Detecting and interpreting conscious experiences in behaviorally non-responsive patients

Lorina Naci*, Leah Sinai, Adrian M. Owen

The Brain and Mind Institute, Department of Psychology, University of Western ON, London N6A 5B7, Canada

ARTICLE INFO

Article history:
Accepted 24 November 2015
Available online xxxx

ABSTRACT

Decoding the contents of consciousness from brain activity is one of the most challenging frontiers of cognitive neuroscience. The ability to interpret mental content without recourse to behavior is most relevant for understanding patients who may be demonstrably conscious, but entirely unable to speak or move willfully in any way, precluding any systematic investigation of their conscious experience. The lack of consistent behavioral responsivity engenders unique challenges to decoding any conscious experiences these patients may have solely based on their brain activity. For this reason, paradigms that have been successful in healthy individuals cannot serve to interpret conscious mental states in this patient group. Until recently, patient studies have used structured instructions to elicit willful modulation of brain activity according to command, in order to decode the presence of willful brain-based responses in this patient group. In recent work, we have used naturalistic paradigms, such as watching a movie or listening to an audio-story, to demonstrate that a common neural code supports conscious experiences in different individuals. Moreover, we have demonstrated that this code can be used to interpret the conscious experiences of a patient who had remained non-responsive for several years. This approach is easy to administer, brief, and does not require compliance with task instructions. Rather, it engages attention naturally through meaningful stimuli that are similar to the real-world sensory information in a patient's environment. Therefore, it may be particularly suited to probing consciousness and revealing residual brain function in highly impaired, acute, patients in a comatose state, thus helping to improve diagnostication and prognostication for this vulnerable patient group from the critical early stages of severe brain-injury.

© 2015 Elsevier Inc. All rights reserved.

Introduction

Decoding the contents of consciousness from brain activity is one of the most challenging frontiers of cognitive neuroscience. Progress has been made in previous studies of healthy individuals. One common type of paradigm uses sensory (e.g., visual or auditory) stimulation to probe the neural code underlying conscious perceptual experiences. For example, Nishimoto et al. (2011) acquired fMRI data as healthy individuals viewed a movie and developed an encoding model for the mechanisms of dynamic visual representation in visual cortex. Another type of paradigm relies on measurement of external behavioral indexes for tagging the temporal evolution of brain activity, and interpreting whether it codes for perception, intention, or action. For example, Soon et al. (2008) used a button-press response to mark the time-point preceding the participants’ conscious intention to act. The authors demonstrated that brain activity at that time-point predicted subsequent decisions, even though participants were not conscious of any future action.

Yet, arguably, the ability to interpret mental content without recourse to behavior is most relevant for understanding patients who may be demonstrably conscious, but entirely unable to speak or move willfully in any way, precluding any systematic investigation of their conscious experience (Plum and Posner, 1966). Indeed, in most cases, it is impossible to know whether these patients are conscious or not. Upon failing to respond to commands in repeated clinical examinations, behaviorally non-responsive patients are deemed unaware of themselves or the environment, and diagnosed as vegetative state (VS) (The Multi-Society Task Force, 1994). In cases where patients show inconsistent but reproducible signs of awareness, they are diagnosed as minimally conscious state (Giacino et al., 2002). The clinical assessment of these patients’ inconsistent behaviors, which are often limited by motor constraints (Owen and Coleman, 2008), is difficult and results in high misdiagnosis rates (~43%) (Schnakers et al., 2009; Andrews et al., 1996; Childs et al., 1993). In the following sections, we review approaches to decoding conscious thought from brain activity, honing in to paradigms that can be used to detect and interpret conscious experiences without recourse to behavior in this patient group. We revisit the findings of Naci et al. (2014) and show some novel experimental results that extend the previous findings.

The lack of consistent behavioral responsivity engenders unique challenges to decoding any conscious experiences that non-responsive patients may have solely based on their brain activity. For this reason,
the aforementioned paradigms that have been successful in healthy individuals cannot serve to interpret conscious mental states in this patient group. In particular, in the absence of behavior, brain activity alone must serve as a marker of conscious awareness and, thus, it is necessary to strictly rule out that the activity in question is due to any other factors, such as exogenously-driven processes. Therefore, paradigms that probe brain activity in sensory cortex (e.g., as in Nishimoto et al., 2011), which reflects (at least in part) exogenous, stimulus-driven brain responses, cannot provide markers of conscious awareness for non-responsive patients. Moreover, in the absence of any behavioral manifestations of conscious states, paradigms that rely on behavioral indexes (e.g., Soon et al., 2008) cannot interpret the relation of brain activity to external events. Instead, independent means for the interpretation of brain activity in these patients are necessary.

