

Published in final edited form as:

Neuroimage. 2011 April 1; 55(3): 1044–1053. doi:10.1016/j.neuroimage.2010.12.085.

Spatio-temporal activity in real time (STAR): Optimization of regional fMRI feedback

Jeremy F. Magland¹, Christopher W. Tjoa², and Anna Rose Childress²

¹Department of Radiology, University of Pennsylvania School of Medicine, Philadelphia, PA 19104, USA

²Department of Psychiatry, University of Pennsylvania School of Medicine, Philadelphia, PA 19104, USA

Abstract

The use of real-time feedback has expanded fMRI from a brain probe to include potential brain interventions with significant therapeutic promise. However, whereas time-averaged blood oxygenation level-dependent (BOLD) signal measurement is usually sufficient for probing a brain state, the real-time (frame-to-frame) BOLD signal is noisy, compromising feedback accuracy. We have developed a new real-time processing technique (STAR) that combines noise-reduction properties of multi-voxel (e.g., whole-brain) techniques with the regional specificity critical for therapeutics. Nineteen subjects were given real-time feedback in a cognitive control task (imagining repetitive motor activity vs. spatial navigation), and were all able to control a visual feedback cursor based on whole-brain neural activity. The STAR technique was evaluated, retrospectively, for five a priori regions of interest in these data, and was shown to provide significantly better (frame-by-frame) classification accuracy than a regional BOLD technique. In addition to regional feedback signals, the output of the STAR technique includes spatio-temporal activity maps (movies) providing insight into brain dynamics. The STAR approach offers an appealing optimization for real-time fMRI applications requiring an anatomically-localized feedback signal.

Keywords

Real-time feedback; fMRI; Self-regulation; Neurofeedback; Blood oxygen level-dependent; Spatio-temporal activity

1. Introduction

Helping individuals to control their brain function through biofeedback has long-standing appeal. Brain biofeedback began by utilizing EEG (electroencephalogram, e.g., cortical rhythms, slow or evoked cortical potentials, etc.) [1] which features good temporal sensitivity, but has relatively poor spatial resolution and limited ability to probe deeper

© 2011 Elsevier Inc. All rights reserved.

Direct correspondence to: Jeremy F. Magland, Department of Radiology, University of Pennsylvania, 1 Founders Pavilion, MRI Education Center, 3400 Spruce Street, Philadelphia, PA 19104, Phone: 215-662-4711, Jeremy.Magland@gmail.com.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

(subcortical) brain regions critical for processing of affect, mood and motivation. Functional magnetic resonance imaging (fMRI), along with advances in computing power, has re-energized the field of brain biofeedback, with early studies demonstrating the possibility of feedback from subcortical as well as cortical regions – with good spatial localization [2–6]. The techniques employed by Weiskopf [6], De Charms [4], and others [7–12] provide subjects with real-time (RT) updates on their localized brain activity, as determined by blood oxygenation level-dependent (BOLD) signal changes within a target region of interest. Impressively, subjects have demonstrated the ability to regulate brain regions involved in the control of chronic pain [13], tinnitus [8], emotion [14,15], and movement [16].

Although these demonstrations are encouraging, the regional BOLD technique presents significant unresolved challenges. A primary limitation of BOLD signal is its susceptibility, not only to drift [17], but also to physiologic noise, including non-cognitive processes such as motion and respiration as well as cognitive processes that are unrelated to the task(s) of interest. In conventional fMRI, such effects pose less of a problem, as they are averaged out over a typical 10–15 minute scan time, comprising a number of independent task repetitions. The output of such experiments usually consists of a statistical map of spatially-resolved brain activation. However, for real-time feedback fMRI, temporally resolved information (i.e. corresponding to single events) is of greater importance because it is the temporally-resolved feedback that must be used by the subject during training.

LaConte et al [18] addressed the issue of noisy feedback through an alternative approach derived from a classifier discriminating between two contrasting whole-brain states. The classifier (dependent on BOLD signal fluctuations throughout the brain) was developed using machine learning during a pre-scan in which the subject was told to alternate between two distinct sets of thoughts. Because it does not require region of interest (ROI) selection, the method is robust, automatically exploiting patient-specific brain activity patterns. Furthermore, the technique is less susceptible to thermal noise since it uses information obtained from the entire brain (averaging effect), and has the potential to automatically suppress physiologic and motion-related noise via degrees of freedom in the model. These properties make the whole-brain approach well-suited to applications such as communicating with otherwise unresponsive patients with brain injuries [19,20]. The disadvantage is that the approach lacks regional specificity.

Motivated by the need for regional specificity in therapeutics targeting particular brain regions, we have developed a new real-time processing technique that retains the robustness of the whole-brain classifier approach (i.e., with suppression of noise), while providing spatial specificity. As in the whole-brain method, the technique uses a classifier-training period at the beginning of the scan to develop a model for feedback. However, rather than developing a single classifier across the entire brain, thousands of individual (spatially-localized) regression models are developed and then combined using a principal component analysis to obtain real-time tracking of spatially-resolved brain activity that is associated with the task design. Though computationally demanding, this optimization can be incorporated into standard real-time fMRI feedback protocols with conventional computing capabilities, facilitating feedback-based therapeutics.

2. Methods

2.1. Participants

The participants in the study were a diverse group of 19 adults (12 male) who signed witnessed, informed consent to be scanned within protocols approved by the Office of Human Research at the University of Pennsylvania School of Medicine. Ages ranged from

21–50 years; education ranged from 12–24 years. Five were abstinent, stabilized cocaine patients (labeled P1-P5) enrolled in other (non-medication) research protocols (a secondary objective of this optimization study was to determine whether a simple cognitive control task employing real-time fMRI feedback could be tolerated and performed by clinical populations). The remaining subjects are labeled C1-C14.

2.2. Data Acquisition and Processing

All imaging was performed on a 3 Tesla MRI scanner (Tim Trio; Siemens, Erlangen, Germany). Functional data were acquired using 2D echo-planar imaging with the following parameters: TE = 31 ms, TR = 2 s, flip angle = 90°, resolution = (3.6 mm)², FOV = (230 mm)², 32 slices, slice thickness = 4.5 mm with zero inter-slice spacing (image array size 64×64×32). Structural images were also obtained using a Magnetization Prepared Rapid Gradient Echo (MPRAGE) imaging sequence with the following parameters: TR = 1.5 s, TE = 3.7 ms, TI = 900 ms, resolution = (1 mm)³, FOV = 256×192×160 mm³.

Reconstructed functional images (in DICOM format) were exported immediately after acquisition to a stand-alone fMRI processing computer via an internal network connection. All processing was performed using a custom integrated software system running on an internal network of computers (see Fig. 1). The software was written in C++ and used the Qt framework (Nokia Corporation, Finland). Apart from spatial smoothing (only used for the whole-brain classifier) and prospective drift correction (described below), no further image preprocessing was performed. In particular, no temporal smoothing was applied, since this would result in loss of temporal resolution for the real-time feedback. Because the importance of motion correction (or image alignment) for this protocol was uncertain, the effect of motion correction was evaluated retrospectively (see Section 3.2.6).

2.3. Functional Task Design

Each imaging/feedback session consisted of a functional scan consisting of a classifier-training period (5–8 minutes) and a feedback period (ranging between 8 and 24 minutes) immediately following the classifier-training period. We used a real-time feedback protocol with two components known to activate functionally distinct, and geographically separate brain regions [19–22]. During the classifier-training period, subjects were told to alternate between two “thought” tasks: (1) “Imagine hitting a tennis ball, over and over again” (repetitive motor thoughts), and (2) “Imagine moving from room to room in a familiar building” (spatial navigation thoughts). Each task was performed for 30 seconds at a time, separated by 10 seconds of instructions. In this way, the classifier-training period consisted of 4–6 cycles, where each 80-second *cycle* included one period of each task. We will refer to the two tasks as “task A” and “task B”.

During the feedback period (immediately following classifier-training), the participant was provided with real-time whole brain-state feedback in the form of a continuously updating visual feedback marker (Fig. 1). The marker moved vertically according to the real-time fMRI prediction, reflecting the extent to which the subject was engaging in thoughts related to task B versus task A. The subject was instructed to alternate between trying to make the marker go up and down, comprising 6–24 full cycles. Since the study was focused on technique development, and not on demonstrating efficacy of the feedback, we did not attempt to evaluate whether displaying the feedback was beneficial to its control.

Feedback was provided in the form of horizontal solid white bars appearing behind the solid black instructions text box, visible in the margins on either side of the instructions on the projector screen (see Fig. 1). Each subject was told, prior to the scan, that the position of the

marker would reflect his/her brain state. Feedback was visible during the 30-second tasks, but not during instruction periods.

