed) of mean age 29.38 years (S.D. 7.6) took part in the experiment, all with either perfect or corrected eyesight. Of these, two were discarded due to technical issues, and 1 was rejected subsequently because of severe EEG movement artifacts.

Sub	F	df	P	Sub	F	df	P
1^{**}	99.65	242	.00	13	1.56	242	.21
2^{**}	16.11	244	.00	14	0.6	244	.438
7^{**}	7.06	242	.00	10	0.02	240	.89
8^{**}	3.32	242	.04	$6^{\dagger}$	0.84	244	.36
9^{*}	3.00	244	.08	$5^{\dagger}$	5.57	242	.01
$12^{\diamond}$	1.94	244	.16	$4^{\dagger}$	15.48	244	.00

Table 1: *EEG NF relative change index.* **, * *- significant positive change (*$p < 0.1$, $p < 0.05$*).* $^{\diamond}$ *- borderline success.* † *- negative change, no effect on narrative.*

Brain Data Acquisition

EEG data was acquired using a 32-electrode MRI-compatible BrainAmp MR system[1]. Data was recorded at a sampling rate of 5000Hz and collected on a PC running RecView software[1] for gradient and cardioballistic artifact removal. Alpha band (8-12Hz) power was extracted online from electrodes F_3 and F_4 as mentioned in Section 5, sampled in 500ms windows. The mean A_2 asymmetry score was calculated for each window calculated, and this was used to drive NF visuals.

Simultaneous to EEG recording, subjects underwent fMRI measurement with a 3T GE scanner. Scanning of fMRI was based on the echo-planar imaging (EPI) sequence of functional T_2^*-weighted images (TR/TE/flip angle: 3,000/35/90; FOV: $20 \times 20\text{cm}^2$; matrix size: 128×128) divided into 39 axial slices (thickness: 3mm; gap: 0mm) covering the whole cerebrum. A T1-weighted anatomical scan was used for alignment.

Experimental Protocol

As narrative evolution for each subject is driven by neural activity during user support, the length of experimental runs was somewhat variable. A typical run consisted of a NF training session (∼4 min.), a narrative training session (∼8 min., running through an example narrative outside of the MRI), an active session (∼8 min.), and a replay session (∼8 min.). Additional MRI scans of around 20 min. were needed to measure brain anatomy. To determine the controllable asymmetry range for each subject we used the distribution of asymmetry scores from the training session, thus accounting for individual differences in baseline EEG trait asymmetry score.

The active session began with 60 seconds of blank screen followed by a narrative that contained up to two opportunities for support through NF (30 secs. each), dynamically generated by the system as described previously. The replay session consisted of the visualization of the narrative generated by the same subject during an active session, with the interaction mechanism disabled, thus serving as a control baseline for fMRI. This could only be determined after-the-fact due to the narrative variability.

6.1 Analysis

We operated under the assumption that the A_2 at baseline is a stable trait metric [Davidson, 2003] that can be shifted due to affective mental process during the active NF session. To characterize the relative change in the asymmetry during active support window we calculated for each subject the distance between baseline A_2 calculated from the rest period at the beginning of the active session, to both NF windows using a repeated-measures ANOVA.

When comparing this EEG measure against a user's ability to successfully change the narrative using the NF, five subjects had "successful" narrative outcomes combined with significant up-modulation in A_2 scores ($p < 0.1$, 4 with $p < 0.05$). This is shown in Table 1. With these subjects we can be confident that the successful use of the BCI was due to actual modulation of EEG. Another subject had borderline significant up-modulation (subject 12). An additional three subjects had successful outcomes in terms of narrative progression, but no significant up-modulation of A_2 scores, indicating some possible over-sensitivity in the calibration for those subjects (10,13,14). Three subjects showed significant negative relative A_2 scores, so were unsuccessful in the use of the BCI at all († in Table 1).