Patient studies to date have made significant strides towards decoding the brain activity of non-responsive patients. The majority of studies have used structured instructions, such as “imagine playing tennis,” “imagine navigating around your house,” or “attend to the word ‘yes’/’no’,” to distinguish between stimulus- and endogenously-driven brain activity (Naci and Owen, 2013; Chen et al., 2013; Fernández-Espejo and Owen, 2013; Cruse et al., 2012; Bardin et al., 2012; Cruse et al., 2011; Bardin et al., 2011; Monti et al., 2010; Owen et al., 2006). When tested with these paradigms, a number of entirely behaviorally non-responsive patients have demonstrated willful modulation of brain activity according to task instruction, thus demonstrating their conscious awareness. Therefore, studies to date have established that brain activity can be used as a proxy for behavior, and that a significant minority of non-responsive patients are demonstrably aware, despite the lack of behavioral response (Cruse et al., 2011; Monti et al., 2010). However, the extent of preserved mental life in these patients, especially with respect to whether they experience the world similarly to healthy individuals, remains unknown.

Decoding executive function common to different individuals

To address this question, we recently asked (Naci et al., 2014) whether there is a common neural code that can account for how different individuals might form similar conscious experiences, and, if so, whether it could be used to interpret those experiences without recourse to behavior. Initially, we quantified the neural correlates of conscious experience common to different individuals using fMRI and behavioral investigations, and then applied the same approach in behaviorally non-responsive patients with unknown levels of consciousness, to examine and quantify their experience of the world in the absence of self-report.

In particular, we used executive function as an empirical window for quantifying human conscious experience. As an endogenously driven meta-cognitive process that is integral to our conscious experience of the world, executive function can serve as a marker of conscious awareness in non-responsive patients. Prior to this work, laboratory tests of executive function had not been related to the open-ended nature of conscious experiences. To solve this problem, we used a naturalistic paradigm and investigated executive function during movie viewing. By their very nature, engaging movies are designed to give viewers a shared conscious experience driven, in part, by the recruitment of similar executive processes, as each viewer continuously integrates their observations, analyses and predictions, while filtering out any distractions, leading to an ongoing involvement in the movie’s plot. Previous work with healthy participants had shown that, when different individuals watch the same movie, correlated changes of brain activity across them are observed (Hasson et al., 2010, 2008, 2004; Bartels and Zeki, 2004, 2005; Bartels and Zeki, 2005; Bartels et al., 2008). However, it remained unknown whether any of these correlated activity fluctuations reflect similar executive function across different individuals in response to the evolving executive demands of the movie plot.

Similarly to these previous studies (Hasson et al., 2010, 2004; Bartels and Zeki, 2004; Bartels and Zeki, 2005; Bartels et al., 2008), we found that healthy participants viewed a highly engaging short movie by Alfred Hitchcock – the so-called ‘Master of Suspense’ – in the fMRI scanner, they displayed highly correlated activity across the brain (Figs. 1A–D), including supraventral frontal and parietal regions, which support executive function (Barbey et al., 2012; Ptak, 2011; Duncan, 2010; Hampshire and Owen, 2006; Sauseng et al., 2005).

To investigate directly how the executive demands of the movie drove the fronto-parietal activity, we assessed the movie’s executive load in two independent behavioral experiments. First, the extent to which the movie made demands on executive functioning was quantified with a ‘dual-task’ procedure that has been used previously to investigate executive performance, both in patients with frontal lobe damage (Pettet et al., 2013; Baddeley et al., 1996; Vikki et al., 1996) and in healthy volunteers (Mizuno et al., 2012; Collette et al., 2005; D’Esposito et al., 1995). This framework assumes that, because executive function is a finite resource, in moments when the load on one execute-dly demanding task (i.e., the movie) is greatest, the performance of a second execute-dly demanding task will be impaired, yielding a direct, quantitative, measure of the executive demands of the first task across time.

Adopting this framework, we used a dual-task for determining the executive demands of the movie in a behavioral experiment, independently of the fMRI data, and subsequently, mapped them onto brain activity via statistical parametric modeling. The dual-task consisted of simultaneous performance of the Sustained Attention to Response Task (SART) – which measures sustained attention and quantifies executive function (Robertson et al., 1997) – and movie viewing. SART was an optimal choice of an execute-dly demanding task that could be performed while simultaneously watching the movie. It operates on the principle that insufficient attention to a task can result in slips of action, as automatic, unintended action sequences are triggered inappropriately. These automatic actions result in performance errors that can be detected in the SART reaction times. Specifically, in the SART, participants are required to respond with a button press to a series of randomly presented (‘go’) digits, but withhold responses to one pre-specified (‘no go’) digit. A signature of the SART is that a shortening of reaction times indicates a decrement in executive control, or, conversely, an increment in response automaticity (Whyte et al., 2006; Manly et al., 1999; Robertson et al., 1997). In particular, a shortening reaction time predicts an increased likelihood of a subsequent incorrect response to a ‘no go’ digit, and correlates with electrophysiological measures of waning attention (Datta et al., 2007).