2.4. Automated Pixel Weighting and Masking

In order to effectively restrict the whole-brain classifier to pixels of interest (i.e. within the anatomic brain) an automated weighting algorithm was implemented on the basis of the functional data itself. Each pixel was preferentially included in the classification model on the basis of its temporal signal-to-noise ratio (TSNR), computed as

$$w_i = \text{TSNR}_i = \frac{\mu_i}{\sigma_i}, \quad (1)$$

where μ_i and σ_i are the mean and standard deviation, respectively, of the i^{th} pixel computed over all classifier-training frames. In this way, pixels outside the brain (with low signal) were given low weight, as were pixels with high signal variation, such as in the eyes and ventricles. To reduce computation time, only 50% of pixels (those with the highest weight) were retained in the classification model. Since the head only occupies a portion of the entire field of view, this conservative threshold essentially only excludes pixels outside the brain. We note that this threshold is not a critical feature of the algorithm since its only purpose is to reduce computation time by removing pixels which are expected to contribute very little to the model.

2.5. Real-Time Classification Algorithm

Brain-state classification was performed using a model of the form

$$F_n = \sum_{i=1}^M c_i w_i (X_{i,n} - \mu_i), \quad (2)$$

where F_n is the *non-drift-corrected* feedback signal at the n^{th} frame, c_i are the model coefficients, w_i are the weights described above, $X_{i,n}$ are the pixel intensities at the n^{th} frame, μ_i are the mean pixel intensities as above, and $i=1..M$ ranges over all retained pixels. The Partial Least Squares (PLS) regression algorithm was used to define the coefficients c_i in equation (2) to fit the box design function

$$\sum_{i=1}^M c_i w_i (X_{i,n} - \mu_i) \approx \begin{cases} +1 & i \in \text{Task A} \\ -1 & i \in \text{Task B} \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

for $n = 1..N_{\text{training}}$ (N_{training} denoting the number of training frames). Our choice to use PLS regression for whole-brain classification was motivated by the algorithm's speed as well as its robustness with respect to very large sets of predictor variables [23, 24]. The idea of PLS regression is to derive a small number of orthogonal components (or factors) that have high covariance with the dependent variable. These components, which are in turn linear combinations of the original predictor variables, are then linearly combined to predict the dependent variable in a standard regression. In this study we chose to use four PLS components as we found that using more than these did not improve the models.

2.6. Drift Correction

We compared three different techniques (two retrospective and one prospective) for low-frequency drift correction in our study. Prospective correction for baseline drift (i.e., the method used for real-time feedback) was performed using the following simple recursion:

$$\begin{aligned}\tilde{F}_n &= F_n - B_n \\ B_n &= B_{n-1} + \alpha(F_n - B_{n-1}) \quad (n > 1) \\ B_1 &= 0,\end{aligned}\tag{4}$$

where \tilde{F}_n is the adjusted feedback, and the parameter $0 < \alpha < 1$ determines how quickly the baseline is adjusted. Larger values of α result in a more rapid baseline correction but can also have an artificial attenuating effect on the feedback signal. Because this was a real-time study, we needed to select the value of α prospectively. In the absence of extensive pilot data for this task, α was selected heuristically as 0.025, the rationale being that this is equal to the reciprocal of the number of frames per cycle (i.e., $0.025 = 1/40$). Two retrospective techniques were also investigated, including simple cycle mean subtraction in which the mean of each 40-frame cycle was subtracted from the data prior to analysis, and linear detrending in which sliding-window linear fits were subtracted from the data (the duration of the sliding window was varied between 20 and 100 seconds). In addition, various values of the prospective correction parameter α were also compared in the retrospective analyses.

2.7. Spatio-Temporal Activity Mapping

In addition to the whole brain state classifier used to provide real-time feedback for this study, spatially-resolved activity patterns were also generated for each functional scan. The technique, which we call *spatio-temporal activity in real time (STAR)*, yields a time-resolved map of brain activity associated with the tasks. The output of the STAR processing (see Fig. 2) includes a training STAR map, available toward the end of the initial classifier-training period, serving as a functional localizer for manual selection of a target region of interest. Output during the feedback period includes dynamic (frame-by-frame) activation maps (STAR movies) as well as regional time series that can be used for real-time regional neurofeedback.

Note that in this study, only the *whole-brain RT feedback* was generated at the time of the scan. Other processing outputs (i.e., *training and feedback STAR maps, regional STAR RT feedback* and *STAR movies*) were computed retrospectively. These techniques are described in more detail in the sections 2.8–2.10.

2.8. STAR Algorithm

The STAR algorithm is a method for generating a (spatially-resolved) statistical model for discriminating between the two cognitive tasks on the basis of training data acquired during the initial (classifier-training) portion of the scan. At each spatial position r and time frame t during the feedback period, the model estimates a probability $\text{Pr}(r;t)$, representing the likelihood that the brain (nearby r) is in state B (versus A). For display and reporting purposes these probabilities are transformed into log-odds ratios (logits) and multiplied by a sign (± 1) according to

$$Lo(r;t) = v(r) \cdot \log\left(\frac{\text{Pr}(r;t)}{1 - \text{Pr}(r;t)}\right).\tag{5}$$

The sign factor ($v(r) = \pm 1$) is, for practical reasons, chosen to ensure that $Lo(r;t)$ is *positively* correlated with the underlying BOLD signal during the training period. During feedback, positive values of $Lo(r;t)$ can then be interpreted as being associated with increased BOLD signal. Without this correction, all quantities would be positively correlated with the task design function.

The model is developed on the initial training data using a three-step process. In the first step, a linear regression is performed at each pixel r to predict the task (A or B) on the basis of the pixel intensity (BOLD) data within a 3×3 pixel neighborhood of r . For this study we used PLS regression (with 4 components), but other methods can be used as well. Prior to regression, the time series for each pixel was drift-corrected using the method described above, except that the initial baseline value was selected as the mean of the first 40 frames. The first step yields 9 regression coefficients per pixel. In the second step, these coefficients are scaled to produce predicted log-odds ratios using K -fold cross-validation, where K is the number of classifier-training cycles. Specifically, individual cycles are excluded, one at a time, from the classifier-training data, and the model is developed on the basis of the remaining data and applied to each excluded cycle. The set of cross-validated predictions are then used to obtain the scaling factor

$$C_r = \frac{|m_B - m_A|}{s^2}, \quad (6)$$

where m_A and m_B are the within-task mean predictions, and s^2 is the within-task estimated variance of the predictions. Thus, the coefficients are scaled by C_r to obtain a model for estimating the log odds ratio $Lo(r;t)$ during the feedback period. Note that formula (6) is motivated by the case where the within-task predictions are normally distributed (with equal variance).

The third and final step makes use of principal component analysis (PCA) to reduce the noise in the predicted logits. Let C_i ($i=1..L$) be the first L principal components of the set

$$S_{\text{training}} = \{Lo_{\text{training}}(r;\cdot)\} \quad (7)$$

of cross-validated predicted logits during classifier-training. In this study we used $L = 20$ components as this seemed sufficient to capture most (i.e., >95%) of the variation in the initial data sets. Then

$$Lo_{\text{training}}(r;t) \approx \sum_{i=1}^L a_{r,i} C_i(t), \quad (8)$$

where $a_{r,i}$ are the component loadings. The components are then approximated in terms of the predicted logits (e.g., using PLS regression) as

$$C_i(t) \approx \sum_r b_{i,r} Lo_{\text{training}}(r;t), \quad (9)$$

where the sum is over all pixels. Finally, the *filtered* predicted logits (i.e. STAR output) for the feedback frames is defined as

$$\text{STAR}(r;t) := \sum_{i=1}^L a_{r,i} \bar{C}_i(t), \quad (10)$$

where

$$\bar{C}_i(t) := \sum_r b_{i,r} Lo(r;t). \quad (11)$$

2.9. Static and Dynamic STAR Maps

The output of the STAR algorithm is a spatio-temporally resolved map (dynamic STAR map, or movie) of logits predicting the brain state (A or B) throughout the functional scan. The dynamic map can be viewed in real time, during the scan (or retrospectively), to visualize the various brain regions' dynamic participation in the tasks. In addition, we derive a static activation map (STAR map), similar to a parametric map, using the formula

$$\text{STARMAP}(r) := \text{logit} \left[\frac{1}{N_{AF} + N_{BF}} \left(\sum_{t \in AF} \text{logit}^{-1}(-\text{STAR}(r;t)) + \sum_{t \in BF} \text{logit}^{-1}(\text{STAR}(r;t)) \right) \right], \quad (12)$$

where AF and BF are the sets of time points in feedback tasks A and B , respectively, and $N_{AF} + N_{BF}$ is the total number of such time points. Put another way, this formula computes the average predicted probability of correct classification across all feedback cycles and then converts the result back to logit units for display. Similarly, we derive a training STAR map by summing over frames in the classifier-training portion of the scan, using $Lo_{training}(r; t)$ in place of $STAR(r; t)$.