What we aim to show is that the true performance of the BCI is illustrated through the relative A_2 scores, and that no empathy-related affective activity was occurring when contra-indicated by the EEG scores. While the BCI itself still appears to possibly benefit from further tuning with regard to sensitivity, it provided the correct outcomes for significant changes in EEG.

MRI Analysis

Analysis of fMRI data was performed with the SPM5[2] MATLAB tool. This includes preprocessing of fMRI data: (a) slice timing correction to the middle slice, (b) correction for head movement by realignment of all images to the mean image of the scan using rigid body transformation with six degrees of freedom, (c) normalization of the images to Montreal Neurological Institute (MNI) space by co-registration to the EPI MNI template via affine transformation, and (d) spatial smoothing of the data to 6mm FWHM. Finally, the first six images of each scan were discarded to allow for T_2^* equilibration effects. Statistical analysis was based on individual maps of activation obtained from a general linear model (GLM). The GLM included regressors that model epochs of active support during the live narrative session and epochs during replay of the support sessions within the replay of the previously generated movie. All regressors were convolved with a canonical hemodynamic response function (see model response in figure 6 B.). To reduce the effect of physiological artifacts and nuisance variables, six motion parameters were introduced as covariates in the model. T-statistical maps were obtained by contrasting hemodynamic responses during epochs of active support versus replay of these epochs.

7 Results

For further analysis we are comparing two groups: *successful*—those who significantly increased the A_2 score during the support period, and *unsuccessful*—those who did not modulate it significantly or in fact, reduced it. Since