A group of healthy participants (N = 27) simultaneously watched the movie and performed an auditory version of the SART. At the end of the dual-task, participants answered 14 multiple-choice questions, by selecting one of four answer options. These assessed each participant’s encoding of basic facts about the movie, and therefore, indirectly, their overall attention to the movie throughout its duration, as they simultaneously performed the SART. Group-averaged SART performance followed the canonical pattern. Reaction times in trials immediately preceding an incorrect button-press response (i.e., to a ‘no go’ trial) were significantly shorter than those immediately preceding correctly withheld responses (t(18) = 4.37; p < 0.0005), suggesting that errors were due to lapses of attention (failure to inhibit a response to a ‘no go’ trial). As the dual-task paradigm employed in the study measured interference between fundamentally similar cognitive processes, within the constraints of our dual-task, these errors are caused by competing executive demands that arise during movie viewing. Only such high-level cognitive processes could interfere with the mental processes recruited by SART, or conversely, by the very definition of the dual-task paradigm, any lower-level processes cannot cause interference, or performance decrements in the SART. Hence, the continual assessment of SART performance fluctuation while participants simultaneously viewed the movie did indeed yield a direct, continuous measure of the movie’s executive load.

Please cite this article as: Naci, L., et al., Detecting and interpreting conscious experiences in behaviorally non-responsive patients, NeuroImage (2015), http://dx.doi.org/10.1016/j.neuroimage.2015.11.059
In order to correctly map the dual-task behavioral performance on the fMRI data, only reaction times from subjects who successfully performed both tasks simultaneously were included. At the single-subject level, 15/27 participants showed both a SART effect ($p < 0.05$) and over 70% accuracy in the post movie questionnaire, and were deemed to have correctly performed both components of the dual-task. The group-averaged SART RTs from these subjects were included in the fMRI analysis. SART reaction times significantly ($p < 0.05$; FWE cor) predicted activity in a brain network involving frontal and parietal regions in the independent group of participants, who had watched the movie without a secondary task in the fMRI scanner (Fig. 2B). This result confirmed that correlated activity in these regions was driven by the executive load of the movie, as indexed by an entirely independent behavioral measure acquired in a separate group of healthy participants.

Second, to confirm that these modulations in frontal and parietal activities reflected a common conscious experience across individuals, we developed a qualitative measure of the subjective experience of viewing the movie that reflected, albeit indirectly, its ongoing executive load. In a second behavioral experiment, a different group of participants ($N = 15$) watched the movie, and rated how ‘suspenseful’ it was every 2 s from ‘least’ to ‘most suspenseful’ (Fig. 2C). Suspense ratings were highly similar [$t(14) = 25.3; p < 0.0001$] across different participants confirming their common experience. Moreover, the group-averaged suspense ratings significantly ($p < 0.05$; FWE cor) predicted activity in a similar fronto-parietal network in the independent group who had watched the movie in the fMRI scanner, with frames rated as ‘highly suspenseful’ predicting stronger activity in this network (Fig. 2D). This finding was consistent with a previous electroencephalography (EEG) study showing that inter-subject correlation of the EEG signal predicted the subjective experience of individuals watching audio-visual television content (Dmochowski et al., 2014).

In classic Hitchcock movies, such as the one used in this study, suspense arises through understanding the plots’ twists and turns, which are highlighted by the dramatic music soundtrack. To further test whether the perception of suspense indeed depended on comprehending the narrative, or whether it could be elicited on the basis of the soundtrack alone without any comprehension of the movie’s plot we acquired fMRI data in a fourth experiment (previously...
Nevertheless, as we argue above, these measures were motivated by a conceptual and methodologically rigorous approach, and their convergent findings corroborated our hypothesis. Both these independent measures explained activity in similar fronto-parietal networks that were statistically similar to the one observed in the movie fMRI data, as well as to the canonical executive network observed in a large number of neuroimaging studies (Yarkoni et al., 2011).

In summary, the results from the four experiments in healthy participants suggested that the movie’s executive demands drove brain activity in frontal and parietal regions, and, further, that the correlation of this activity across individuals underpinned their similar experience. Each individual’s fronto-parietal brain activity could be predicted from the rest of the group’s (Fig. 4), thus these data suggest that the fronto-parietal activity represents a reliable neural index of the similarity of his/her cognitive experience to the others’. This suggested that similar conscious experiences in different individuals are supported by a common neural code.

### Decoding conscious experience in behaviorally non-responsive patients

In a subsequent study, we used this approach to examine and quantify the movie watching experiences of two entirely behaviorally non-responsive, severely brain-injured patients with unknown levels of consciousness (Naci et al., 2014). Probing of brain function in individual brain-injured patients critically depended on the single-subject level reliability of brain responses in healthy participants. A prior set of leave-one-out analyses in healthy participants had shown that each participant’s auditory, visual, and fronto-parietal activities could be significantly (p < 0.05; FWE cor) predicted by the time-course of the corresponding activity in the rest of the group (Figs. 4A–C). The stereotyped brain activity underlying similar executive function across individual healthy participants enabled model-based predictions that could be applied to individual non-responsive patients.

