

# The Brain's Silent Messenger: Using Selective Attention to Decode Human Thought for Brain-Based Communication

Lorina Naci, Rhodri Cusack, Vivian Z. Jia, and Adrian M. Owen

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

The interpretation of human thought from brain activity, without recourse to speech or action, is one of the most provoking and challenging frontiers of modern neuroscience. In particular, patients who are fully conscious and awake, yet, due to brain damage, are unable to show any behavioral responsiveness, expose the limits of the neuromuscular system and the necessity for alternate forms of communication. Although it is well established that selective attention can significantly enhance the neural representation of attended sounds, it remains, thus far, untested as a response modality for brain-based communication. We asked whether its effect could be reliably used to decode answers to binary (yes/no) questions. Fifteen healthy volunteers answered questions (e.g., “Do you have brothers or sisters?”) in the fMRI scanner, by selectively attending to the appropriate word (“yes” or “no”). Ninety percent of the answers were decoded correctly based on activity changes within the attention network. The majority of volunteers conveyed their answers with less than 3 min of scanning, suggesting that this technique is suited for communication in a reasonable amount of time. Formal comparison with the current best-established fMRI technique for binary communication revealed improved individual success rates and scanning times required to detect responses. This novel fMRI technique is intuitive, easy to use in untrained participants, and reliably robust within brief scanning times. Possible applications include communication with behaviorally nonresponsive patients.

## Introduction

The question of whether we can convey our thoughts without recourse to speech or action has preoccupied scientists for decades. In particular, patients who are fully conscious and awake, yet are unable to show any behavioral responsiveness due to brain damage (Owen et al., 2006; Monti et al., 2010; Cruse et al., 2011), expose the limits of the neuromuscular system and the necessity for alternate forms of communication. Over the past 30 years, electroencephalography (EEG) studies have provided significant insights into how “neural responsiveness” might be used to drive motor-independent communication. More recently, functional magnetic resonance imaging (fMRI) studies have investigated brain–computer interfaces (BCIs) based on the hemodynamic brain response (Owen et al., 2006; Boly et al., 2007; Monti et al., 2010; Bardin et al., 2011; Bardin et al., 2012; Sorger et al., 2012).

Whether they aim for binary or multiple-choice communication, fMRI BCI methods rely on one-to-one mapping between a task (e.g., motor/spatial navigation imagery, inner speech, mental rotation), which can be readily executed at will to generate strong brain activity, and a unique communication token, such as

a word (e.g., “yes” or “no”) or a letter of the alphabet. These arbitrary task-to-response mappings engender several limitations including complex instructions and reliance on participant pretraining, high demands on short-term memory resources, and relatively long scanning times needed to convey a response. The most successful binary method (Boly et al., 2007) deployed to date uses a dual-task paradigm and has been used to communicate with untrained healthy volunteers (Monti et al., 2010), as well as behaviorally nonresponsive patients (Owen et al., 2006; Monti et al., 2010). To convey either a “yes” or a “no,” a participant engages in motor (playing tennis) or spatial navigation (navigating around one’s home) imagery. Although highly reliable in healthy volunteers, this method places considerable demands on short-term memory resources (Monti et al., 2010) and requires relatively long scanning intervals, which can become limiting factors, especially for brain-injured patients. Moreover, some healthy participants do not generate robust activation to mental imagery tasks, thus reducing the efficacy of this method in this group (Guger et al., 2003; Boly et al., 2007).

We investigated selective auditory attention as a more intuitive, yet thus far untested, response modality for robust, reliable, and highly accurate brain-based communication. It is well established that selective attention can significantly enhance the neural representation of attended sounds (Bidet-Caulet et al., 2007), although most studies focus on group-level changes rather than individual responses. Hence, the first aim was to investigate whether the effect of selective attention on the brain response to attended words (e.g., “yes”/“no”) could be observed robustly in individual participants. The second aim was to determine whether this effect could be used for accurate binary communication. We hypothesized that this method would be not only

Received Dec. 5, 2012; revised April 15, 2013; accepted April 25, 2013.

Author contributions: L.N. and A.M.O. designed research; L.N. and V.Z.J. performed research; R.C. contributed unpublished reagents/analytic tools; L.N. and R.C. analyzed data; L.N. and A.M.O. wrote the paper.

This work was funded by the DECODER Project, funded by the European Commission in the Seventh Framework Programme (2007–2013), the James S. McDonnell Foundation, and the Canada Excellence Research Chairs Program. We thank Adam McLean for his technical support with the functional magnetic resonance imaging data acquisition.

The authors declare no competing financial interests.

Correspondence should be addressed to Dr. Lorina Naci, The Brain and Mind Institute, Department of Psychology, Western University, London, ON N6A 5B7, Canada. E-mail: lnaci@uwo.ca.

DOI:10.1523/JNEUROSCI.5577-12.2013

Copyright © 2013 the authors 0270-6474/13/339385-09\$15.00/0

more intuitive, and thus easier to use, but also more time efficient than the current best-established fMRI technique for binary communication (Owen et al., 2006; Boly et al., 2007; Monti et al., 2010).

## Materials and Methods

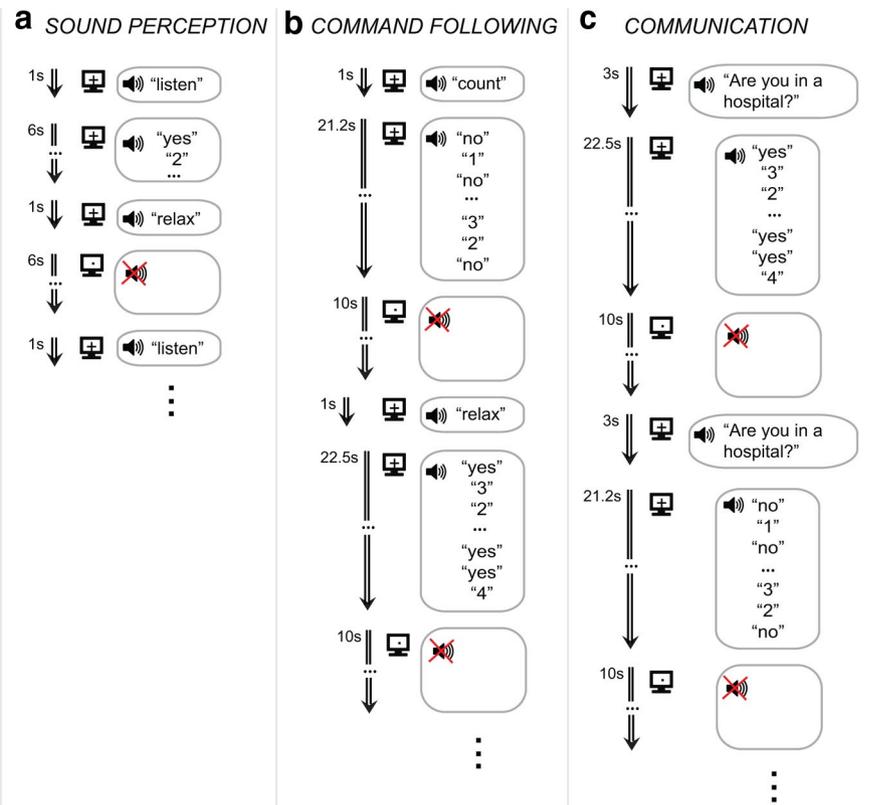
**Volunteers.** Ethical approval was obtained from Western University's Health Sciences Research Ethics Board. Sixteen volunteers (19–31 years; six males) with no history of neurological disorders participated in the study. All volunteers were right handed and native English speakers. They signed informed consent before participating and were remunerated for their time. One female volunteer became uncomfortable in the scanner and could not complete the experiment. Data from this volunteer were excluded from the analyses.