For visual comparison of static maps with a more conventional functional image analysis approach, we also used FEAT (FSL, FMRIB, Oxford, UK) to analyze each subject's session individually [25,26]. For each subject, motion corrected, brain extracted, temporally filtered functional images were entered into a general linear model predicting the design function (default preprocessing options).

2.10. Regional STAR Feedback

In addition to generating static and dynamic activity maps for retrospective (or real-time) exploration, the STAR data can also be used for regional feedback in real-time applications (see Fig. 2). A target region of interest is selected toward the end of the classifier-training run (or retrospectively) on the basis of the training STAR map, defined above. The regional feedback signal (in logit units) is then defined as the average value of $STAR(r; t)$ within the selected ROI.

2.11. Sub-region Analyses

For our sub-region analyses, we selected five *a priori* functional regions of interest directly related to the two "thought" tasks; these have been featured in several prior published reports [19–22]. For the repetitive motor task ("Imagine hitting a tennis ball, over and over again"), we selected the supplementary motor area (SMA), placing a 3×3 pixel rectangle on the image slice corresponding to the maximal functional activation (height and extent) of SMA in the training STAR map. For the spatial navigation task ("Imagine moving from room to room in a familiar building"), we selected two brain regions reliably activated by

this task, the parahippocampal place area, PPA (left and right) and the retrosplenial cortex, RC (left and right), again placing 3×3 pixel rectangles on the image slices showing the maximal functional activation (height and extent) on the training STAR map. In the case of weak or unilateral functional activations, the default ROI placement was based on underlying anatomical structure, with reference to Duvernoy [27]. For the purpose of the current study (comparing the classification accuracy of STAR vs. regional BOLD), ROI placement was done retrospectively. However, in practice, the STAR maps generated by the initial classifier-training run allow immediate selection of specific ROIs for the subsequent feedback period.

2.12. Repeat scans

Scans for two of the subjects were repeated. In one case (C7), the subject reported not being able to see the feedback or instructions due to not wearing eyeglasses. Two days later, the subject was provided with a pair of non-ferromagnetic eyeglasses and the scan was repeated. In a second case (P3), the subject explained (after the first scan) that he had never before held a tennis racket, and therefore could not relate to that task. The scan was repeated the next day with the tennis task being replaced by 'repeatedly casting a fishing line'. In these cases, only the second of the two scans have been included in the plots and statistical analyses of this paper. Thus, we performed all statistical analyses on $n=19$, including the second scan from these two individuals. However, to rule out bias caused by additional practice in these individuals, we also re-ran the statistical tests excluding both individuals ($n=17$) to confirm that the overall conclusions of the study did not differ.

3. RESULTS

3.1. Whole Brain Activity

3.1.1. Real-time Control of Whole-Brain Feedback—After a 5–8 minute classifier-training period, 18 of the 19 subjects were able to reliably control the feedback marker (up and down) using only their thoughts, with an average (per-frame) classification accuracy of 83%, and an average correlation with design of 0.60. Classification accuracy is the percent of time that the feedback marker is on the expected side of the midline of the projector screen throughout the feedback portion of the scan, whereas the latter measure is computed as the correlation of the feedback signal with the design function (−1 during the first task, +1 during the second task, and 0 during instructions). The one subject (C7) who was unable to control the feedback marker reported having difficulty seeing the instructions and feedback due to not wearing eyeglasses in the scanner. In a follow-up scan, using a pair of non-ferromagnetic eyeglasses, the same subject demonstrated improved control of the feedback (classification accuracy 73% over 240 feedback frames).

Whole-brain classifier results for all 21 scans are listed in Table 1. For the column labeled “Classification by 80 sec Cycle”, the values represent proportions of correctly classified cycles, determined by comparing the mean feedback signal between the two task periods. The *frame-by-frame classification accuracy* (hereafter referred to simply as *classification accuracy*) is also reported in this table for each scan. All but two subjects had classification accuracies greater than 70%, with the majority above 80%. Alternate measures for evaluating success in controlling the feedback marker, such as correlation with a design function or *t*-tests, were also considered. However, for simplicity, and since we observed that the various measures were highly correlated with one another, we only report classification accuracy in this paper. For example, Fig. 3 shows the strong positive correlation between classification accuracy and correlation with design ($R^2 = 0.95$). The reason that classification accuracy is numerically higher than correlation with design is simply a matter of ranges. Whereas accuracy theoretically ranges between 0 and 1 (with 0.5

representing no control and 1 representing perfect control), correlation with design ranges between -1 and 1 (with 0 representing no control and 1 representing perfect control).

3.1.2. Speed of Whole-Brain Classifier Formation—The three simulation plots of Fig. 4(B–D) illustrate the rapidity with which the discriminating classifier was formed, or, in other words, how many cycles were needed to later provide meaningful feedback. For example, in (C), only two cycles (160 seconds) of data were used to define the whole-brain classifier, and yet the resulting feedback was comparable to the actual feedback (A) based on six classifier-training cycles. On the other hand, using just one cycle of data (see D) did not yield a good classifier. The number of cycles required to form robust classifiers varied somewhat across individuals in this study. However, in most cases 3–6 minutes (2–4 cycles) of classifier training was sufficient for good cursor control during feedback. In other words, including additional data in the classifier definition model did not significantly improve the discriminatory properties of the feedback. For this reason, the classifier training period was reduced to four cycles after the data from the first six participants were examined (see Table 1).

3.1.3. Alternate Techniques for Whole-Brain Classification—As mentioned above, the PLS algorithm was chosen, in part, due to its low computational cost in handling large sets of predictor variables. Other techniques, most notably support vector machines (SVM) [18], have been shown to also perform well in this context. Whereas a thorough comparison of the various strategies is beyond the scope of this paper, we present here a simple comparison between PLS and SVM, using default regression parameters in the SVMlight software [28]. Fig. 5 reveals the high degree of correlation in classification results between the two techniques ($R^2 = 0.83$). The average classification accuracy using PLS was slightly, but significantly, higher than for SVM ($83\% > 79\%$; $p = 0.003$ in a 2-tailed paired t -test), suggesting that PLS is at least comparable to SVM in this application. The p -value was unchanged (to this precision) when the two repeat scans were excluded ($n=17$).

3.2. Regional Activity

3.2.1. Regional STAR Feedback—Although the whole-brain method provided the highest classification accuracy (83% on average), Fig. 6 and Table 2 show that the STAR method outperformed the unprocessed BOLD signal for regional feedback in all five sub-regions. This is evident in the median box plots (Fig. 6) as well as in the mean STAR classification accuracies (Table 2), which ranged between 69% and 79%. The mean STAR classification accuracies were significantly higher than the BOLD accuracies in all five *a priori* regions; p -values were less than 0.01 (2-tailed paired t -tests). As can be seen in representative subject C2 (Fig. 7), the STAR time series, in addition to being more accurate in predicting the task, are less noisy than the BOLD signal series. Similar results were found when the two repeat scans were excluded ($n=17$), with p -values increasing slightly as expected due to the reduced sample (largest p -value of 0.027 for SMA).

Fig. 8 demonstrates that regional STAR feedback provides spatially-localized information. The expectation is that classification accuracy for anatomical regions involved in the same function (e.g., PPA and RC for spatial navigation) should be inter-correlated. Indeed, this is demonstrated in plots A–C ($R^2 = 0.71$ – 0.84). In contrast, classification accuracies for regions involving distinct functions (e.g., SMA for motor repetition versus PPA for spatial navigation) should be relatively less correlated, as demonstrated in plots D–E ($R^2 = 0.26$ – 0.35).

3.2.2. Static STAR Maps—STAR maps (both training and feedback) are shown in Fig. 9 for two subjects (one with a clear response to both tasks, and one lacking a clear response to

the motor task) for purpose of illustration. As described above, the five *a priori* regions of interest were selected on the basis of the *training* STAR map (i.e. obtained during the initial portion of each scan), by selecting supra-threshold functional activations (“hot spots”) within the expected anatomic areas, provided that supra-threshold activations were present. In the absence of a hot spot (e.g., the SMA for Subject C11), the *a priori* anatomical region was used. Note that because the whole-brain classifier approach only demands that the two task-related brain states be distinct, the lack of clear, region-specific activation for one task or the other did not undermine the ability of the classifier to provide feedback, and both these subjects (C2 and C11) showed excellent control of the whole-brain feedback (90% and 92% accuracy, respectively). Presumably, the contrasts between the two brain states (for C2 and C11) were equally robust even though the imagined motor state may not have been distinguishable from rest in C11.

The color scale for the STAR maps is in logit (log odds ratio) units, allowing these maps to be directly compared between subjects and also between training/feedback, even though different numbers of cycles were involved in obtaining the logits. Specifically, Subject C2 had 6 training and 18 feedback cycles, whereas Subject C11 had 4 training and 6 feedback cycles (see Table 1). The colored pixels (i.e. $|\text{logit}| > 0.1$) are those most strongly predictive of task, as determined by the cross-validation and PCA filtering procedures of the STAR algorithm.