[1] Brain Products GmbH. http://www.brainproducts.com

[2] http://www.fil.ion.ucl.ac.uk/spm

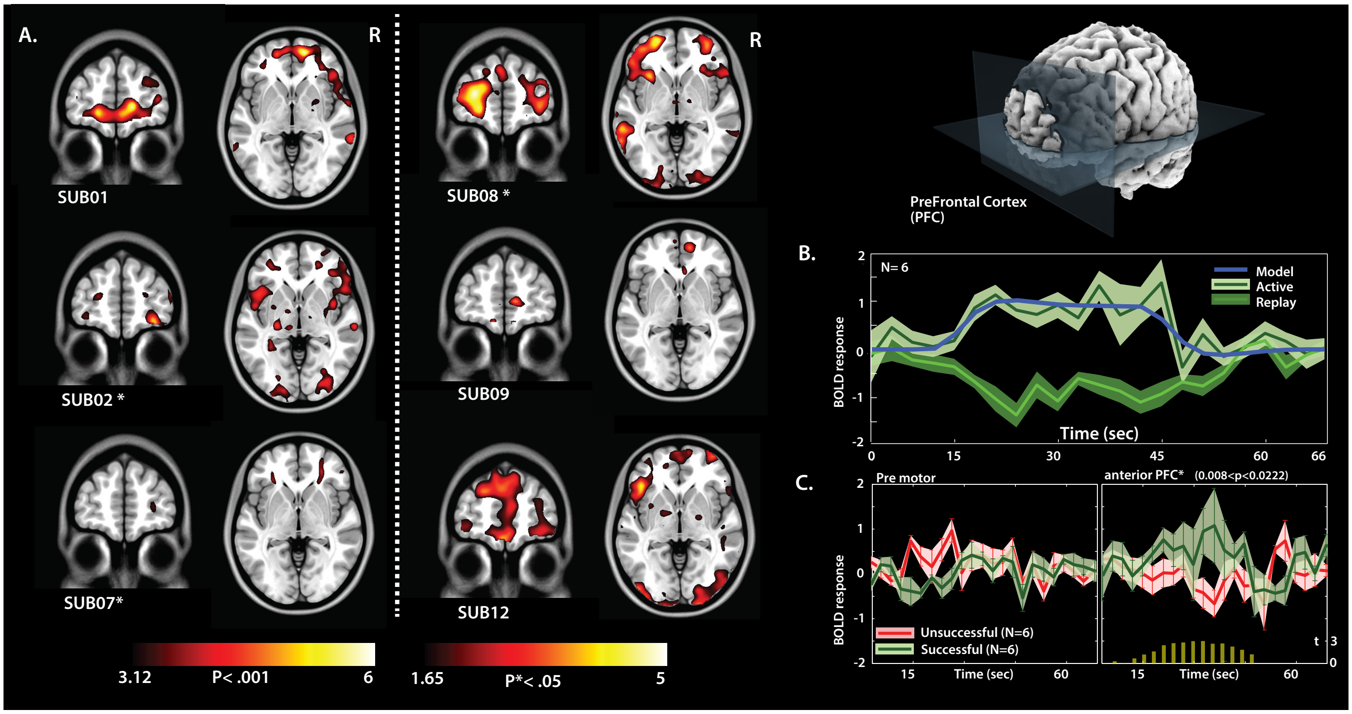

Figure 6: fMRI validation of the frontal asymmetry NF paradigm. **A.** Slice views (coronal and horizontal as indicated in top-right) of fMRI activation maps overlaid on a template anatomical scan (SPM5). Slices are shown for 6/12 participants who were highly successful in modulating their EEG alpha asymmetry index during active support periods. The parametric activation maps were obtained by whole brain contrasts of active more than replay ($p < .001$, $p* < .05$). **B.** Time course of averaged estimated effect ($n = 6$) obtained from the contrast of active vs. replay from peak activation in a priori region of interest in the right medial prefrontal Cortex (see 3D location, top-right). **C.** Comparisons between successful and unsuccessful participants (green and red plots, respectively) in % signal change during the support period obtained from a relevant region of interest localized at the vmPFC, and from a non-relevant region in the premotor cortex. There is a significant difference in BOLD response between the groups for the vmPFC, with successful group showing greater change (yellow squares indicate a sliding window of 24 sec in variable significance $0.006 < p < 0.0324$), while the premotor area shows little difference in activation.

the BCI principle is a priori focused on changing the narrative positively, for validation of the method, we concentrated on the six individuals who were successful in up-modulating their A_2 score as well as having a positive narrative outcome.

7.1 Behavioral Validation

All of the subjects correctly identified both the antagonist and protagonist. Not surprisingly the successful group perceived that they had a greater influence on the outcome of the unfolding narrative. Additionally, we inspected the reported subjective state of all subjects, the consensus emotion in the unsuccessful group was frustration (4 out of 6), while the successful group reported more approachable behavior (i.e., empathy and positive emotions).

Subjects quite clearly identified the protagonist of the story as "kind" and the antagonist as "vicious", the only dissenting opinion being two subjects who characterized the feature character as "neutral" and not "kind". Personal perception of the extent to which the viewer was helpful or able to make a difference to the story was split with successful subjects agreeing that they were helpful to the main character and had an impact on the story.

These results, along with informal feedback, indicate that subjects did understand the dynamics of the narratives and that subjective perception of their effectiveness was aligned with successfulness of response as indicated by MRI.

7.2 Neural Validation of the Interactive Experience

Whole-brain General Linear Model (GLM) analysis of the fMRI on the 6 individuals who were successful in A_2 up-modulation revealed enhanced activation during the periods of user support via NF relative to the same periods during passive replay in a cluster of regions in the pre-frontal cortex (PFC). These prefrontal loci include anterior and medial aspects of Brodman Area 10 and 11, known to be involved in cognitive and emotional control processes. Figure 6 shows the significant increased activation obtained in these PFC loci, confirming that successful up-modulation of EEG alpha asymmetry resulted in relevant regional recruitment. The whole-brain GLM analysis also provided additional indications for successful support-related regional activation in the middle temporal gyrus and the anterior insula. Only three of the successful supporters activated these regions at a threshold of $p < 0.001$ (uncorrected), but none of the unsuccessful supporters did so.