**Stimuli.** All instructions and auditory stimuli were recorded from a Canadian English male speaker and presented binaurally. The stimuli were the words “one,” “two,” “three,” “four,” “five,” “six,” “seven,” “eight,” “nine,” “yes,” “no,” and the sentences “Do you have any brothers or sisters?” and “Are you over 21 years old?” All waveforms were normalized to match their RMS amplitude using the Praat software ([www.praat.org](http://www.praat.org)) (Boersma and Weenink, 2012). A white fixation was presented on a black screen during sound presentation, and a blank black screen was presented during silent intervals.

Before scanning, computer-based versions of the tasks that elicited button-press responses were tested behaviorally, in an independent group of volunteers ( $n = 16$ ). Error rates and reaction times were measured. Task parameters were optimized to engage attention sufficiently, so as to produce detectable behavioral effects at a single-subject level, which might improve its detectability with fMRI. However, behavioral output was irrelevant to the fMRI experiment, where the index of “success” was a significant neural change contingent on the task manipulation. Hence, no behavioral output was measured in the scanner.

**fMRI task design.** The first level of the selective attention paradigm examined the cortical response during passive sound perception. Volunteers passively listened to single-word stimuli and performed no active task for the duration of the session. Trials had an on/off design, with miniblocks/sequences of words (6 s) followed by silence (6 s) (Fig. 1a). The instruction to listen/relax (1 s) cued the participants at the start of each sound perception/relaxation interval. The words were presented in a pseudorandom order. There were 24 on/off trials of 14 s each, lasting 5.6 min, including the delivery of word cues.

The second level tested an individual's ability to upregulate his or her brain activity by selectively paying attention to specific words. We used a counting task to manipulate auditory attention, in particular because counting has been shown to engage sustained attention (Ortuño et al., 2002). Moreover, mental calculation tasks have been successfully used in BCI applications (Lee et al., 2009), as they have been found to elicit robust activations at the single-subject level, thus fulfilling one of the criteria for successful application of the BCI technique in individual users. This second session also served to localize, for each volunteer, the brain regions most responsive to the attention manipulation; thus, we refer to it as the “localizer” session. We expected that due to variations in brain morphology, functional organization, and possibly also different task strategies, the foci of activation due to attention would vary slightly between individuals. Thus, each individual's native attention network, determined based on functional activation in the localizer session, was



**Figure 1.** fMRI paradigm. *a–c*, The figure illustrates the design of each component of the fMRI paradigm: (*a*) sound perception, (*b*) attention localizer, and (*c*) communication.

used to constrain subsequent analyses of the brain responses during the communication sessions.

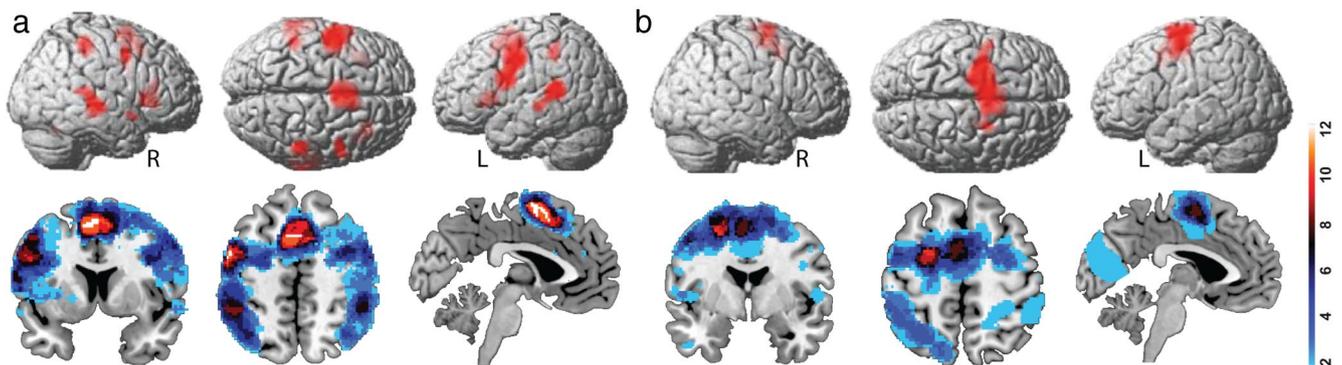
In BCI paradigms, where functional activation serves as a proxy for the participant's behavior, a high level of confidence in the fMRI results is necessary to avoid false positives. To ensure robust and reliable effects for each individual, we tested the results against a priori hypotheses, motivated from the participant's previously established response to similar stimuli. Specifically, initially we build a hypothesis during the localizer session, and subsequently tested it during each communication session. The localizer session allowed us to identify each individual's attention network and to hypothesize that all or parts of this network would be activated, when the subject attended to an answer (yes or no), during the communication session. Subsequently, in the communication sessions, we tested whether these effects were manifested when the participant freely chose to attend to a word (either yes or no) in response to a binary question.

The trials in the localizer session had an on/off design, with the attention interval 21.2 s/22.5 s, followed by the “relaxation” interval (10 s) (Fig. 1*b*). During each attention/relaxation interval, the volunteer heard a sound sequence containing repeated presentations of a target word, either “yes” or “no,” interspersed with repeated presentations of the digits 1 to 9. The words were repeated several times and presented in a pseudorandomized order; no more than two consecutive repetitions of the same word occurred. Each sequence started with either the word “yes” or the word “no.” The opposite word was presented in the next sequence. Thus, the yes and no sequences appeared in pairs, though, outside of the pair, two sequences presenting the same word could follow each other. Each sequence was composed of 40 words with 9 or 11 presentations of the target word, had 200 ms gaps between words, and lasted between 21,200 and 22,500 ms, depending on whether it presented the word “no” (350 ms) or the word “yes” (450 ms). The instruction “count,” presented at the start of the attention interval, signaled to the volunteer to count the number of times either of the target words (“yes” or “no”) occurred within the following sequence. The purpose of the digits (1 to 9) was to act as close distractors to the target number, thus increasing task

**Table 1. Individual activation peaks for each session**

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Total	Peak coordinates (x, y, z, mean ± SD)
<b>Sounds &gt; Silence</b>																	
Superior temporal gyrus (L/R)	+	+	+		+	+	+	+	+	+	+	+	+	+	+	14 of 15	L, -63 ± 6, -20 ± 5, 6 ± 2; R, 68 ± 3, -14 ± 12, 1 ± 5
Middle temporal gyrus (L/R)		+		+		+					+		+	+	+	7 of 15	L, -66 ± 3, -22 ± 7, 2 ± 6; R, 68, -25 ± 1, -5 ± 1
<b>Count &gt; Relax</b>																	
Middle temporal gyrus (L)	+	+	+													3 of 15	-61 ± 8, -40 ± 5, 11 ± 8
Postcentral gyrus (L)	+								+			+	+			4 of 15	-55 ± 5, -4 ± 4, 39 ± 14
Precentral gyrus (L/R)		+	+	+	+		+					+		+		7 of 15	L, -4 ± 2, 3 ± 5, 59 ± 3; R, 57 ± 4, 2 ± 6, 44 ± 8
Pre-SMA (L/R)						+	+		+	+	+		+		+	7 of 15	L, -4 ± 2, 3 ± 5, 59 ± 3; R, 8, 10, 54
Inferior parietal gyrus (L)								+								1 of 15	L, -40, -56, 54
Superior frontal gyrus (R)				+												1 of 15	28, 0, 60
Inferior frontal operculum (R)						+										1 of 15	48, 8, 26
Inferior parietal gyrus (R)								+								1 of 15	38, -46, 52
Middle frontal gyrus (R)											+			+		2 of 15	48, 18 ± 20, 46 ± 14
<b>Tennis &gt; Relax</b>																	
Pre-SMA (L/R)	+	+	+		+		+	+	+		+	+	+	+	+	12 of 15	L, -7 ± 6, -4 ± 3, 64 ± 6; R, 12, -12, 74
Inferior parietal lobule			+			+	+	+			+			+	+	7 of 15	L, -34, -58, 56; R, 50, -54, 46

Coordinates of the peaks of the two most highly activated clusters for each volunteer during sound perception, selective attention, and motor imagery sessions ( $p < 0.05$ , FDR corrected).