The overlap between participating regions in the training and feedback STAR maps reflects the robustness of the procedure, since these data were derived from temporally independent portions of the scans. Pixels that are highlighted in the training map, but not in the feedback map, correspond to false-positives, since such pixels were estimated as significant in the classifier-training scan, but did not validate during the feedback portion. However, note that the brightest (i.e. most reliable) pixels in the training map remain in the feedback map, particularly in the *a priori* regions of interest. The results of a conventional statistical analysis (FEAT) is shown for subject C2 in Fig. S8 demonstrating a high level of agreement with the STAR approach.

3.2.3. Dynamic STAR Movies—Example STAR movies (or dynamic STAR maps) showing the time-resolved brain activity for subjects C2 and C11 are provided in animations Sup1 and Sup2 at 5 frames per second ($\times 10$ speedup). Image sequences for these movies are also shown in Figs. Sup3 and Sup4 with every five frames averaged together. As already described above for the static STAR maps, subject C2 (Animation Sup1) shows robust activations associated with both tasks in the *a priori* regions of interest (SMA during “tennis”, and PPA/RC during “room-by-room” navigation), while subject C11 (Animation Sup2) does not evidence clear SMA activation during “tennis” thoughts. Though these example STAR movies were created retrospectively, the STAR program enables this viewing function in real-time, allowing real-time visual examination of the brain regions participating in the task. This enables flexible selection or de-selection of regions to be used for subsequent feedback during the scan. This kind of capability – whether using static STAR maps or dynamic STAR movies -- could be useful in ‘shaping’ paradigms, as the feedback can be incrementally shifted or ‘weighted’ to encourage participation by a target region.

To demonstrate the feasibility of performing the STAR classification in real time, an additional subject was scanned, and provided with regional STAR feedback (as opposed to whole-brain feedback) based on a manually selected ROI. A real-time screen capture of the feedback portion of the experiment (including the data available at the console as well as the feedback display) is shown in animation Sup5.

3.2.4. Computation Time—The model classification time for the STAR algorithm (i.e. the one-time processing at the end of the classifier-training period) was approximately 9 seconds (Intel Core 2 Duo CPU 2.53GHz). Most importantly, this computation time was within the duration of the 10-second instruction period. Specific details for the STAR computation are beyond the scope of this paper. However, we point out that it was necessary to perform initial model computations during the acquisition of classifier-training data (i.e. performing preliminary processing on data as it was acquired) in order to achieve this reduced model computation time. Once the model was defined, the total feedback computation time per frame was less than 0.2 seconds for regional STAR and less than 0.02 seconds for whole-brain feedback. Image reconstruction and data transfer took approximately 0.5 seconds for a total lag time of 1–3 seconds (considering acquisition time of 2 seconds per frame). If needed, this lag time can be reduced using parallel processing and faster data acquisition.

3.2.5. Drift correction evaluation—A comparison of outcomes from three drift correction techniques is shown in Fig. Sup6. The time constant for the prospective drift corrections (i.e., TR/α) was varied between 10 and 100 seconds, and the window duration for retrospective linear detrending was varied between 20 and 100 seconds. The retrospective method yielded the highest classification accuracies, with peak linear detrending performance for sliding-window duration of 60 seconds. The prospective correction method produced comparable results, with only 3–5% lower classification accuracies for $\alpha = 0.025$ used in this study. Furthermore, the prospective correction plot shows that $\alpha = 0.025$ is a near-optimal choice of drift correction parameter, and that the results are steady over a range of α values. Similar results were found for the regional BOLD method (although not displayed) with classification accuracies improving by less than 4% when comparing between retrospective and prospective drift correction methods.

3.2.6. Motion correction—As described above, no image pre-processing steps, aside from prospective drift correction and spatial smoothing for the whole-brain classifier, were performed. In particular, no motion correction (a.k.a. real-time alignment) was applied in our analyses. Because motion correction would add additional processing time (and the potential for unforeseen problems) to the real-time feedback, we elected to evaluate the importance of motion correction in our protocol, retrospectively. We used a freely available software package, MCFLIRT (FSL, FMRIB, Oxford, UK). Briefly, the software attempts to correct for subject motion by registering all images in the acquired time-series to the middle volume [29]. Mean classification accuracies before and after motion correction (for whole brain and for the five sub-regions) are compared in Fig. Sup7. Although post-correction accuracies were on average slightly higher, none were statistically significant (2-tailed paired t -tests).

4. DISCUSSION

We have described a novel technique (STAR) for providing robust regional real-time fMRI feedback, and have evaluated the method retrospectively in 19 subjects using a paradigm based on alternating sets of thoughts (motor repetition vs. spatial navigation). All subjects were able to achieve rapid and accurate cursor control with the (PLS-based) whole-brain feedback, with an average (per frame) classification accuracy of 83%. Retrospectively, the regional STAR feedback within five *a priori* regions of interest (SMA, PPA-R, PPA-L, RC-R, RC-L) was also determined to be relatively robust. With average classification accuracies above 70%, the STAR technique performed significantly better than a regional BOLD feedback approach for all five regions, while maintaining spatially localized information.

Our approach addresses the need for noise suppression in real-time feedback applications, particularly when the feedback is localized to a small sub-region of the brain. Conventional fMRI involves averaging of data collected over an entire scan in order to generate a parametric map showing locations of task-related activity. In contrast, real-time feedback applications require measurement of brain activity at every data frame (TR on the order of 2 seconds), and are thus more susceptible to noise. By combining information from pixels throughout the brain, the STAR method is able to achieve the desired noise reduction without significantly impacting regional specificity.

The STAR approach can be incorporated into a localized self-regulation training protocol (in which participants are instructed to alternate between two cognitive tasks) in the following manner. First, regions of interest are selected from the training STAR map (functional localizer), which is developed during the initial (3–6 minute) portion of the scan. The training STAR map displays (cross-validated) logits, or predicted probabilities, with bright pixels corresponding to regions where real-time feedback is expected to be robust (i.e. discriminating well between the two states). If a suprathreshold activation (bright spot) appears within a target region of interest (e.g., the insula for learning control of an emotional arousal response), then the operator would manually select that region. Next, the subject is shown STAR feedback from the selected region(s), and attempts to increase control of the feedback cursor over time. The progress can be seen (in real time during the scan) by the operator as well as by the participant using either the feedback time series or the dynamic feedback STAR map. Alternatively, if no suprathreshold pixels (bright spots) appear within the targeted anatomic region, the operator may elect to extend the classifier-training period or restart the scan (perhaps with updated instructions or with modified experimental parameters).

As mentioned in section 2.3, this study did not focus on evaluating feedback efficacy. Therefore, although the subjects could maintain good control of the feedback marker presented during the scan, it remains unclear whether they could have performed equally well without seeing the feedback marker at all. This will be the topic of future studies, where we expect various factors, including task difficulty, nature of the feedback, and fatigue/boredom will play a role in predicting efficacy.

A limitation of the technique is that it depends on a training pre-scan (or classifier-training portion of the scan) to develop the model. Fortunately, at least in the present application, this seems to only require 3–6 minutes of scan time. The large benefit of this small time investment is that the investigator can immediately determine which brain regions are (indeed) participating in the task, and can then choose the exact anatomical regions to be used for the feedback run, tailored to the individual. This empirical approach enables precise selection of regions for feedback, without requiring *a priori* omniscience regarding ROIs. This empirical determination of ROIs can facilitate the training process for complex tasks in which the relative participation of hypothesized brain regions, for a given individual, is uncertain -- and may actually be the target of the investigation. However, we note that the technique cannot track any activity during feedback which is not (at least to some extent) exhibited during the initial classifier-training period, possibly limiting the application of the methodology. This is in contrast to the BOLD method which does not require acquisition of training data. Nevertheless, in cases where there does not appear to be sufficient activity during training, the classifier-training session could either be restarted or lengthened in order to obtain a sufficiently robust classification model for feedback.

The current approach is best suited for localized brain regions that are geographically distinct (rather than overlapping, or anatomically interconnected), for which the subject has the potential ability to modulate activity (with respect to 30 second task periods), and whose

activity does not “carry over” beyond the instructed task period. Brain regions whose activity, once triggered, is persistent or even recruiting (e.g., limbic regions triggered by exposure to drug or sexual cues [30,31]) pose a challenge for all “comparison-based” real-time approaches, including (the simple subtraction between alternating states in) regional BOLD, classifier-based approaches, and STAR. For brain regions that “stay on” past the task period, other kinds of feedback strategies that do not depend upon baseline or comparator states (e.g., connectivity) may offer future utility as a feedback signal. Furthermore, the present technique cannot, in its current form, be applied to event-related task designs, or other paradigms where the shape of the hemodynamic response function must be considered.