Intriguingly, signals obtained from the peak of activation within the PFC in each of the successful participants suggests that they not only increased their activity during the active user support, but also decreased it during the same period of the replay session (see Figure 6 B). To test the anatomical specificity of this regional effect we calculated time courses of activation during the active support window for each group in two distant loci: one in a task relevant area in the anterior aspect of the PFC (BA10, MNI: 26, 58, 6, selected based on the overlap of successful activation maps at $p < 0.05$). The other in a non-task relevant area in the right pre-motor (BA 6, MNI: 56,6,48, selected based on the overlap of unsuccessful activation maps, at $p < 0.3$) (see Figure 6 C). A direct comparison between these traces showed that only activation change in the PFC loci clearly distinguished between successful and unsuccessful individuals (sliding-window independent t-test $0.008 < p < 0.0222$ FDR corrected).

8 Discussion

It has been previously shown that mPFC is intricately involved in emotional regulation processing [Shimamura *et al.*, 2013; Etkin *et al.*, 2011], including representation of emotional value [Kawasaki *et al.*, 2001; Lerner *et al.*, 2009], and contextual interference [King *et al.*, 2005]. Furthermore, we showed that the modulation of the A_2 EEG signal is not derived from premotor, a commonly used marker in BCI for motor control [Gunduz *et al.*, 2011; Ramos-Murguialday *et al.*, 2012]. This points to the uniqueness of our task in possibly recruiting internally generated empathic related processes rather than perceptual-attention driven processes. The role of empathy has attracted significant interest in media psychology as well as interactive narrative and has provided a unifying framework for narrative visualization, user experience and user interaction through NF.

In our experiments, we have endeavored to provide generic instructions to our subjects, such as "mentally supporting" the main character. As a consequence, users have reported various strategies for generating positive thoughts, but the fMRI validation of our experiments have confirmed the selective activation of medial pre-frontal areas known to be involved in cognitive affective control mechanism, to the exclusion, in particular, of any motor area.

Our results therefore indicate the possible regulation of directed empathetic support through the use of EEG-NF with minimal training and potential to use this in an integrated fashion within an Interactive Narrative system. With more sophisticated training strategies, and a greater number of training sessions and investigation of the most successful strategies for empathetic thought, a higher and more reliable success rate could be achieved.

From a methodological perspective our work is a proof-of-concept implementation of a Brain-Computer Interface to an Interactive Narrative. The use of EEG-NF facilitates the control of complex user responses, while preserving the voluntary element of interaction and being compatible with the diversity of narrative generation. Users adapted remarkably well to this new interaction mechanism, with more than half of them succeeding in modulating NF input after only two short training sessions preceding the actual experiments. This contrasts with previous work in clinical applications of frontal asymmetry NF that reported multiple training sessions over several days or weeks.

Considering that our experiments took place inside an MRI scanner, hardly a user-friendly environment, the level of success achieved is certainly promising. Subjects showed significant variability in their basal frontal asymmetry, which currently forces individual calibration to define thresholds for successful EEG-NF. This could potentially be mitigated by a more dynamic reaction to baseline asymmetry over time.

References

[Allen *et al.*, 2001] John J. B. Allen, Eddie Harmon-Jones, and Cavender James H. Manipulation of frontal EEG asymmetry through biofeedback alters self-reported emotional responses and facial EMG. *Psychophysiology*, 38(4):685–693, June 2001.

[Cheong and Young, 2006] Yun-Gyung Cheong and R. Michael Young. A computational model of narrative generation for suspense. In *Proceedings of the 21st National Conference on Artificial intelligence - Volume 2*, AAAI'06, pages 1906–1907. AAAI Press, 2006.

[Davidson, 2003] Richard J. Davidson. Darwin and the neural bases of emotion and affective style. *Annals of the New York Academy of Sciences*, 1000(1):316–336, 2003.

[Etkin *et al.*, 2011] Amit Etkin, Tobias Egner, and Raffael Kalisch. Emotional processing in anterior cingulate and medial prefrontal cortex. *Trends in Cognitive Sciences*, 15(2):85–93, 2011.