**Figure 2.** Brain regions significantly activated by response generation, with selective attention and motor imagery. **a, b**, Regions significantly activated by selective attention (**a**) and those activated by motor imagery (**b**) in the volunteer group ( $n = 15$ ). The bottom image in each panel displays the overlay of significant activations from each volunteer ( $p < 0.05$ ; FDR voxelwise corrected). The color bar to the right indicates the number of participants, corresponding to each color, who activate the same voxels. Warm colors depict high overlap.

difficulty during the count trials and enabling suppression of any automatic task activity during the relax trials. To minimize eye movements, a white fixation was presented on a black screen during sound presentation, and volunteers were asked to keep eyes on the fixation. At the end of a sound sequence, a blank black screen and no sound were presented for 10 s. This provided a break from stimuli presentation and task performance. The instruction to “relax” followed, and signaled to the volunteer to ignore the sounds and not count during this time. There were five count/relax intervals, or trials, lasting 5.8 min.

The third level tested the ability to communicate answers to binary (yes/no) questions, simply by attending to the appropriate word (Fig. 1c). There were two independent “communication” sessions, each presenting either the question “Do you have any brothers or sisters?” or “Are you over 21 years old?” These two questions were chosen to generate a relatively even spread of yes and no responses among our volunteer group. The volunteers had to convey the answer to each question by selectively paying attention to the answer word and ignoring the occurrences of the opposite word. To selectively attend to a word, volunteers used a similar strategy to that used in the localizer session: counting the number of times they heard the answer word (“yes” or “no”), if the sequence presented the answer, and relax or do not count if the sequence presented the

opposite word. Similarly to the localizer session, the yes and no sequences were organized in pairs, and a break period of 10 s (blank screen and silence) followed each sequence. Unlike the localizer session, in the communication session, the selection of the target word to be attended (“yes” or “no”) was self-guided, rather than dictated by instructions. Each volunteer determined the target online, at the start of the session, depending on his or her answer to the specific question. The same question and a short instruction were repeated at the start of each sound sequence. There were five yes/no trials, lasting 7 min, including repeated delivery of the question and instructions. At the end of the experiment, volunteers provided verbal yes/no answers to the questions asked in the scanner. These were used to cross-validate the answers that were derived based on the fMRI activation.

In addition, a session of a well-established motor imagery task (imagining playing tennis) (Owen et al., 2006) was acquired for comparison with the selective attention task. For the purpose of this comparison, the motor imagery and spatial navigation imagery tasks used in the aforementioned dual-task paradigm were considered equivalent with regard to their BCI performance. In the interest of brevity, only one session—the motor imagery task—was acquired and compared to the selective attention task in terms of individual success rate, and the amount of

scanning time required to convey binary (yes/no) responses. The design of this session was kept identical to those previous studies using this paradigm (Owen et al., 2006; Boly et al., 2007; Monti et al., 2010). The session had an on/off design. Volunteers imagined playing tennis (30 s) every time they heard the word “tennis,” and relaxed (30 s) every time they heard the word “relax.” In this session, volunteers were asked to keep their eyes closed, to reduce sensory input and facilitate mental imagery (Kosslyn et al., 2001). There were five tennis/relax intervals, or trials, lasting 5.3 min, including the delivery of word cues.

The five experimental sessions—sound perception, selective attention localizer, selective attention communication (twice), and motor imagery—were presented in pseudorandom order.

**fMRI data acquisition.** Scanning was performed using a 3 Tesla Siemens Tim Trio system with a 32-channel head coil, at the Robarts Research Institute in London, Ontario, Canada. Functional echoplanar images were acquired [33 slices; voxel size,  $3 \times 3 \times 3$  mm; interslice gap, 25%; TR, 2000 ms; TE, 30 ms; matrix size,  $64 \times 64$ ; flip angle (FA),  $75^\circ$ ]. The selective attention paradigm had 150 scans for the passive listening, 175 scans for the localizer, and 220 scans for each communication session. One hundred and sixty scans were acquired for the tennis imagery session. An anatomical volume was obtained using a T1-weighted 3D MPRAGE sequence (32 channel coil; 33 slices; voxel size,  $1 \times 1 \times 3$  mm; interslice gap, 50%; TR, 2300 ms; TE, 4.25 ms; matrix size,  $64 \times 64$ ; FA,  $75^\circ$ ). Volunteers lay supine in the scanner looking upward into a mirror box that allowed them to see a projector screen behind their head. Noise-cancellation headphones (Sensimetrics, S14) were used for sound delivery.

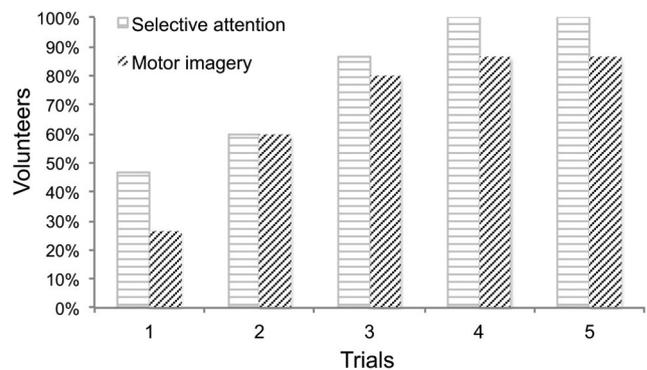
**fMRI data analyses.** The imaging data were analyzed using SPM8 (Wellcome Institute of Cognitive Neurology; <http://www.fil.ion.ucl.ac.uk/spm/software/spm8/>). Preprocessing was performed using AA software ([www.cusacklab.org](http://www.cusacklab.org)). The processing steps were as follows: correction for timing of slice acquisition, motion correction, normalization to a template brain, and smoothing. The data were smoothed with a Gaussian smoothing kernel of 10 mm FWHM (Peigneux et al., 2006). Spatial normalization was performed using SPM8s segment-and-normalize procedure, whereby the T1 structural was segmented into gray and white matter and normalized to a segmented MNI152 template. These normalization parameters were then applied to all echoplanar images. The time series in each voxel was high-pass filtered with a cutoff of 1/128 Hz to remove low-frequency noise, and scaled to a grand mean of 100 across voxels and scans in each session.