Although the advantage over regional BOLD was found to be statistically significant in all regions, we note that we have only explored a few of the most common pre-processing options available for the BOLD signal time series. Other techniques such as physiologic noise reduction and advanced motion correction may further improve the performance of the BOLD method.

As a final note, for the majority of subjects, the visual feedback was perceived as helpful for maintaining cognitive control in the “tennis” vs. “spatial navigation” tasks. However, some individuals found it relatively easy to alternate brain states in this task, without feedback. For these subjects, the addition of feedback was initially perceived as unnecessary, and in some cases, distracting. This suggests that task difficulty and subject ability may determine whether real-time fMRI feedback is perceived as beneficial. Future real-time studies should take this into account, matching (when possible) task difficulty and feedback format to the target population.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by National Institutes of Health research grants K25-EB007646, R33-DA026114, P50-DA12756, and P60-DA005186, and by VISN 4 MIRECC.

References

1. Wolpaw JR, et al. Brain-computer interfaces for communication and control. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology* 2002;113(6):767–791. [PubMed: 12048038]
2. Yoo S-S, Jolesz FA. Functional MRI for neurofeedback: feasibility study on a hand motor task. *Neuroreport* 2002;13(11):1377–1381. [PubMed: 12167756]
3. Weiskopf N, et al. Physiological self-regulation of regional brain activity using real-time functional magnetic resonance imaging (fMRI): methodology and exemplary data. *NeuroImage* 2003;19(3):577–586. [PubMed: 12880789]
4. deCharms RC, et al. Learned regulation of spatially localized brain activation using real-time fMRI. *NeuroImage* 2004;21(1):436–443. [PubMed: 14741680]
5. Posse S, et al. Real-time fMRI of temporolimbic regions detects amygdala activation during single-trial self-induced sadness. *NeuroImage* 2003;18(3):760–768. [PubMed: 12667853]
6. Weiskopf N, et al. Self-regulation of local brain activity using real-time functional magnetic resonance imaging (fMRI). *Journal of Physiology, Paris* 2004;98(4–6):357–373.
7. Cohen MS. Real-time functional magnetic resonance imaging. *Methods (San Diego, Calif.)* 2001;25(2):201–220.

8. Haller S, Birbaumer N, Veit R. Real-time fMRI feedback training may improve chronic tinnitus. *European Radiology*. 2009
9. Johnston SJ, et al. Neurofeedback: A promising tool for the self-regulation of emotion networks. *Neuroimage* 2010;49(1):1066–1072. [PubMed: 19646532]
10. Papageorgiou, T., et al. Neurofeedback of two motor functions using supervised learning-based real-time functional magnetic resonance imaging; Conference Proceedings: ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference; 2009. p. 5377-5380.
11. Rota G, et al. Self-regulation of regional cortical activity using real-time fMRI: the right inferior frontal gyrus and linguistic processing. *Hum Brain Mapp* 2009;30(5):1605–1614. [PubMed: 18661503]
12. Yoo S-S, et al. Neurofeedback fMRI-mediated learning and consolidation of regional brain activation during motor imagery. *International Journal of Imaging Systems and Technology* 2008;18(1):69–78. [PubMed: 19526048]
13. deCharms RC, et al. Control over brain activation and pain learned by using real-time functional MRI. *Proceedings of the National Academy of Sciences of the United States of America* 2005;102(51):18626–18631. [PubMed: 16352728]
14. Caria A, et al. Volitional Control of Anterior Insula Activity Modulates the Response to Aversive Stimuli. A Real-Time Functional Magnetic Resonance Imaging Study. *Biological Psychiatry* 2010;68(5):425–432. [PubMed: 20570245]
15. Caria A, et al. Regulation of anterior insular cortex activity using real-time fMRI. *NeuroImage* 2007;35(3):1238–1246. [PubMed: 17336094]
16. Lee J-H, et al. Brain-machine interface via real-time fMRI: preliminary study on thought-controlled robotic arm. *Neuroscience Letters* 2009;450(1):1–6. [PubMed: 19026717]
17. Yan L, et al. Physiological origin of low-frequency drift in blood oxygen level dependent (BOLD) functional magnetic resonance imaging (fMRI). *Magnetic Resonance in Medicine: Official Journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine* 2009;61(4):819–827. [PubMed: 19189286]
18. LaConte SM, Peltier SJ, Hu XP. Real-time fMRI using brain-state classification. *Human Brain Mapping* 2007;28(10):1033–1044. [PubMed: 17133383]
19. Boly M, et al. When thoughts become action: an fMRI paradigm to study volitional brain activity in non-communicative brain injured patients. *NeuroImage* 2007;36(3):979–992. [PubMed: 17509898]
20. Owen AM, et al. Detecting awareness in the vegetative state. *Science (New York, N.Y.)* 2006;313(5792):1402–1402.
21. Monti MM, et al. Willful modulation of brain activity in disorders of consciousness. *The New England Journal of Medicine* 362(7):579–589. [PubMed: 20130250]
22. Weiskopf N, et al. Principles of a brain-computer interface (BCI) based on real-time functional magnetic resonance imaging (fMRI). *IEEE Transactions on Bio-Medical Engineering* 2004;51(6):966–970. [PubMed: 15188865]
23. Krishnan A, et al. Partial Least Squares (PLS) methods for neuroimaging: A tutorial and review. *NeuroImage*.
24. McIntosh AR, et al. Spatial pattern analysis of functional brain images using partial least squares. *NeuroImage* 1996;3(3 Pt 1):143–157. [PubMed: 9345485]
25. Woolrich MW, et al. Temporal autocorrelation in univariate linear modeling of FMRI data. *Neuroimage* 2001;14(6):1370–1386. [PubMed: 11707093]
26. Smith SM, et al. Variability in fMRI: a re-examination of inter-session differences. *Human Brain Mapping* 2005;24(3):248–257. [PubMed: 15654698]
27. Duvernoy, HM. *The Human Brain: Surface, Three-Dimensional Sectional Anatomy with MRI, and Blood Supply*. Springer; 1999.
28. Joachims, T. Making large-scale SVM learning practical, in *Advances in Kernel Methods - Support Vector Learning*. Cambridge, MA: MIT Press; 1999.
29. Jenkinson M, et al. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage* 2002;17(2):825–841. [PubMed: 12377157]

30. Childress AR, et al. Prelude to passion: limbic activation by "unseen" drug and sexual cues. *PLoS One* 2008;3(1):e1506–e1506. [PubMed: 18231593]
31. Childress AR, et al. Limbic activation during cue-induced cocaine craving. *The American Journal of Psychiatry* 1999;156(1):11–18. [PubMed: 9892292]