FMRI data was acquired from the two patients, as they freely viewed the Alfred Hitchcock movie. We reasoned that, if one of the patients engaged in executive processing while watching the movie, he would exhibit similar brain activity patterns in frontal and parietal regions to the healthy participants. Conversely, we could use the healthy participants’ fronto-parietal activity as a benchmark for assessing the presence of
executive function, and, therefore, as an index of conscious experience, in two clinically similar, behaviorally non-responsive patients.

Initially we assessed auditory and visual functions to test whether the patients’ sensory processing was similar to the healthy controls. Subsequently the patients’ executive function was evaluated. The time-course of the auditory/visual/fronto-parietal network in the healthy group was used as a regressor in the SPM model of each patient’s movie data. Patient 2 showed brain activity in auditory and visual cortices that was correlated to that of the healthy group (p < 0.05; FWE cor) suggesting intact processing of both auditory and visual information in the movie (Fig. 5A). However, these responses cannot provide evidence of conscious awareness, as stimulus selective responses are observed in sensory regions in the absence of consciousness in anesthetized human (Macdonald et al., 2015) and nonhuman primates (Supér et al., 2001).

Importantly, activity in a network of frontal and parietal regions that is known to support executive processing (Barbey et al., 2012; Ptak, 2011; Duncan, 2010; Sauseng et al., 2005; Hampshire and Owen, 2006) was significantly correlated with that of healthy participants (p < 0.05; FWE cor). To further test whether the frontal and parietal activations observed in Patient 2 truly reflected executive processes related to the movie plot, we assessed the extent to which it correlated with the quantitative and qualitative measures of the movie’s executive load. Both of these measures, derived in healthy participants, significantly (p < 0.05; FWE cor) predicted the patient’s activity in the same frontal and parietal regions (Fig. 5). In summary, this patient demonstrated a highly similar fronto-parietal brain response to that of the three independent groups of healthy participants. The patient’s brain activity in frontal and parietal regions was tightly correlated with the healthy participants over time, and also, it reflected the executive demands of specific events in the movie, as measured both qualitatively and quantitatively in healthy individuals (Figs. 5B & C).

These neuroimaging results were striking in light of the patient’s enduring lack of behavioral responsiveness observed in repeated assessments at his bedside over the 16-year period. However, they aligned with the patient’s positive outcome in an independent command-following task, the results of which were unknown at the time of the movie experiment (Naci and Owen, 2013). However, it was impossible to determine, based on the results of the command-following task and the patient’s behavioral assessments, whether he maintained conscious experiences comparable to those of healthy individuals in response to real-world events in his environment.

In the previous sections, we have argued that in the movie paradigm the time-course of the correlated fronto-parietal activation pattern in healthy individuals, or the evolution of its peaks and troughs over time, is the code that represents at the neural level the cognitive experience of the movie’s executive demands. This code does not read off the precise details of a person’s thoughts. Rather, it can reveal whether two individuals have a highly similar cognitive experience when exposed to the same information. Thus, the time-course of the fronto-parietal activation provides a template for decoding whether patients who due to brain injury cannot provide self-report, have similar cognitive experiences to healthy individuals in response to the executive demands of the movie.

Therefore, the similarity of Patient 2’s fronto-parietal activation to that of healthy controls suggested that he experienced the movie similarly to the healthy individuals who watched it in the scanner. Importantly, the patient’s fronto-parietal activation was highly sensitive to the movie’s executive load determined by a second group of controls with a dual-task procedure outside of the scanner. This suggested specifically that the patient had a similar conscious response to the movie’s executive demands as healthy individuals. An independent study corroborated this interpretation. Naci et al. (2015) found that the fronto-parietal activity in response to naturalistic rich stimulation, such as movies, was extinguished in deeply anesthetized unconscious individuals, further suggesting that the patient’s brain responses in these regions could not be realized without the presence of covert conscious awareness. Moreover, the patient’s fronto-parietal activation was also highly sensitive to the subjective experience of the movie’s suspense, determined by a third group of controls. This suggested specifically

![Fig. 5. Decoding executive function in one behaviorally non-responsive patient. Healthy group: (A) group-level auditory (purple) and visual (blue) ICs. (B–C) The healthy group’s activity predicted by the quantitative (B) qualitative (C) executive measure (green) is overlaid on the group fronto-parietal IC (red); overlap areas are displayed in yellow. Patient 1: (A) the healthy group’s auditory IC predicted significant activity in Patient 1’s auditory cortex (purple). (B–C) No evidence of visual responses, or executive processing similar to the healthy participants was observed. Patient 2: (A) the healthy group’s auditory and visual ICs predicted significant activity in Patient 2’s auditory (purple) and visual (blue) cortex, respectively. (B–C) The quantitative (B) and qualitative (C) executive measures predicted activity (green) in the patient’s frontal and parietal regions. Overlap with activity predicted by the healthy group’s fronto-parietal IC (red) is displayed in yellow. Taken from Naci et al. (2014).](http://dx.doi.org/10.1016/j.neuroimage.2015.11.059)
that the patient had a similar subjective perception of suspense during the movie as healthy individuals. Sensitivity to the executively demanding and suspenseful moments of the movie depended on understanding of the movie plot, which required a sophisticated cognitive repertoire, including language understanding, working memory, reasoning capacities, and sustained attention.