Before analyses, the first five scans of each session were discarded, to account for T1 relaxation and allow volunteers to adjust to the noise of the scanner. The preprocessed data were then analyzed in SPM8 using the general linear model. Fixed-effect analyses were performed in each subject, corrected for temporal autocorrelation using an AR(1) + white noise model. Two event types for each session, corresponding to the two on/off periods, were defined as follows: sound perception (sound/silence; 6 s/6 s), selective attention localizer (count/relax; 21.2 s/22.5s), communication (“no” sequence/“yes” sequence; 21.2 s/22.5 s), and motor imagery (tennis/relax; 30 s/30 s). The silent periods (10 s) in the selective attention paradigm provided an implicit baseline for the relaxation of the BOLD signal to baseline level, as well as a period of true rest for the participants in between periods of sound presentation. Moreover, the relax trials, where participants were instructed to pay no attention, served as a complex baseline to the count trials, where participants were instructed to pay attention to specific words. To control for any sensory confounds, the stimuli in the two trial types were exactly the same, and they differed only in the instructions. By comparing the BOLD activation in these two trial types, we were able to localize the effects of attention for each participant. Events for each of the regressors were modeled by convolving boxcar functions with the canonical hemodynamic response function. Also included in the general linear model were nuisance variables, namely, the movement parameters in the three directions of motion and three degrees of rotation, as well as the mean of each session. Linear contrasts were used to obtain subject-specific estimates for each effect of interest. The contrasts containing these parameter estimates for each voxel were entered in the second stage of analysis, treating volunteers as random effects and using one-sample *t* test across the 15 volun-

**Table 2. Group activation peaks in the attention localizer (Count > Relax)**

Brain areas	Side	x	y	z	z value	p value
Postcentral	L	-54	-6	42	5.41	0.001
	L	-62	2	22	4.05	0.007
Precentral	L	-50	2	28	4	0.008
	R	48	2	46	3.85	0.011
Superior temporal	R	42	4	32	3.11	0.04
	L	-56	-38	14	5.2	0.001
	R	54	-24	-4	4.38	0.004
	R	70	-32	10	3.48	0.02
Middle temporal	R	60	-42	12	3.2	0.034
	L	-66	-26	0	3.39	0.024
	L	-50	-24	-4	3	0.048
Pre-SMA	L	0	6	60	4.91	0.002
	R	6	14	50	4.51	0.003
Inferior frontal triangularis	R	32	30	2	4.06	0.007
Insula	R	30	26	12	3.9	0.01
	R	36	22	4	3.88	0.01
	L	-30	24	4	3.8	0.012
	L	-36	14	6	3.5	0.019
Inferior parietal	L	-46	-40	46	4.05	0.007
	R	48	-34	50	3.97	0.008
Superior temporal pole	R	54	6	-10	3.44	0.021

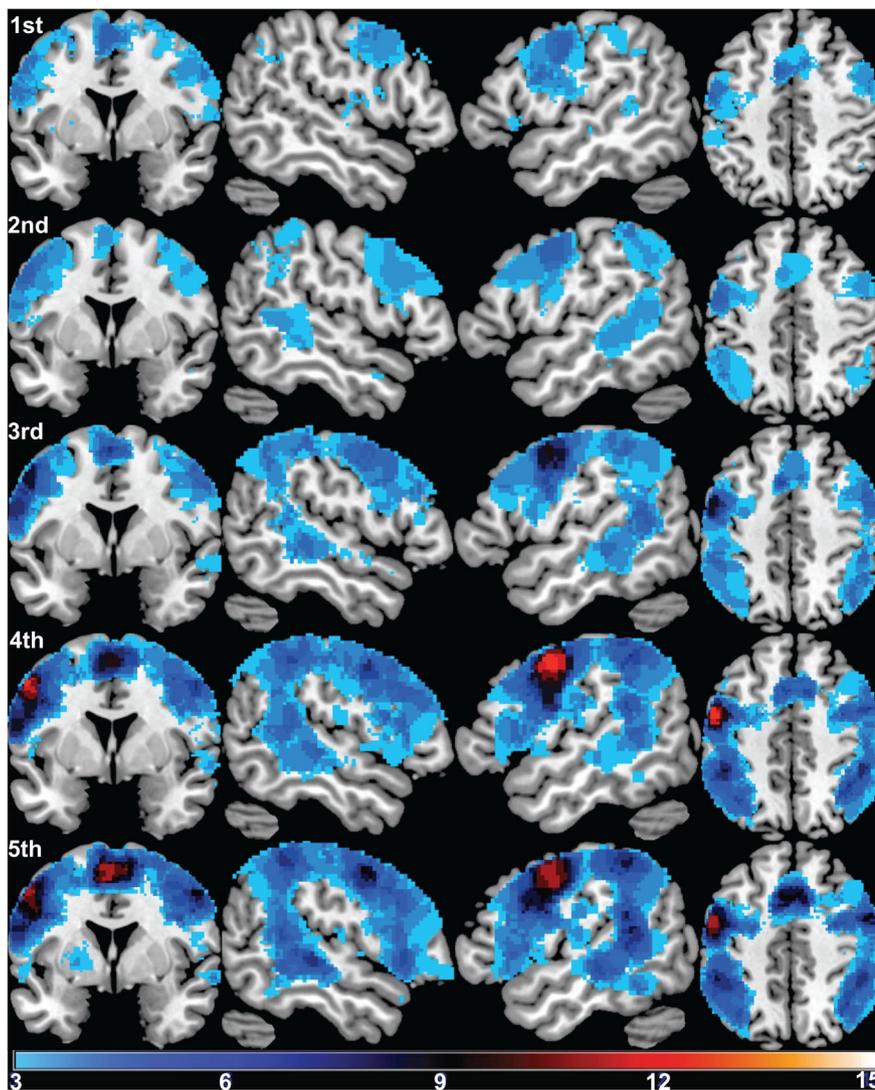
Significant whole-brain results for the group data ( $p < 0.05$ , FDR corrected). L, Left; R, right.



**Figure 3.** The proportion of volunteers that showed significant activation with increasingly more time, in each task. The figure displays significant whole-brain activation to the selective attention and motor imagery tasks at one of five time intervals in the same areas where activation was observed for the duration of the entire task. Each interval contains an added on/off trial, from one, at the start of each session, to five, at the end. At the end of each scanning session, 100 versus 87% of volunteers showed significant activation for each task, respectively.

teers. For the sound perception, localizer, and motor imagery sessions, whole-brain analyses were used to determine significant activation at the individual/fixed effects and at the group/random effects levels. Only clusters or voxels that survived at a  $p < 0.05$  threshold, corrected for multiple comparisons [false discovery rate (FDR)] (Worsley et al., 1996), were reported.

The comparison between the selective attention and the motor imagery tasks were based on data from whole-brain analyses of each data set. In further analyses, regions of interest (ROIs) analyses were used to analyze brain responses during the communication sessions of the selective attention paradigm. For each individual data set, activations at the fixed effects level, from the Count > Relax contrast (localizer session), were used to derive two ROIs. Each ROI was defined as a 10 mm sphere with center coordinates at the peak voxel of each of the two most strongly activated significant clusters. (For Volunteers 10 and 15, a single ROI was derived based on the single significant cluster observed in the whole-brain data analysis of each.) These independently defined, subject-specific ROIs were used to test for significant activations in each communication session in the fixed effects-level yes–no and no–yes contrasts. Contrasts in both directions were performed because we did not have a priori hypotheses about which word the volunteer would attend



**Figure 4.** Overlay of significant activations from each volunteer, during the selective attention session, with increasingly more time on task. The figure displays activation in the first-level model (Count > Relax;  $p < 0.05$ ; FDR cluster corrected) for each volunteer, at one of five time intervals, from the first, at the top, to the fifth, at the bottom. The figure shows both the individual variation in activation pattern at each interval and their spatial overlap. The first time interval shows significant activation in 7 volunteers, the second in 9, and the third in 13, and the fourth and fifth intervals each show activation in 15 volunteers. The color bar at the bottom indicates the number of participants, corresponding to each color, who activate the same voxels. Warm colors depict high overlap.

to. The MarsBaR SPM toolbox (<http://marsbar.sourceforge.net/>) was used to test ROI activations (Brett et al., 2002). For each volunteer, each answer decoded by fMRI analysis was compared to that provided in the volunteer's verbal report. The binomial test was used to assess decoding accuracy at the group level.