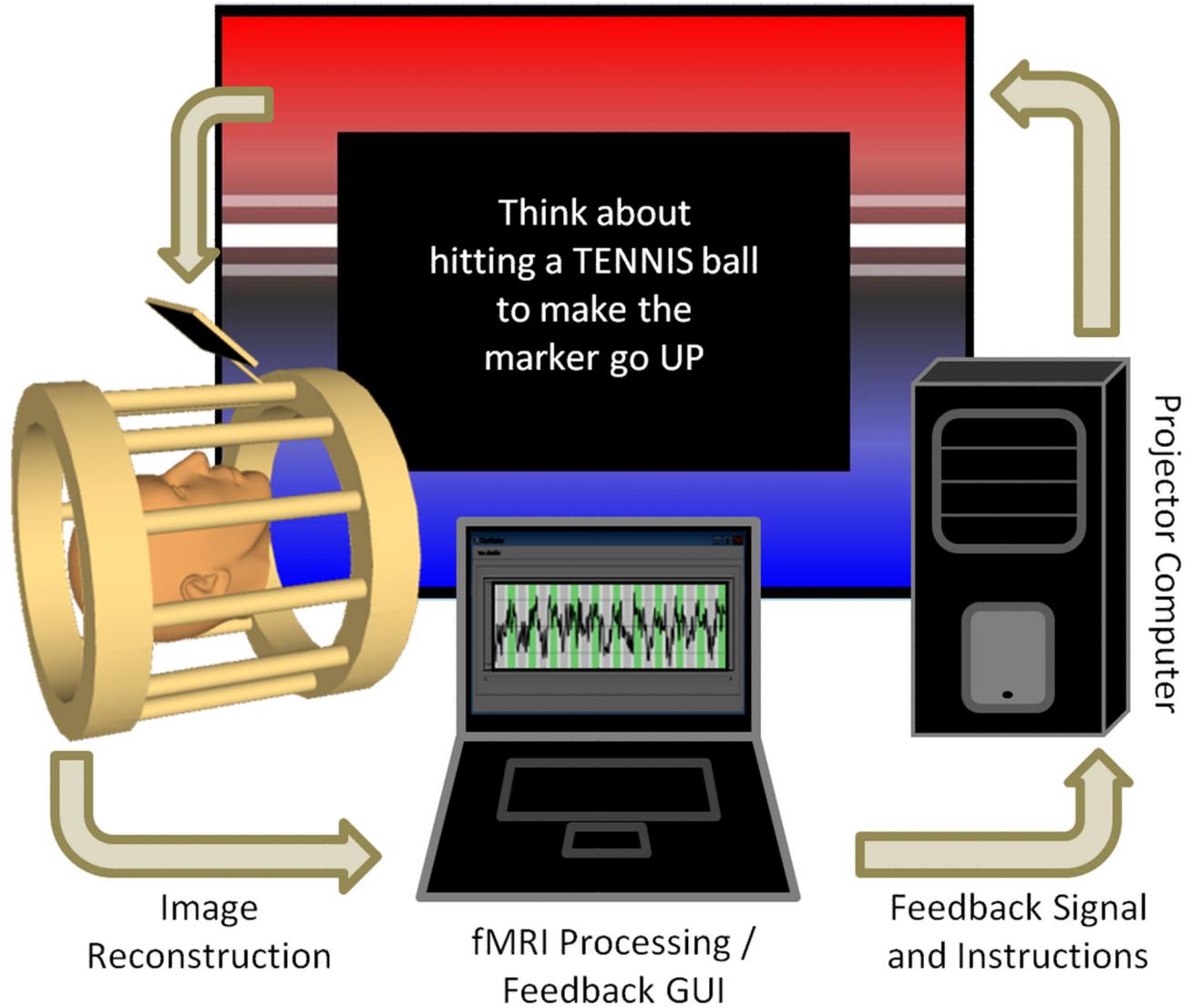


Figure 1. Real-time feedback loop comprising functional data acquisition, real-time data processing, and display of brain-state information back to the participant.

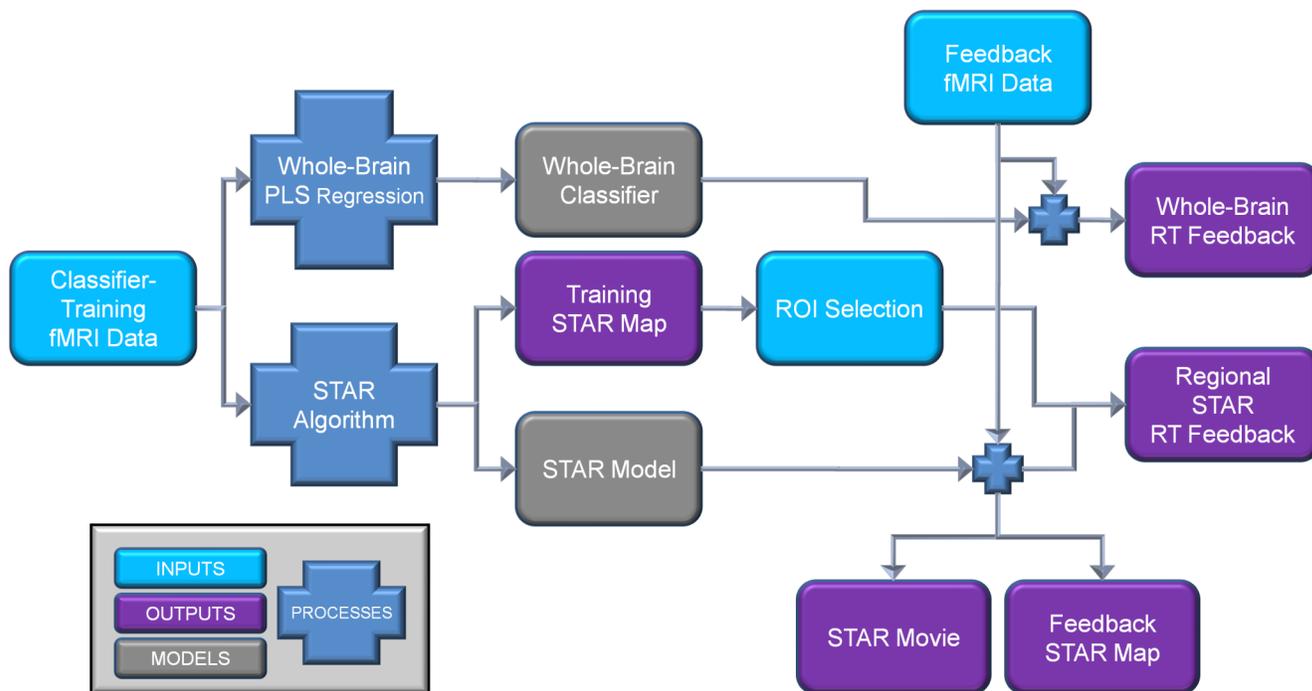


Figure 2. Flow diagram for the STAR processing pipeline.

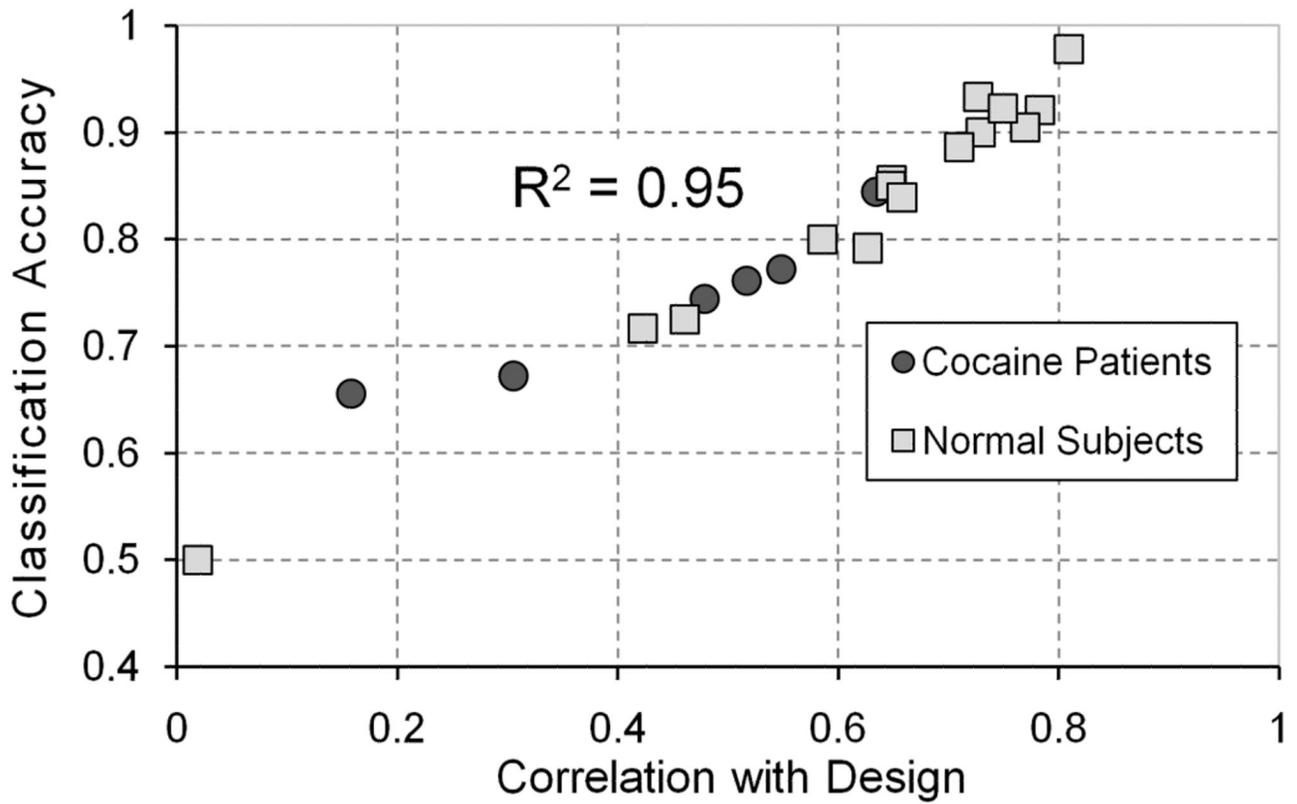


Figure 3.

Comparison of two measures of accuracy for whole-brain state feedback across 19 subjects. Note the high degree of correlation between the two measures ($R^2 = 0.95$). Classification accuracy is used as the primary measure of interest for this work, as it is more intuitive, reflecting the ability of the classifier to categorize the brain state from frame to frame (TR to TR) within each cycle.

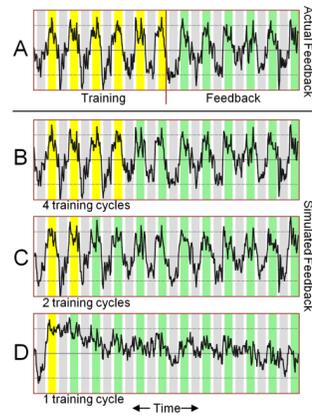


Figure 4.