Therefore, these data suggested not only that Patient 2 was consciously aware, but also that he had preserved cognitive abilities that enabled him to continuously engage in complex thoughts about real-world events unfolding over time. In this context, these are likely to include updating the contents of working memory (e.g., to follow the plot), relating events in the movie to past experiences (e.g., to appreciate that a gun is a dangerous weapon), and coding the foreshadowing cues (i.e., events that might have future relevance to the plot) characteristic of movies of this type (Naci et al., 2014).

Despite a highly similar clinical and behavioral profile, Patient 1 showed no fMRI evidence of executive processing. This was consistent with the lack of similarity between her visual response and that of healthy controls (Figs. 5A–C). Independent fMRI analysis showed that only a couple of isolated voxels in the patient's visual cortex responded significantly (p < 0.05; FWE cor) to basic visual properties, such as the visual contrast in the movie, suggesting that she retained very limited visual processing. Moreover, as the visual activity of all individual controls in the study showed a very high similarity (t(11) = 11.9; p < 0.00001) to one another, the lack of a comparable visual response in the patient strongly suggested that she had highly impaired visual function. This further suggested not only that the patient did not adequately process the movie's visual features, but also that consequently no audio-visual integration could take place, both of which were critical processing steps for understanding the film.

In summary, the patient's data suggested that Patient 1 had preserved auditory processing but highly impaired visual function, consistent with influential studies showing impaired visual function in a high proportion of patients with disorders of consciousness (The Multi-Society Task Force, 1994). Therefore, these data highlighted the need for developing auditory-only paradigms suitable for detecting executive function in covertly aware patients. However, it is impossible to determine based on these negative results, whether the patient was truly unconscious, whether she was aware but could not understand the movie due to impaired visual function, or whether she lacked complex visual function as well as conscious awareness. Moreover, negative findings in non-responsive patients must be interpreted with caution and cannot be used as conclusive evidence for lack of awareness, also because false negative findings in functional neuroimaging studies may sometime occur even in healthy volunteers, due to the normal variance in the signal to noise ratio of the BOLD measurement reflecting the experimental manipulation. Nevertheless, the aforementioned index of executive processing did reveal significant and similar changes in the fronto-parietal network in each and every healthy participant who watched the movie, suggesting that its neural signature is reliably present in all adult and conscious humans.

Caution must also be exercised when interpreting significant positive results that do align with a-priori predictions, as in the case of Patient 2. Spurious results in patients can sometime result from neuroimaging analyses, such as for example, from the normalization of a patient's native brain space to the healthy controls' template (Beisteiner et al., 2010; Crinion et al., 2007; Brett et al., 2001). However, the current approach did not involve normalization to a healthy template, nor did it constrain the patient's expected brain activity based on the localization of the effect in healthy controls. Instead, the time-course of brain activity in healthy controls served to build a strong prediction for the temporal evolution of brain activity in the patients, should they retain covert awareness. The precise location of a patient's brain activity was expected to deviate from that of the healthy controls. Not only is this naturally the case for individual healthy participants (e.g., Figs. 4A–C), but also, importantly, it is to be expected in brain-injured patients as a result of structural and concomitant functional reorganization of the brain. Nevertheless, a spatial heuristic based on the controls' data informed the interpretation of the patients' results, helping to infer the nature of the underlying residual brain function. In summary, drawing comparisons in the temporal domain enabled direct relation of the healthy controls' activation to that of brain-injured patients, while avoiding stringent spatial constraints on the patients' functional anatomy.

In fMRI studies such as these, due to the absence of external cues, the functional brain activity serves as the only indicator of the patient's mental states. Hence, it is especially important to achieve a high level of confidence in the interpretation of brain activity, in order to avoid false positives. In this study, converging results corroborated our interpretation of the brain activation observed in Patient 2. The Patient's residual brain function was assessed hierarchically, from sensory (e.g., auditory, visual) to higher order (e.g., executive) processes (Figs. 5A–C). As executive function in response to the movie depended on the integrity of more basic cognitive processes, such as visual and auditory perception, evidence for preserved structural and functional integrity along the cognitive hierarchy suggested that the patient was capable of integrated thought. Moreover, when assessing further the integrity of the patient's executive function, independent predictions from three independent control groups were considered. Converging patient results from these three independent tests provided strong evidence for the presence of executive function and helped interpret the patient's ongoing conscious experience during movie viewing.