In addition to the standard analyses mentioned above, where all the scans collected in each session were included, analyses were performed to determine how much acquisition time would be necessary to elicit robust activation at the single-subject level with each task. Each individual data set from each of the selective attention and motor imagery sessions was divided into five sets of increasingly bigger numbers of scans. The first time interval included the first on/off trial in each task (i.e., count/relax localizer, answer/relax, tennis/relax), the second included the first two, the third included the first three, the fourth included four, and the fifth included the entire data set. There were 34, 68, 102, 136, and 170 scans per set of the localizer; 42, 84, 126, 168, and 210 scans per set of each communication session; and 31, 62, 93, 124, and 155 scans per set of the motor imagery session. The analyses performed for each time interval, for each individual data set, were the same as described above: whole-

brain analyses for the localizer and imagery sessions, and ROI analyses for the communication sessions.

## Results

In the sound perception session (contrast Sound > Silence), passive listening activated bilateral auditory cortex (superior and middle temporal gyri) in all volunteers significantly more than silence (Table 1).

In the localizer session (contrast Count > Relax), whole-brain analysis revealed significant activations in bilateral inferior parietal, superior temporal, premotor, and inferior prefrontal gyri significantly more than rest in the group data (Fig. 2a; Table 2). Furthermore, the effect of selective attention was observed at the single-subject level. Each volunteer activated a subset of these regions, and the peak activation foci varied slightly between volunteers (Table 1). As the same sounds were presented during both the count and relax trials, these activations do not reflect basic auditory processing. Indeed, the fact that significant differences were detected when the external stimuli remained formally identical, confirms that the volunteers were able to selectively attend to a specific word (yes/no). The foci of significant activations, including the superior temporal, premotor, and inferior prefrontal frontal cortex, have been implicated in speech perception and production (Pelle et al., 2010). This suggests that volunteers were able to selectively attend to either word by activating its linguistic representation, possibly through a subvocal rehearsal process. Additionally, activations in the prefrontal and parietal regions are consistent with studies implicating these regions in multiple cognitive demands, including attention (Duncan, 2010; Ptak, 2011; Frank and Sabatinelli 2012), not dissimilar to those imposed by the counting task used here.

Motor imagery (imagining playing tennis) activated a large cluster centered in the left pre-SMA ( $x = -8, y = -2, z = 72$ ;  $p = 0.001$ , FDR clusterwise corrected), in the whole-brain analysis of group data (Fig. 2b). This result closely replicates those of previous studies (Owen et al., 2006; Boly et al., 2007; Monti et al., 2010). Significant activation in the pre-SMA, as revealed by whole-brain analyses, was observed for 12 of 15 volunteers (Table 1). One in 12 volunteers showed subthreshold activation ( $p = 0.06$ , FDR corrected). Nevertheless, we classified this volunteer as a positive case, because activation was observed in the anatomically appropriate region (pre-SMA), as predicted a priori from previously published data. As we did not apply any volume reduction technique to enhance effects in this a priori motivated region (i.e., masking, small volume correction, ROI analysis), but performed whole-brain analyses, effects in this region are more conservatively thresholded than in previous reports, and are

**Table 3. Individual decoding results for the answers to each question**

	Volunteer number															Total	Proportion decoded
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15		
Q1	+	+	+	+	+	+	+	–	+	–	+	+	+	+	+	13/15	86.67%
Q2	+	+	+	+	+	+	+	+	–	+	+	+	+	+	+	14/15	93.33%
Total	2	2	2	2	2	2	2	1	1	1	2	2	2	2	2	27/30	90%

Activation in independently defined regions of interests, native to each volunteer, decoded both answers in 12 of 15 and one answer in 3 of 15 of the volunteers. Overall, 90% of all the answers, conveyed solely on the basis of brain activation, were correctly decoded.

likely to be at least as robust. Activation in the inferior parietal lobule, the second region reported by the original group study on motor imagery (Boly et al., 2007), was observed in 7 of 15 volunteers. Two volunteers did not show any activation to motor imagery. In summary, 13 of 15, or 87%, of the volunteers showed task-specific activation to motor imagery.

How much scanning time is needed on the selective attention task to generate robust brain activation that can be detected with fMRI? Forty-seven percent of volunteers generated significant activation within one on/off trial (1.1 min) in the same areas where activation was observed for the duration of the entire task. A lower proportion, 25%, of volunteers generated significant activation to the motor imagery in about the same time, again in the same areas as those observed for the entire task (Fig. 3). All of the successful volunteers showed task-specific significant activation, for each task, in four on/off trials (4.5 min/4.1 min).

Do volunteers perform the attention task in the early time intervals (i.e., within 1.1 min) in the same way as they do with more time on task, or do the early activations reflect general effort, and, therefore, functionally undifferentiated brain activity? The subset of volunteers who could do the task in one on/off trial (i.e., exhibited significant task-related activity) activated the same regions across the different intervals on task (Fig. 4). This suggested that they engaged task-appropriate cognitive functions, such as language perception and production/subvocalization, counting, and, more generally, selective sustained attention, soon after the start of the task. With more time on task, the number of volunteers that showed significant activation in these task-specific regions increased, and so did the overlap between the individual activation maps (Fig. 4).

Across all of the communication sessions of the selective attention paradigm, the direction of significant activation (contrast “yes” sequence vs “no” sequence) in either one or both subject-specific and independently defined ROIs successfully decoded 90% of the answers ( $p = 0.0001$ ). For 3 of 15 volunteers, one of two answers was decoded, and for the majority (12 of 15), both answers were successfully decoded (Table 3). For the 3 of 30 answers that were not decoded, activation failed to reach the threshold of statistical significance (Fig. 5). Across volunteers, the subject-specific regions of interest overlapped most commonly in the bilateral precentral gyrus and pre-SMA (Table 4). These results suggested that the effect of self-guided selective attention to either word (“yes” or “no”) could be successfully decoded for each individual volunteer, and thus could be used as a reliable method for brain-based communication.

Given that it is possible to determine, in some cases, that a participant can pay attention based solely on his or her brain activity, within approximately 1 min of task compliance in the scanner, how much scanning time is needed to decode the answer to a question? For the first question, 1.4 min of scanning was sufficient to decode the answer in 54% of volunteers; 2.8 min was sufficient in 85%, 4.2 min was sufficient in 85%, 5.6 min was sufficient in 92%, and 7 min was sufficient in 100% of volunteers