[Frijda, 1988] Nico H. Frijda. The laws of emotion. *American Psychologist*, 43:349–358, 1988.

[Gilroy *et al.*, 2012] Stephen W. Gilroy, Julie Porteous, Fred Charles, and Marc Cavazza. Exploring passive user interaction for adaptive narratives. In *IUI '12: Proceedings of the 2012 ACM International Conference on Intelligent User Interfaces*, pages 119–128. ACM, 2012.

[Gratch, 2000] J Gratch. Social situated planning. In *AAAI Fall Symposium on Socially Intelligent Agents - The Human in the Loop*, 2000.

[Gunduz *et al.*, 2011] Aysegul Gunduz, Peter Brunner, Amy Daitch, Eric C Leuthardt, Anthony L Ritaccio, Bijan Pesaran, and Gerwin Schalk. Neural correlates of visual–spatial attention in electrocorticographic signals in humans. *Frontiers in Human Neuroscience*, 5, 2011.

[Hammond and Baehr, 2009] D. Corydon Hammond and Elsa Baehr. Neurofeedback for the treatment of depression: Current status of theoretical issues and clinical research. In Thomas H. Budzynski, Helen Kogan Budzynski, James R. Evans, and Andrew Abarbanel, editors, *Introduction to Quantitative EEG and Neurofeedback: Ad-*

vanced Theory and Applications, pages 295–314. Academic Press, 2009.

[Hand and Varan, 2009] Stacey Hand and Duane Varan. Interactive stories and the audience: Why empathy is important. *Computers in Entertainment (CIE)*, 7(3):39:1–39:14, September 2009.

[Henriques and Davidson, 1991] Jeffrey B. Henriques and Richard J. Davidson. Left frontal hypoactivation in depression. *Journal of Abnormal Psychology*, 100(4):535–545, 1991.

[Janicke and Raney, 2012] Sophie H. Janicke and Arthur A. Raney. Exploring how we enjoy antihero narratives: A comparison of fans and nonfans of 24. Presented at the annual meeting of the International Communication Association, Boston, 2012.

[Jhala and Young, 2005] Arnav Jhala and R. Michael Young. A discourse planning approach to cinematic camera control for narratives in virtual environments. In *Proceedings of the 20th National Conference on Artificial intelligence - Volume 1*, AAAI'05, pages 307–312. AAAI Press, 2005.

[Kawasaki *et al.*, 2001] Hiroto Kawasaki, Ralph Adolphs, Olaf Kaufman, Hanna Damasio, Antonio R Damasio, Mark Granner, Hans Bakken, Tomokatsu Hori, and Matthew A Howard. Single-neuron responses to emotional visual stimuli recorded in human ventral prefrontal cortex. *Nature Neuroscience*, 4(1):15–16, 2001.

[King *et al.*, 2005] JA King, T Hartley, HJ Spiers, EA Maguire, and N Burgess. Anterior prefrontal involvement in episodic retrieval reflects contextual interference. *Neuroimage*, 28(1):256–267, 2005.

[Lerner *et al.*, 2009] Yulia Lerner, David Papo, Andrey Zhdanov, Libi Belozersky, and Talma Hendler. Eyes wide shut: amygdala mediates eyes-closed effect on emotional experience with music. *PLoS One*, 4(7):e6230, 2009.

[Lotte, 2012] Fabien Lotte. On the need for alternative feedback training approaches for bci. In *Berlin BCI Workshop on Advances in Neurotechnology*, 2012.

[Nijholt *et al.*, 2009] Anton Nijholt, Danny Plass-Oude Bos, and Boris Reuderink. Turning shortcomings into challenges: Brain-computer interfaces for games. *Entertainment Computing*, 1(2):85–94, April 2009.