Future directions

With the exception of the aforementioned study by Naci et al. (2014), to date, neuroimaging studies that have probed consciousness in behaviorally non-responsive patients have tested whether any given patient could follow commands, and therefore demonstrate conscious awareness, via his/her brain activity (Naci and Owen, 2013; Fernández-Espejo and Owen, 2013; Chennu et al., 2013; Cruse et al., 2012; Bardin et al., 2012; Cruse et al., 2011; Bardin et al., 2011; Monti et al., 2010; Owen et al., 2006; see Casali et al., 2013 for a different approach). However, the requirement that a patient must be able to produce brain responses as prescribed by study instructions, in order to demonstrate that he/she is aware, is likely too stringent for many patients who are aware, but, due to the effects of brain injury, fail to comply with structured instructions (Chung et al., 2013; Giacino et al., 2002; McDowell et al., 1997). The discrepancy between the high proportion of non-responsive patients (~43%) who are routinely misdiagnosed through bedside assessments (Schnakers et al., 2009) and those who are able to demonstrate willful brain-based responses (17–19%) (Cruse et al., 2011; Monti et al., 2010) suggests that command-following neuroimaging techniques may lack the sensitivity to detect conscious awareness in a subset of patients.

By contrast, the approach by Naci et al. (2014) is unconstrained by any task commands, but, rather, captures attention naturally, and, therefore, might be more effective for detecting conscious awareness. Similar naturalistic paradigms in the auditory modality may be particularly suited to testing highly impaired patients with very limited residual brain function, including chronic patients with absent visual function and patients in coma. Coma is an acute state of unconsciousness occurring immediately after a brain injury, during which the patient exhibits no evidence of arousal or awareness. Comatose patients are understood to be unarousably unconscious, do not open their eyes, and only exhibit reflex responses to stimulation (Young, 2009). Some coma patients go on to make a good recovery, while others progress into vegetative or minimally conscious states. The first days and weeks following a serious brain injury are a time of considerable prognostic uncertainty, and this complicates decisions faced by health care providers and families, including whether to continue or withdraw life-sustaining therapies. While pilot studies in comatose patients have
revealed residual cognitive function, no patient to date has demonstrated covert awareness through command following paradigms. Although naturalistic paradigms have not yet been used to test comatose patients, due to their low-effort nature, they are likely to be more effective for identifying the subset of these patients, who remain aware, yet have not recovered sufficient cognitive resources to modulate their brain activity according to commands.

With this in mind, we have developed in healthy individuals an auditory-only naturalistic paradigm for the assessment of executive function in patients with impaired vision or eyes-closed comatose patients (previously unpublished; see Supplementary data). fMRI data was acquired from a group of healthy individuals (N = 15) as they listened to a short audio-story and a baseline version of the same auditory stimulus. The audio-story (5 min) was an auditory excerpt from the opening scene of the movie ‘Taken.’ Similarly to the movie, in this audio-story both speech and sound effects play a critical role in the development of its plot. The baseline stimulus was created by spectrally rotating the frequencies in the intact stimulus in order to retain the spectro-temporal characteristics of natural speech while rendering it unintelligible.

When healthy participants (N = 15) listened to the audio-story in the fMRI scanner, they displayed highly correlated (p < 0.05; FWE cor) activity across the brain (Fig. 6A), including fronto-parietal regions, known to support executive function (Barbey et al., 2012; Ptak, 2011; Duncan, 2010; Hampshire and Owen, 2006; Sauseng et al., 2005). By contrast, significant (p < 0.05; FWE cor) inter-subject correlation during the baseline was restricted to the auditory cortex, and a small cluster in right inferior prefrontal cortex (Fig. 6B). Formal comparison between the intact and baseline stimulus revealed that the intact story elicited significantly (p < 0.05; FWE cor) more inter-subject correlation than the baseline bilaterally, in parietal, temporal, motor, and dorsal/ventral frontal/prefrontal cortex (Fig. 6C). These results confirmed that the inter-subject correlation elicited by the story in high-order supramodal regions could not merely reflect modulations in the feed-forward processing load imposed by variations in the auditory information, nor any automatic attention effects triggered by the similarity of auditory stimuli across participants.