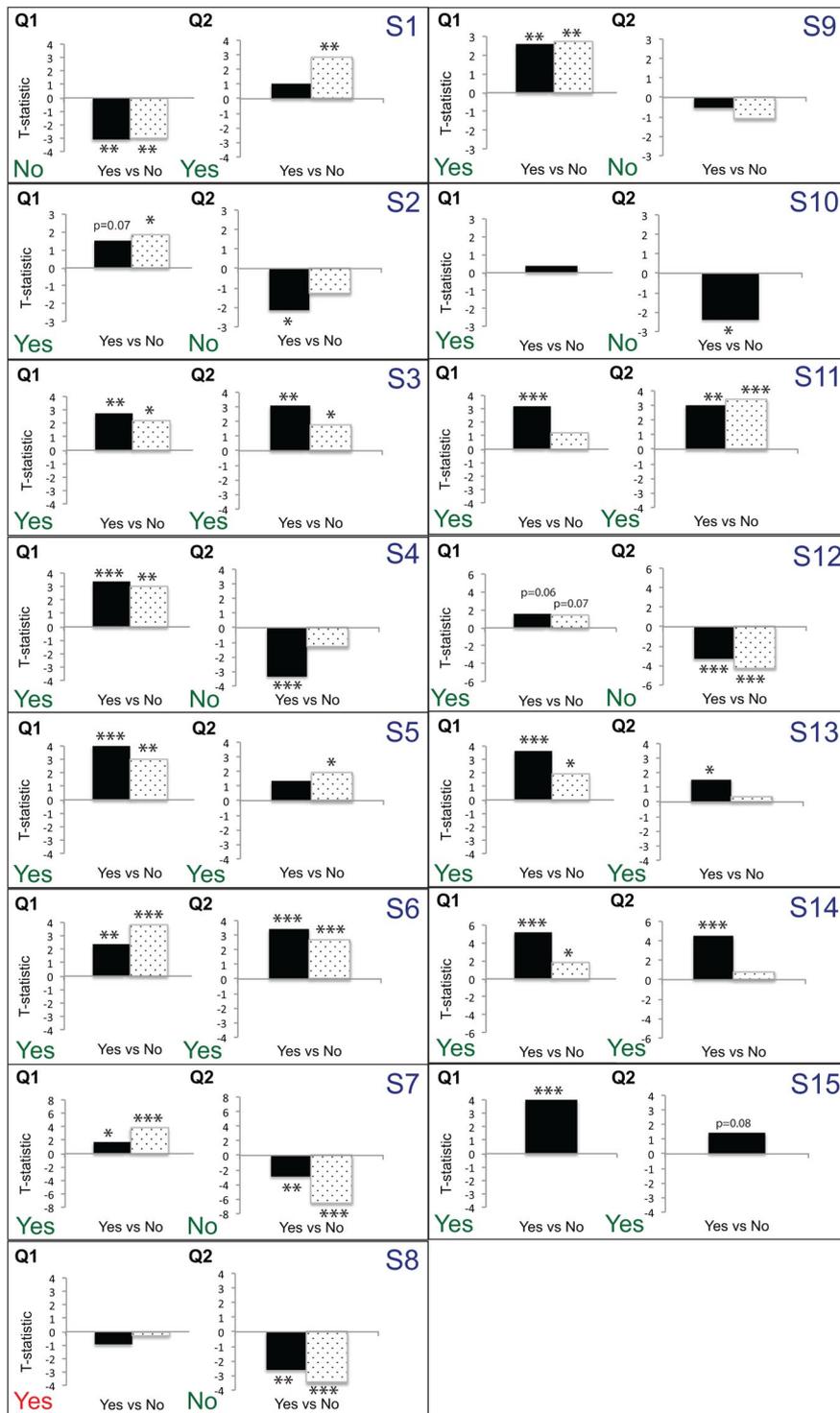
(Fig. 6). The high proportion of volunteers that can respond in brief scanning intervals (i.e., 1.4 min) suggests that this method lends itself to BCI applications. The proportions for the first communication session were comparable with those in the localizer session, whereas in the second communication session, performance decreased, perhaps due to fatigue from increased overall time in the scanner. Within each communication session, the proportion of successful volunteers continued to increase with more time on task, suggesting that some volunteers needed the entire duration of the session to produce sufficiently robust brain activity to convey the answers to the questions.

## Discussion

For the first time, we show that the neural effects of selective auditory attention can be reliably and robustly replicated in individual participants so as to convert intentions into brain-based communication. We report a novel and intuitive fMRI method for binary communication that allows a participant to respond without any behavioral action, by simply attending to the word (yes or no) he or she wishes to convey. The results confirmed that this technique delivers highly robust, reliable, and accurate communication within brief scanning times (less than 5 min) in untrained volunteers.

Critically, each volunteer used top-down attention to select and convey answers to binary (yes/no) autobiographical questions. The power of selective attention to magnify the neural representation of relevant sounds (Tiitinen et al., 1993; Woldorff et al., 1993), and suppress other (Ghatan et al., 1998; Kawashima et al., 1999), irrelevant ones, has been well established in previous studies. However, thus far, it has not been investigated as a method of communication for fMRI-based BCI applications. This may, in part, be due to the highly variable selectivity of auditory attention (Fritz et al., 2007). Because the features to be attended can vary widely (e.g., auditory pitch, frequency or intensity, tone duration, timbre), the effect of attention is likely to be observed in a number of diverse neural regions, determined by the specific demands of the attention manipulation.

We observed activity enhancements in a network of regions including inferior parietal, superior/middle temporal, and precentral/postcentral gyri, subtending to the pre-SMA, as well as the inferior frontal gyrus and the insula. All of these regions have been implicated in a variety of cognitive processes elicited by our attention manipulation, namely, mental arithmetic. In particular, the frontoparietal network, involved in supramodal attention demands (Duncan, 2010), may be recruited by the attention demands of the counting task. The recruitment of the association auditory cortex (middle/superior temporal gyri) has been observed in attention-guided selective processing of linguistic stimuli (Hugdahl et al., 2003), much like the attention to specific words required by our mental task. Furthermore, the superior temporal, premotor, and inferior prefrontal frontal cortex, have been implicated in speech perception and production (Pelle et



**Figure 5.** Individual answers decoded from brain activity. Each panel depicts the *t* statistic graphs for the two questions (Q1, left; Q2, right), for each volunteer (S1–S15). Each bar represents the *t* statistic for the yes versus no comparison for each ROI (first, solid black; second, back and white bar). For Volunteers 10 and 15, a single ROI was derived (solid black bar), based on the single significant cluster observed in the whole-brain data analysis. The respective correct answers, reported verbally by the volunteer, appear at the bottom of each graph. The answers appear in green (29 of 30), if the brain activation suggested the same answer as the verbal report, and in red (1 of 30) otherwise. In three volunteers (S8, S9, and S10), in half of the sessions, activation failed to reach statistical significance, and decoding could not be carried out. In only one of these three sessions, the direction of activation suggested the wrong answer (S8; Q1). \**p* < 0.05; \*\**p* < 0.01; \*\*\**p* < 0.001.

al., 2010), processes likely recruited by a subvocal rehearsal of the target words, which may facilitate their maintenance in short-term memory. Finally, the insula has been implicated in salience processing (Menon and Uddin, 2010) and may serve here to fa-

ilitate access to attention and working memory resources, when a salient event (i.e., a specific word, “yes” or “no”) is detected. Notwithstanding variations in morphology, functional organization, and possibly also task strategy among different volunteers, we observed robust and replicable effects of attention in some or all regions of this bilateral network, at the level of individual participants. This suggested that the paradigm is suitable for BCI applications.

One region that displayed high overlap between the individual foci of attention was the pre-SMA. Similarly, significant pre-SMA activations were observed in the majority of volunteers during the motor imagery paradigm (Owen et al., 2006), which showed focal recruitment, for the most part, limited to this region. The involvement of the pre-SMA in both tasks can be explained based on its role in a variety of cognitive functions (Nachev et al., 2007), including sustained cognitive control (Nachev et al., 2005) and attention to intention (Lau et al., 2004), two processes critical to both paradigms. Importantly, the difference between the two in the extent of neural recruitment suggests that the attention paradigm invokes a larger number of discrete cognitive processes compared to motor imagery. The motor imagery, rather than the spatial navigation task (Owen et al. (2006)), was used in this study to provide a conservative comparison with the novel selective attention task. Boly et al. (2007) found that motor imagery provided the most robust results out of four mental imagery tasks tested, including spatial navigation.