Actual training/feedback cycle (A) and retrospective simulations (B–D) to test the minimal number of training cycles needed to train a brain classifier yielding useful real-time feedback. Gray and green bars represent tasks A and B, respectively, with 30 seconds per task separated by 10 seconds of instructions (white). Yellow bars represent task B during the classifier-training phase. A robust classifier was formed after only two cycles (160 seconds) of training, with minimal additional improvement beyond 4 training cycles (320 seconds). Illustration is based on data from subject C1, but rapid classifier formation was the norm across all subjects.

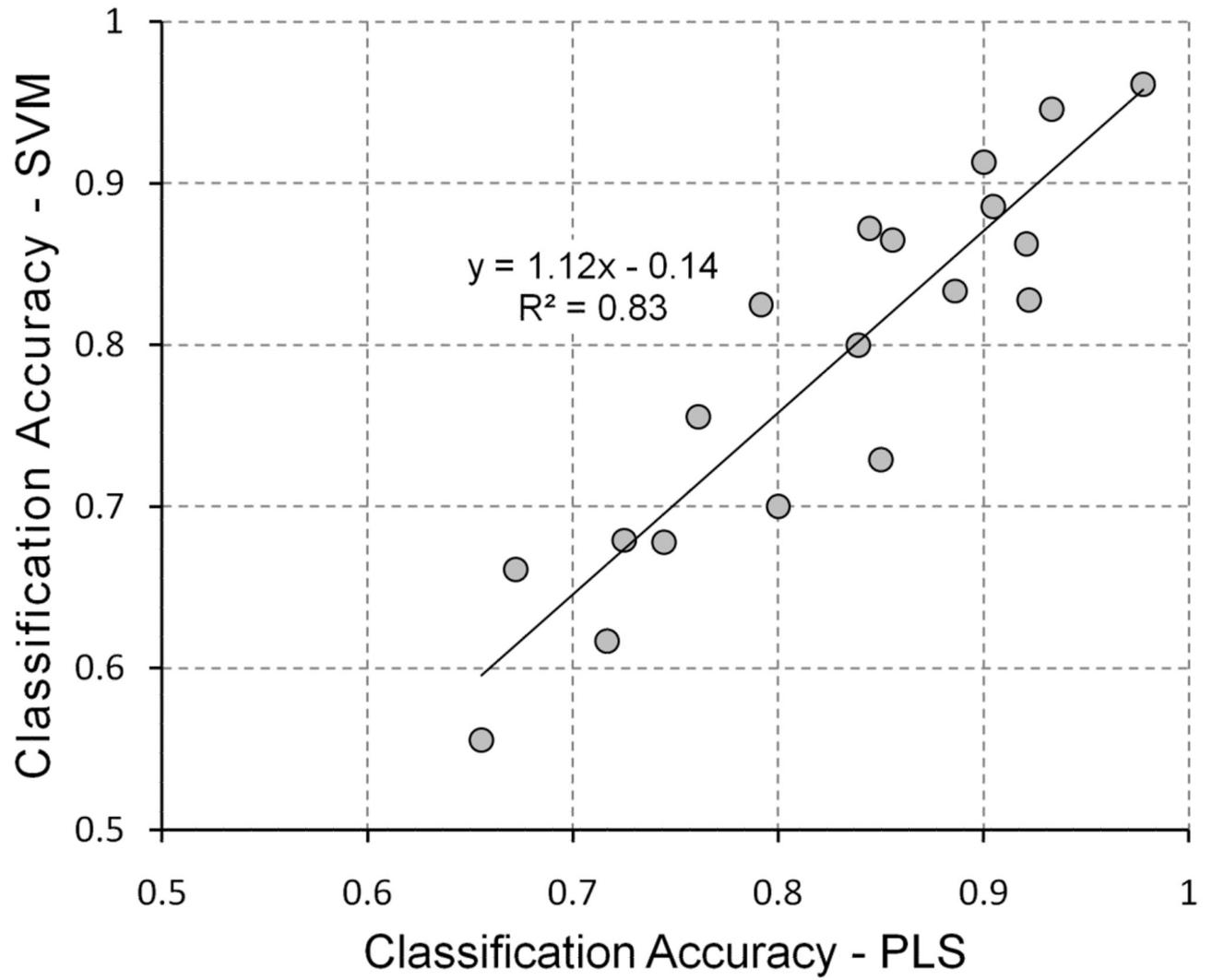


Figure 5.
Whole-brain classification accuracy for SVM vs. PLS regression methods.

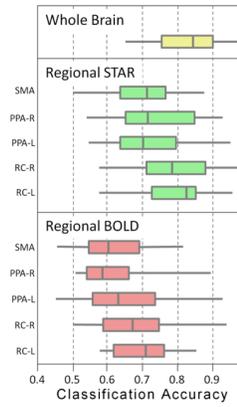


Figure 6. Group results for whole-brain, regional STAR, and regional BOLD techniques. Midlines represent medians; boxes denote the range between first and third quartiles.

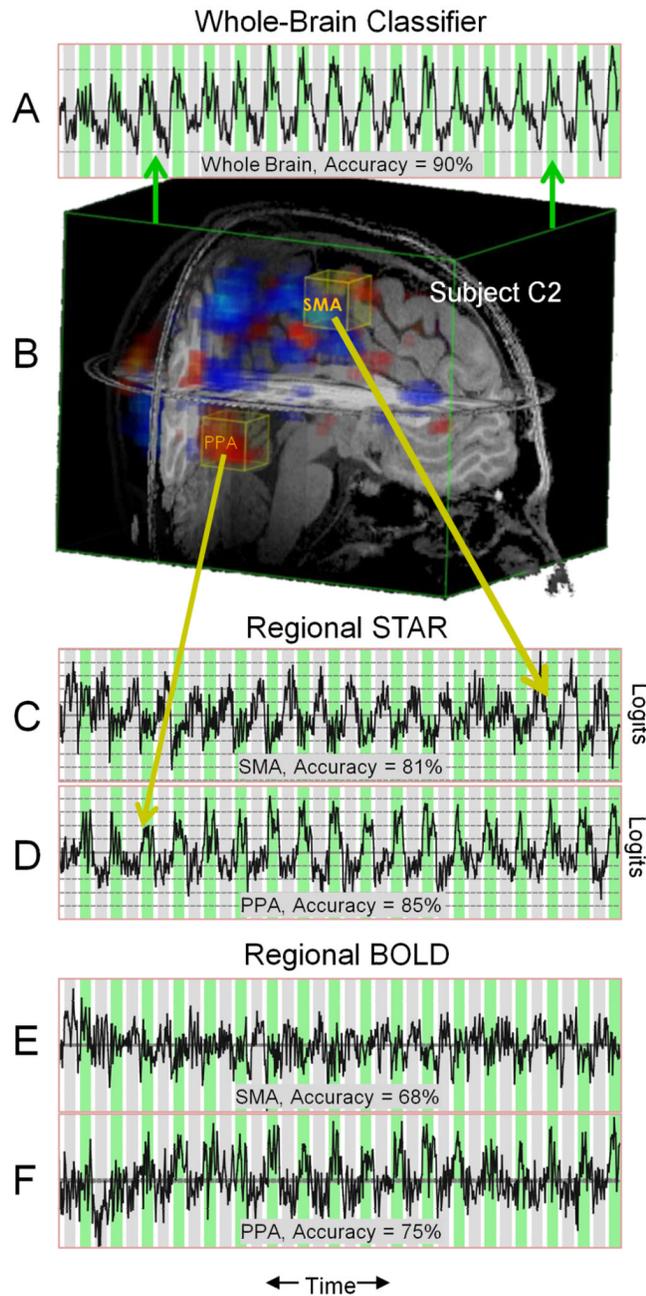


Figure 7.

Whole-brain and regional feedback time series for subject C2. As in Fig. 4, Gray and green bars represent tasks A and B, respectively. Whereas whole-brain feedback (A) has the highest level of accuracy and appears least noisy, regional feedback (C–E) has the advantage of providing spatially-localized information. The STAR approach to regional feedback (C–D) is more accurate and contains less noise than a pure BOLD approach (E–F). Horizontal lines in (C–D) mark integral logit units ranging from -5 to $+5$.