[Plass-Oude Bos *et al.*, 2010] Danny Plass-Oude Bos, Boris Reuderink, Bram van de Laar, Hayrettin Gürkök, Christiam Mühl, Mannes Poel, Anton Nijholt, and Dirk Heylen. Brain-computer interfacing and games. In D. Tan and A. Nijholt, editors, *Brain-Computer Interfaces*, pages 149–178. Springer, 2010.

[Pope and Palsson, 2004] Alan T. Pope and Olafur S. Palsson. Helping video games "rewire" our minds. Technical report, National Aeronautics and Space Administration, Hampton, VA. Langley Research Center, 2004.

[Porteous *et al.*, 2010] Julie Porteous, Marc Cavazza, and Fred Charles. Applying Planning to Interactive Storytelling: Narrative Control using State Constraints. *ACM Transactions on Intelligent Systems and Technology (ACM TIST)*, 1(2):1–21, 2010.

[Ramos-Murguialday *et al.*, 2012] Ander Ramos-Murguialday, Markus Schürholz, Vittorio Caggiano, Moritz Wildgruber, Andrea Caria, Eva Maria Hammer, Sebastian Halder, and Niels Birbaumer. Proprioceptive feedback and brain computer interface (BCI) based neuroprostheses. *PloS One*, 7(10):e47048, 2012.

[Raz *et al.*, 2012] G. Raz, Y. Winetraub, Y. Jacob, S. Kinreich, A. Maron-Katz, G. Shaham, I. Podlipsky, G. Gilam, E. Soreq, and T. Hendler. Portraying emotions at their unfolding: A multilayered approach for probing dynamics of neural networks. *NeuroImage*, 60(2):1448–1461, 2012.

[Renard *et al.*, 2010] Yann Renard, Fabien Lotte, Guillaume Gibert, Marco Congedo, Emmanuel Maby, Vincent Delannoy, Olivier Bertrand, and Anatole Lécuyer. OpenViBE: An open-source software platform to design, test, and use brain-computer interfaces in real and virtual environments. *Presence: Teleoperators and Virtual Environments*, 19(1):35–53, February 2010.

[Reuderink *et al.*, 2013] Boris Reuderink, Christian Mühl, and Manne Poel. Valence, arousal and dominance in the EEG during game play. *International Journal of Autonomous and Adaptive Communications Systems*, 6(1/2013):45–62, 2013.

[Riedl and Young, 2010] M.O. Riedl and R.M. Young. Narrative planning: Balancing plot and character. *Journal of Artificial Intelligence Research*, 39:217–268, 2010.

[Rosenfeld, 2000] J. Peter Rosenfeld. An EEG biofeedback protocol for affective disorders. *Clinical EEG electroencephalography*, 31(1):7–12, 2000.

[Shimamura *et al.*, 2013] Arthur P Shimamura, Diane E Marian, and Andrew L Haskins. Neural correlates of emotional regulation while viewing films. *Brain imaging and behavior*, 7(1):77–84, 2013.

[Tan, 1996] Ed S. Tan. *Emotion and the Structure of Narrative Film: Film as an Emotion Machine*. Mahwah, N.J: L. Erlbaum Associates, 1996.

[Tikka *et al.*, 2012] Pia Tikka, Aleksander Väljamäe, Aline W. de Borst, Roberto Pugliese, Niklas Ravaja, Mauri Kaipainen, and Tapio Takala. Enactive cinema paves way for understanding complex real-time social interaction in neuroimaging experiment. *Frontiers in Human Neuroscience*, 6, 2012.

[Tullett *et al.*, 2012] Alexa M. Tullett, Eddie Harmon-Jones, and Mochael Inzlicht. Right frontal cortical asymmetry predicts empathic reactions: Support for a link between withdrawal motivation and empathy. *Psychophysiology*, 49:1145–1153, 2012.

[Vorderer *et al.*, 2004] Peter Vorderer, Christoph Klimmt, and Ute Ritterfeld. Enjoyment: At the heart of media entertainment. *Communication Theory*, 14(4):388–408, 2004.