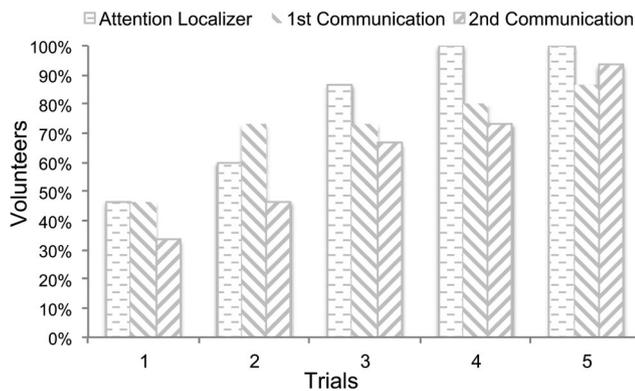
Formal comparison between the selective attention and motor imagery tasks revealed better performance in the individual success rates and amount of scanning time needed to elicit robust responses in the selective attention paradigm. First, 100% of volunteers showed significant task-appropriate activity to the selective attention task, compared to 87% to the motor imagery. This result is consistent with previous studies showing that a proportion of healthy volunteers do not produce reliable brain activation during mental imagery tasks (Boly et al., 2007). Conversely, the high success rate in the selective attention task may be explained by its reliance on a number of interrelated cognitive processes, which may make it robust to weak or inconsistent brain activation in any given process.

Second, even for volunteers who needed the entire session to convey a response, the selective attention task required half the time that a dual-task paradigm, involving motor imagery, would require (Owen et al., 2006; Monti et al., 2010). In the

**Table 4. Regions that encode brain-based responses, for each volunteer**

Region of interest	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Total	Peak Coordinates (x, y, z, mean ± SD)
Middle temporal gyrus (L)	+	+	+													3 of 14	−61 ± 8, −40 ± 5, 11 ± 8
Postcentral gyrus (L)	+								+			+	+			4 of 14	−55 ± 5, −4 ± 4, 39 ± 14
Precentral gyrus (L/R)		+	+	+	+		+					+		+		7 of 14	L, −4 ± 2, 3 ± 5, 59 ± 3; R, 57 ± 4, 2 ± 6, 44 ± 8
Pre-SMA (L/R)						+	+		+	+	+		+		+	7 of 14	L, −4 ± 2, 3 ± 5, 59 ± 3; R, 8, 10, 54
Inferior parietal gyrus (L)								+								1 of 14	L, −40, −56, 54
Superior frontal gyrus (R)				+												1 of 14	28, 0, 60
Inferior frontal operculum (R)						+										1 of 14	48, 8, 26
Inferior parietal gyrus (R)								+								1 of 14	38, −46, 52
Middle frontal gyrus (R)											+			+		2 of 14	48, 18 ± 20, 46 ± 14

L, Left; R, right.



**Figure 6.** The proportion of volunteers that showed significant activation in the different sessions of the selective attention paradigm, with increasingly more time on task. The figure displays significant whole-brain activation during the attention localizer and each communication session at one of five time intervals, in the same areas where activation was observed for the duration of the entire task. For each session, the proportions are calculated as the ratio of volunteers that showed significant activations to the total number of successful volunteers, 15, 13, and 14, respectively, for each session.

latter paradigm, a volunteer can convey either a “yes” or a “no” by engaging in two different types of mental imagery, during two independent scanning sessions, each lasting 5 min (total, 10 min). In contrast, in the selective attention paradigm, half the time is necessary, because the volunteer can convey either word within the same scanning session (within 4.5 min).

In summary, comparison with the motor imagery task, the best-established fMRI technique for communication with behaviorally nonresponsive patients, suggests that the selective attention paradigm is highly suitable to BCI applications, not only for healthy individuals, but also for this patient group. Indeed, for applications with behaviorally nonresponsive patients, the difference in the underlying mental functions that each method uses lends to complementarity between the two. By probing different aspects of a patient’s spared cognition, each may confer advantages or disadvantages to the patient’s performance, depending on his/her spared capacities. The selective attention method demands sustained and selective attention to external events, as it relies on continuous monitoring and processing of fast-paced auditory information, as well as simple mental arithmetic (counting), which is rendered more difficult by the presence of highly relevant distracters. On the other hand, the motor imagery method does not require monitoring of external events. Rather, it relies on sustained attention to an internally generated mental process (motor imagery), at a self-guided pace. It is possible to

imagine that highly distractible patients, who may not be able to suppress bottom-up attention to stimuli in their environment, will find the selective attention task very challenging. Conversely, the generation of sustained motor imagery may prove harder for some patients, as was the case for 13% of the healthy volunteers studied here.

The selective attention task enabled almost half of the volunteers (47%) to convey their answer to a question within only 1.4 min of scanning, suggesting that this technique is suited for BCI communication within a reasonable amount of time. Moreover, the variation in scanning time required for response selection across volunteers further suggested that the application of this method in the real-time fMRI environment (Caria et al., 2012), where scanning time can be individually tailored, would lead to improved efficiency for individual participants. Although real-time communication through BCI devices is typically considered the domain of more time-sensitive neuroimaging technologies, such as EEG (Min et al., 2010), the brief scanning times obtained with this paradigm suggest it holds promise for effective brain-based communication in the MRI environment. The rapid response detection combined with the lack of a pretraining requirement render its overall time demands considerably shorter than those of binary EEG-based BCI systems. The temporal latency these systems require to decode brain responses is not dissimilar to existing fMRI techniques. Due to variability in spontaneous EEG activity, a long training phase is needed before participants can effectively generate commands for a BCI system (Iversen et al., 2008). Even if participant training times are reduced by powerful machine learning algorithms, relatively long acquisition times for classifier training persist (e.g., 20–30 min) (Blankertz et al., 2007). By comparison, the fMRI paradigm reported here requires, at most, one-quarter of that time (i.e., 1.1–4.5 min for 47–100% of participants) to acquire a localizer that identifies the specific brain areas significantly responsive in each individual.

Furthermore, this fMRI technique compares favorably to binary EEG-based BCI systems with regard to classification accuracy. For example, using techniques of binary classification, a previous EEG study by Miner et al. (1998) reported 64–87% accuracy rates for four participants. Guger et al. (2003) reported >59% accuracy for 93.3% of 99 participants, and 90–100% accuracy for only 6.4% of the participants. More recently, Cruse et al. (2011) used motor imagery tasks with EEG to detect command following (but not communication) and reported accuracies of 60–91% for 75% of participants, and 44–53% for 25% of participants. By comparison, this fMRI technique achieved 100% accuracy for 80% of participants.

Although MRI is more costly than EEG, fMRI BCIs may offer novel and unique opportunities (Naci et al., 2012), especially for patients in a complete locked-in state (CLIS), who do not respond with EEG-based systems. CLIS patients, who have entirely lost all motor abilities (Bauer et al. 1979), have not been able to communicate via EEG BCIs (Birbaumer et al., 2008). The lack of a priori knowledge about their level of cognitive capacities and communicative intent, as well as their varying levels of arousal, constitute major hurdles in building appropriate BCI systems for communicating with this patient group. Hence, the strengths of the technique reported here, especially its ease of use, robustness, and rapid detection, may maximize the chances that any nonresponsive patient will be able to achieve brain-based communication.