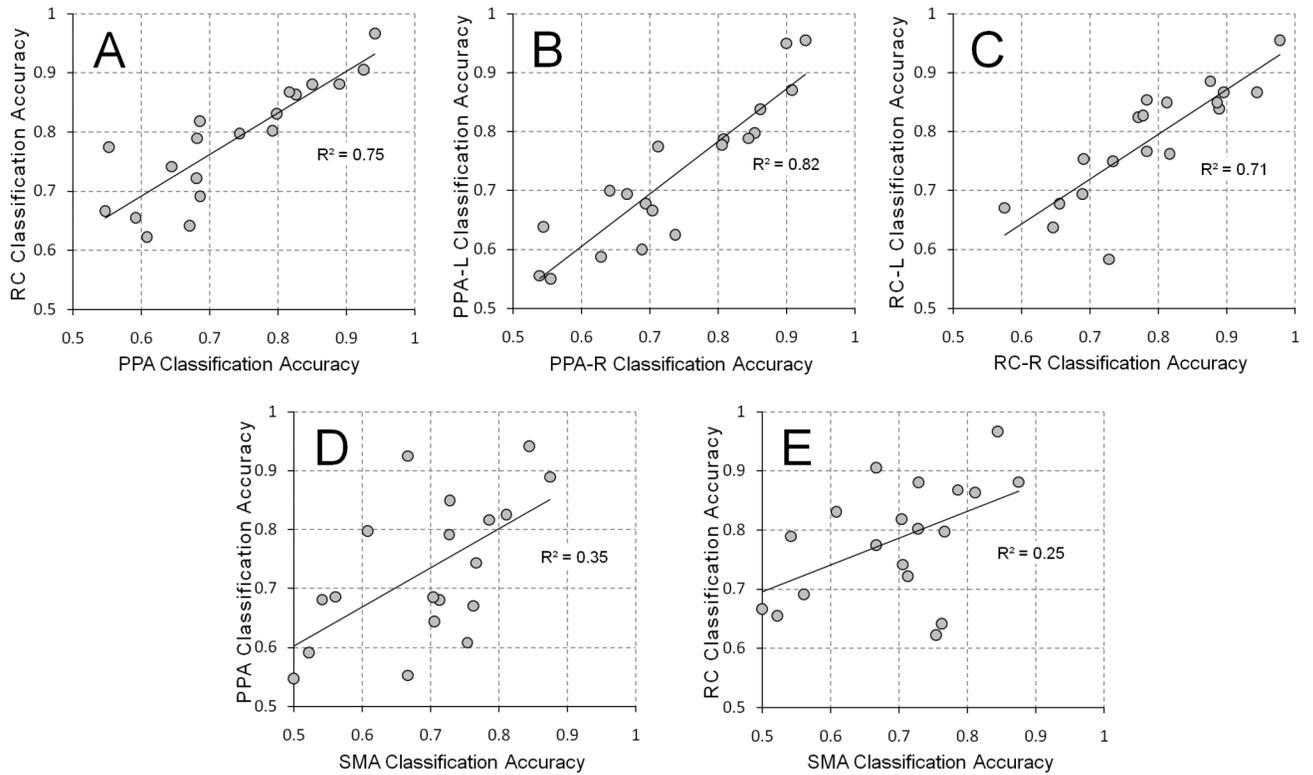


Figure 8. Pairwise comparisons of regional STAR feedback results between regions of interest. For A, D, and E, left and right classification accuracies were averaged for PPA and RC. As expected, the correlation is highest in A–C where the regions being compared are involved in the same functional tasks.

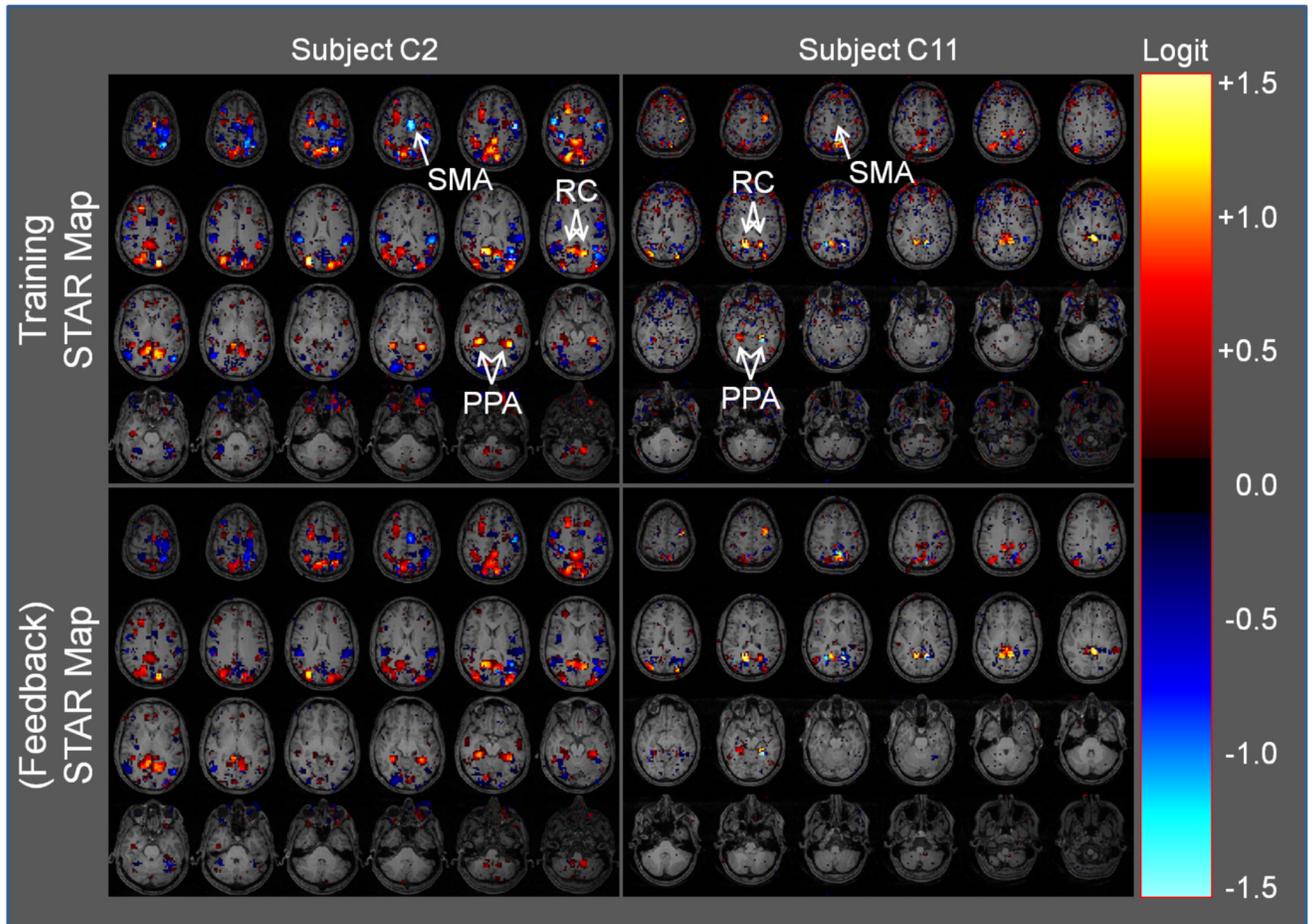


Figure 9. Static STAR maps (both training and feedback) for two subjects (one with a clear response to both tasks, and one lacking a clear response to the motor task). The color scale is in logit (log odds ratio) units, displayed with threshold $|\text{logit}| > 0.1$. There is clear overlap for suprathreshold pixels between the training and feedback maps, and these coincide with the *a priori* anatomical regions of interest, reflecting the robustness of the technique.

Whole-brain classification accuracy (by cycle and by frame) for all 21 scans. For two of the 19 subjects (C7 and P3) a second (repeat) scan was obtained, as explained in the text.

Table 1

Scan	# Training Cycles	Classification by 80 sec Cycle		Classification by 2 sec Frame	
		Classif. Accuracy	# Cycles Tested	Classif. Accuracy	# Frames Tested
C1	6	100%	18	86%	540
C2	6	100%	18	90%	540
C3	4	100%	12	89%	360
C4	4	100%	8	85%	240
C5	4	100%	8	92%	240
C6	4	100%	7	90%	210
C7-1	4	63%	8	50%	240
C7-2	4	88%	8	73%	240
C8	4	100%	8	80%	240
C9	4	100%	8	79%	240
C10	4	100%	8	93%	240
C11	4	100%	6	92%	180
C12	4	100%	6	84%	180
C13	4	100%	8	72%	240
C14	4	100%	6	98%	180
P1	6	100%	6	76%	180
P2	6	83%	6	67%	180
P3-1	6	100%	6	77%	180
P3-2	6	100%	6	84%	180
P4	4	83%	6	66%	180
P5	4	100%	6	74%	180

Table 2

Mean classification accuracy for five regions using BOLD and STAR methods. In each of the five regions, a 2-tailed *t*-test determined that the STAR approach performed significantly better than BOLD.

Region	Mean Classification Accuracy		t-score	p-value
	BOLD	STAR		
SMA	62.3% ± 2.4%	69.7% ± 2.5%	2.94	9.4E-03
PPA-R	61.7% ± 2.4%	73.8% ± 2.9%	4.90	1.6E-04
PPA-L	64.8% ± 2.8%	72.8% ± 2.8%	5.69	4.6E-05
RC-R	67.5% ± 2.6%	78.6% ± 2.5%	4.50	1.4E-04
RC-L	69.4% ± 2.0%	78.5% ± 2.2%	5.18	1.2E-04