To relate the correlated activity fluctuations to different and specific aspects of the audio-story experience we used sensor independent component analysis (ICA) (Calhoun et al., 2009; Beckmann and Smith, 2005, 2004). Group ICA revealed spatially distinct networks in sensory specific (i.e., visual, auditory, motor) cortex, and regions of the frontal and parietal lobes. To further confirm that this fronto-parietal network reflected correlated activity across different participants, we performed single-subject ICAs (Bartels and Zeki, 2004; Carties and Zeki, 2005; Cartels et al., 2008), which revealed a high correlation between the single-subject time-courses for the fronto-parietal component [t(14) = 11.82, p < 0.001; Fig. 7]. The slightly lower inter-subject correlation in the fronto-parietal as compared to auditory regions [t(14) = 16.94, p < 0.001] was consistent with a previous fMRI study showing that during naturalistic stimulation individual time courses were highly consistent in sensory projection cortices and more variable elsewhere (Malinen et al., 2007). By contrast, single-subject time-courses were not correlated during the baseline stimuli, confirming that the processing of higher-level properties of the audio-story, including its narrative, must be driving the correlated activity in these regions across participants. Future studies will specify how the executive demands of the audio-story drive these fronto-parietal activity patterns.

Subsequently, we determined whether this audio-story paradigm was suitable for generating reliable prediction at the individual level, which would suggest that it might be useful for testing individual patients in future applications. A set of leave-one-out analyses showed that 15/15 participants’ auditory and 14/15 participants’ fronto-parietal activity could be significantly (p < 0.05; FWE cor) predicted by the time-course of the corresponding activity in the rest of the group (Figs. 8A & B). These data suggested that the audio-story elicits common executive processes in individual healthy controls that can be used to probe executive function in individual patients.

In summary, like the movie paradigm, this auditory paradigm is easy to administer, brief, and does not require compliance with task instructions. Rather, it engages attention naturally through meaningful sounds and speech that are similar to the real-world auditory information in a patient’s environment. This approach is highly suited to a proportion of brain-injured behaviorally non-responsive patients who have lost visual, but retain auditory function (The Multi-Society Task Force, 1994).

Discussion

Decoding the contents of consciousness in behaviorally non-responsive patients poses unique challenges. Until recently, patient studies have used structured instructions to elicit willful modulation of brain activity according to command, in order to decode the presence of willful brain-based responses in this patient group. These studies have shown that despite the apparent absence of external signs of consciousness, a significant minority patients with disorders of consciousness can respond to commands by willfully modulating their brain activity (Naci and Owen, 2013; Fernández-Espejo and Owen, 2013; Cruze et al., 2012; Cruze et al., 2011; Monti et al., 2010; Owen et al., 2006), and even respond to ‘yes’/‘no’ questions, by performing mental tasks (Naci and Owen, 2013; Fernández-Espejo and Owen, 2013; Monti et al., 2010). However, the extent of preserved mental life in any such responsive patient remained unknown.

In previous work (Naci et al., 2014), we have used multisensory naturalistic paradigms, such as free viewing of an audio-visual movie, to demonstrate that a common neural code supports conscious experiences in different individuals, and moreover, that this code may be used to interpret conscious experiences. In this study, the term ‘neural code’ refers to a system that allows us to decode, or to recognize and interpret, the cognitive experiences of different individuals, and thus it may decipher information about an individual’s cognitive experience that would otherwise be inaccessible in the absence of self-report, as is the case for behaviorally non-responsive patients. Unlike previous methods (Naci et al., 2013; Soon et al., 2008; Kamitani and Tong, 2005) that have used behavioral states to determine how brain activity underlies

Fig. 6. Brain-wide inter-subject correlation of neural activity while listening to the audio-story and auditory baseline. (A) The audio-story elicited significant (p < 0.05; FWE cor) inter-subject correlation within primary and association auditory cortices. A small cluster was also observed in the right inferior prefrontal cortex. None was observed in dorsal prefrontal and parietal cortices. (B) The baseline elicited significant (p < 0.05; FWE cor) inter-subject correlation within primary and association auditory cortices. A small cluster was also observed in the left inferior prefrontal cortex. None was observed in dorsal prefrontal and parietal cortices. (C) The audio-story elicited significantly (p < 0.05; FWE cor) more inter-subject correlation than the auditory baseline derived from the same stimulus, in parietal, temporal, motor, and dorsal/ventral frontal/prefrontal cortices. Warmer colors depict higher t-values of inter-subject correlation.

Please cite this article as: Naci, L., et al., Detecting and interpreting conscious experiences in behaviorally non-responsive patients, NeuroImage (2015), http://dx.doi.org/10.1016/j.neuroimage.2015.11.059
conscious states in healthy individuals, this approach interprets brain activity, and concomitant mental states, without recourse to an individual's behavior. Thus, it is uniquely suited to investigating conscious experience in individuals whose status as conscious agents is uncertain and cannot be tested through behavior or introspective report. When applied to patients with unknown levels of consciousness, this approach provided strong evidence that one patient, who had remained non-responsive for 16 years, was consciously aware and could continuously engage in complex thoughts about real-world events unfolding over time. Future patient cohort studies are necessary in order to establish the clinical prevalence of such highly preserved levels of awareness in this patient group.