## References

- Bardin JC, Fins JJ, Katz DI, Hersh J, Heier LA, Tabelow K, Dyke JP, Ballon DJ, Schiff ND, Voss HU (2011) Dissociations between behavioural and functional magnetic resonance imaging-based evaluations of cognitive function after brain injury. *Brain* 134:769–782. [CrossRef Medline](#)
- Bardin JC, Schiff ND, Voss HU (2012) Pattern classification of volitional functional magnetic resonance imaging responses in patients with severe brain injury. *Arch Neurol* 69:176–181. [CrossRef Medline](#)
- Bauer G, Gerstenbrand F, Rimpl E (1979) Varieties of the locked-in syndrome. *J Neurol* 221:77–91. [CrossRef Medline](#)
- Bidet-Caulet A, Fischer C, Besle J, Aguera PE, Giard MH, Bertrand O (2007) Effects of selective attention on the electrophysiological representation of concurrent sounds in the human auditory cortex. *J Neurosci* 27:9252–9261. [CrossRef Medline](#)
- Birbaumer N, Murguialday AR, Cohen L (2008) Brain-computer interface in paralysis. *Curr Opin Neurol* 21:634–638. [CrossRef Medline](#)
- Blankertz B, Dornhege G, Krauledat M, Müller KR, Curio G (2007) The non-invasive Berlin brain-computer interface: fast acquisition of effective performance in untrained subjects. *Neuroimage* 37:539–550. [CrossRef Medline](#)
- Boersma P, Weenink D (2012) Praat: doing phonetics by computer. Version 5.3.10. <http://www.praat.org/>.
- Boly M, Coleman MR, Davis MH, Hampshire A, Bor D, Moonen G, Maquet PA, Pickard JD, Laureys S, Owen AM (2007) When thoughts become action: an fMRI paradigm to study volitional brain activity in non-communicative brain injured patients. *Neuroimage* 36:979–992. [CrossRef Medline](#)
- Brett M, Anton JL, Valabregue R, Poline JB (2002) Region of interest analysis using an SPM toolbox. Abstract presented at the 8th International Conference on Functional Mapping of the Human Brain, June 2–6, Sendai, Japan. *Neuroimage* 16(2).
- Caria A, Sitaram R, Birbaumer N (2012) Real-time fMRI: a tool for local brain regulation. *Neuroscientist* 18:487–501. [Medline](#)
- Cruse D, Chennu S, Chatelle C, Bekinschtein TA, Fernández-Espejo D, Pickard JD, Laureys S, Owen AM (2011) Bedside detection of awareness in the vegetative state: a cohort study. *Lancet* 378:2088–2094. [CrossRef Medline](#)
- Duncan J (2010) The multiple-demand (MD) system of the primate brain: mental programs for intelligent behaviour. *Trends Cogn Sci* 14:172–179. [CrossRef Medline](#)
- Frank DW, Sabatinelli D (2012) Stimulus-driven reorienting in the ventral frontoparietal attention network: the role of emotional content. *Front Hum Neurosci* 6:1–5. [Medline](#)
- Fritz JB, Elhilali M, David SV, Shamma SA (2007) Auditory attention—focusing the searchlight on sound. *Curr Opin Neurobiol* 17:437–455. [CrossRef Medline](#)
- Ghatan PH, Hsieh JC, Petersson KM, Stone-Elander S, Ingvar M (1998) Coexistence of attention-based facilitation and inhibition in the human cortex. *Neuroimage* 7:23–29. [CrossRef Medline](#)
- Guger C, Edlinger G, Harkam W, Niedermayer I, Pfurtscheller G (2003) How many people are able to operate an EEG-based brain-computer interface (BCI)? *IEEE Trans Neural Syst Rehabil Eng* 11:145–147. [CrossRef Medline](#)
- Hugdahl K, Thomsen T, Erslund L, Rimol LM, Niemi J (2003) The effects of attention on speech perception: an fMRI study. *Brain Lang* 85:37–48. [CrossRef Medline](#)
- Iversen I, Ghanayim N, Kübler A, Neumann N, Birbaumer N, Kaiser J (2008) Conditional associative learning examined in a paralyzed patient with amyotrophic lateral sclerosis using brain-computer interface technology. *Behav Brain Funct* 4:53. [CrossRef Medline](#)
- Kawashima R, Imaizumi S, Mori K, Okada K, Goto R, Kiritani S, Ogawa A, Fukuda H (1999) Selective visual and auditory attention toward utterances—a PET study. *Neuroimage* 10:209–215. [CrossRef Medline](#)
- Kosslyn SM, Ganis G, Thompson WL (2001) Neural foundations of imagery. *Nat Rev Neurosci* 2:635–642. [CrossRef Medline](#)
- Lau HC, Rogers RD, Haggard P, Passingham RE (2004) Attention to intention. *Science* 303:1208–1210. [CrossRef Medline](#)
- Lee JH, Marzelli M, Jolesz FA, Yoo SS (2009) Automated classification of fMRI data employing trial-based imagery tasks. *Med Image Anal* 13:392–404. [CrossRef Medline](#)
- Menon V, Uddin LQ (2010) Saliency, switching, attention and control: a network model of insula function. *Brain Struct Funct* 214:655–667. [CrossRef Medline](#)
- Min BK, Marzelli MJ, Yoo SS (2010) Neuroimaging-based approaches in the brain-computer interface. *Trends Biotechnol* 28:552–560. [CrossRef Medline](#)
- Miner LA, McFarland DJ, Wolpaw JR (1998) Answering questions with an electroencephalogram-based brain-computer interface. *Arch Phys Med Rehabil* 79:1029–1033. [CrossRef Medline](#)
- Monti MM, Vanhaudenhuyse A, Coleman MR, Boly M, Pickard JD, Tshibanda L, Owen AM, Laureys S (2010) Willful modulation of brain activity in disorders of consciousness. *N Engl J Med* 362:579–589. [CrossRef Medline](#)
- Nachev P, Rees G, Parton A, Kennard C, Husain M (2005) Volition and conflict in human medial frontal cortex. *Curr Biol* 15:122–128. [CrossRef Medline](#)
- Nachev P, Wydell H, O'Neill K, Husain M, Kennard C (2007) The role of the pre-supplementary motor area in the control of action. *Neuroimage* 36:T155–T163. [CrossRef Medline](#)
- Naci L, Monti MM, Kübler A, Sorger B, Goebel R, Kotchoubey B, Owen AM (2012) Brain computer interfaces for communication with non-responsive patients. *Ann Neurol* 72:312–323. [CrossRef Medline](#)
- Ortuño F, Ojeda N, Arbizu J, López P, Martí-Clement JM, Peñuelas I, Cervera S (2002) Sustained attention in a counting task: normal performance and functional neuroanatomy. *Neuroimage* 17:411–420. [CrossRef Medline](#)
- Owen AM, Coleman MR, Boly M, Davis MH, Laureys S, Pickard JD (2006) Detecting awareness in the vegetative state. *Science* 313:1402. [CrossRef Medline](#)
- Peelle JE, Johnsrude IS, Davis MH (2010) Hierarchical processing for speech in human auditory cortex and beyond. *Front Hum Neurosci* 4:51. [Medline](#)
- Peigneux P, Orban P, Baetens E, Degueldre C, Luxen A, Laureys S, Maquet P (2006) Offline persistence of memory-related cerebral activity during active wakefulness. *PLoS Biol* 4:e100. [CrossRef Medline](#)
- Ptak R (2011) The frontoparietal attention network of the human brain: action, saliency, and a priority map of the environment. *Neuroscientist* 18(5):502–515.
- Sorger B, Reithler J, Dahmen B, Goebel R (2012) A Real-Time fMRI-based spelling device immediately enabling robust motor-independent communication. *Curr Biol* 22:1333–1338. [CrossRef Medline](#)
- Tiitinen H, Sinkkonen J, Reinikainen K, Alho K, Lavikainen J, Näätänen R (1993) Selective attention enhances the auditory 40-Hz transient response in humans. *Nature* 364:59–60. [CrossRef Medline](#)
- Woldorff MG, Gallen CC, Hampson SA, Hillyard SA, Pantev C, Sobel D, Bloom FE (1993) Modulation of early sensory processing in human auditory cortex during auditory selective attention. *Proc Natl Acad Sci U S A* 90:8722–8726. [CrossRef Medline](#)
- Worsley KJ, Marrett S, Neelin P, Vandal AC, Friston KJ, Evans AC (1996) A unified statistical approach for determining significant signals in images of cerebral activation. *Hum Brain Map* 4:58–73. [CrossRef](#)