For the purpose of clinical diagnosis, consciousness is defined as awareness of oneself and the environment as demonstrated through behavioral command-following (The Multi-Society Task Force, 1994). However, for a large proportion (~43%) of patients, behavioral testing lacks sufficient sensitivity, leading to an incorrect diagnosis (Schnakers et al., 2009; Andrews et al., 1996; Childs et al., 1993). This article presents

Fig. 7. Inter-subject correlation of neural activity within different functional brain networks. Time-courses of group-level ICA components (ICs) clustered into five spatially distinct brain networks, displayed in different colors on top of each bar. Each bar represents the group-averaged individual time-courses for each IC derived by single-subject ICAs. The single-subject time-courses of the auditory (purple) and fronto-parietal (green) ICs were significantly correlated across individuals.

Fig. 8. Predicting sensory-driven and higher-order processes during the audio-story in individual participants. Single-subject SPM analyses probed individual participants’ responses. The processing of the movie’s auditory (A), and executive (B) information in each healthy participant was significantly (p < 0.05; FWE cor) predicted by the time-course of the respective brain network in the leave-one-out group ICA. S1–S15 = Subject1–Subject15.
a novel approach that uses brain activity as a proxy for behavior, specifically to assess the dimension of consciousness that relates to an individual's awareness of the environment.

When the healthy controls' subjective experience of the movie's suspenseful features was assessed moment-by-moment (every two seconds), we found that their perception of suspense was very similar to one another, mirroring their highly correlated brain responses (Naci et al., 2014). Similarly to each healthy individual, the brain response of Patient 2 tracked the control's perception of suspense, thus showing specific sensitivity to environmental features on an ongoing and temporally precise basis, consistent with a previous EEG study in healthy controls (Dmochowski et al., 2012). Therefore, these results suggested that the conscious perception of some brain-injured non-responsive patients can track that of healthy individuals at a remarkably specific time-scale, as events in their environment unfold moment-by-moment.

New experimental findings extended the results of the audio-visual paradigm to auditory-only naturalistic conditions, such as unconstrained listening to an audio-story. Using this paradigm, it may be possible to identify comatose patients who exhibit clear neural signs of executive function, and therefore awareness, at a time when they are thought to be unconscious. Further, by testing patients longitudinally from the initial moments post injury to recovery, and comparing the residual cognitive profile of those patients who improve with those who deteriorate, we may identify objective prognostic markers that will improve treatment decisions and patient care in crucial early stages of brain injury.

Conclusions

These naturalistic paradigms may be suitable for investigating how brain function tracks clinical progression in other patient groups, such as those with psychiatric conditions. The ease of administration and patient engagement makes them advantageous for patient testing over conventional paradigms. Moreover, they can be advantageous on account of the widespread neural recruitment they elicit across the brain. An fMRI study found that ‘rich’ stimulation, such as free viewing of a movie, lead to stimulus-specific activation in more areas than with conventional stimuli (Bartels and Zeki, 2004). Therefore these naturalistic conditions may be uniquely suited to investigating how the interrelation between a number of brain systems, including sensory-driven (auditory, visual) and higher order (e.g., executive), leads to atypical behavior.

A previous study (Hasson et al., 2009) looked at fMRI responses across the entire cortex in adults with autism during free viewing of a popular audio-visual movie. It found that the neural activity of individuals with autism was characterized by idiosyncratic responses that were both highly variable across autistic individuals, and different from the responses observed within the matched controls. Future studies are needed to investigate whether these idiosyncratic responses reflect functionally distinct contributions and what the underlying brain mechanism might be. The study also found that, within each autistic individual, responses across repeated presentations of the same movie were unique and reliably replicable (Hasson et al., 2009). Thus, repeated testing with the same naturalistic stimuli might provide a window into whether altered perceptions are erratic or consistent and specific to certain individual patients or patient groups. When consistent responses to natural conditions are observed for individual patients, testing these individuals at different clinical time-points might further help to track changes or preservation of a patient’s condition. In these cases, it must be considered that significant reductions in neural correlation may occur naturally upon repeated exposures to the same naturalistic stimulus (Dmochowski et al., 2012).

Finally, future studies will advance the methodology for accurately classifying healthy and disordered function in cases where the differences in the BOLD measurement between clinical groups and controls, but also within a clinical group, are nuanced and subtler as compared to the all or none responses reported to date for severely brain-injured patients with disorders of consciousness.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgments

We thank Conor Wild and Rhodri Cusack for their help with the stimuli. This research was supported by the Canada Excellence Research Chairs (CERC) (215063) program.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.neuroimage.2015.11.059.

References


Please cite this article as: Naci, L., et al., Detecting and interpreting conscious experiences in behaviorally non-responsive patients, NeuroImage (2015), http://dx.doi.org/10.1016/j.neuroimage.2015.11.059

Additional references and data are available online at http://dx.doi.org/10.1016/j.neuroimage.2015.11.059.


