| REPORT DOCUMENTATION PAGE | Form Approved OMB NO. 0704-0188 |
| --- | --- |



| 1. REPORT DATE (DD-MM-YYYY) | 2. REPORT TYPE | 3. DATES COVERED (From - To) |
| --- | --- | --- |
| 08-02-2019 | Final Report | 1-May-2012 - 31-Aug-2015 |

4. TITLE AND SUBTITLE
Final Report: Computational Modeling of Individual Brains

5a. CONTRACT NUMBER
W911NF-12-1-0160

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
611102

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

6. AUTHORS
Bruce Naylor, Andrew Floren, Risto Miikkulainen, David Ress

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES
University of Texas at Austin
101 East 27th Street
Suite 5.300
Austin, TX 78712 -1532

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSOR/MONITOR'S ACRONYM(S)
ARO

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
59476-LS.9

12. DISTRIBUTION AVAILIBILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT
Warfighters have benefited significantly from the enormous advances in digital technology over the past several decades. In contrast, too little of the considerable progress in neuroscience has been applied to improving warfighter performance. We believe this reflects the absence of digital technology that can help bridge the gap between neuroscience and digital systems, a gap that might be filled by constructing a computational model of the neuro-cognitive activity of the warfighter. We propose that such a model could be created by algorithms applied to

15. SUBJECT TERMS
Machine learning, sensitivity analysis, visual stream, attention, performance

| 16. SECURITY CLASSIFICATION OF: a. REPORT | b. ABSTRACT | c. THIS PAGE | 17. LIMITATION OF ABSTRACT | 15. NUMBER OF PAGES | 19a. NAME OF RESPONSIBLE PERSON |
| --- | --- | --- | --- | --- | --- |
| UU | UU | UU | UU | | David Ress |
| | | | | | 19b. TELEPHONE NUMBER |
| | | | | | 650-450-0455 |

Standard Form 298 (Rev 8/98)
Prescribed by ANSI Std. Z39.18

# RPPR Final Report
as of 28-Feb-2019

Agency Code:

Proposal Number: 59476LS **Agreement Number: W911NF-12-1-0160**

**INVESTIGATOR(S):**


**Name:** Bruce F. Naylor
**Email:** naylor@iat.utexas.edu
**Phone Number:** 512
**Principal:** N

**Name:** PhD David Ress
**Email:** ress@utexas.edu
**Phone Number:** 6504500455
**Principal:** Y

**Name:** Risto P. Miikkulainen Phd
**Email:** risto@cs.utexas.edu
**Phone Number:** 5124719571
**Principal:** N

Organization: **University of Texas at Austin**
Address: 101 East 27th Street, Austin, TX 787121532
Country: USA


DUNS Number: 170230239 EIN: 746000203
**Report Date:** 31-Oct-2015 Date Received: 08-Feb-2019
**Final Report** for Period Beginning 01-May-2012 and Ending 31-Aug-2015
**Title:** Computational Modeling of Individual Brains
**Begin Performance Period:** 01-May-2012 **End Performance Period:** 31-Aug-2015
**Report Term:** 0-Other
Submitted By: PhD David Ress Email: ress@utexas.edu
Phone: (650) 450-0455
**Distribution Statement:** 1-Approved for public release; distribution is unlimited.

**STEM Degrees:** 0 **STEM Participants:** 3

**Major Goals:** Aim 1: Stimulus design and fMRI measurements
The power of the synthesized neurometric models depends upon the design of the stimuli and the quality of the fMRI measurements. We will use 3D virtual worlds for our stimuli because they provide interactivity, complete control over the content, and are closer to what brains are designed to process than still imagery. These worlds will be designed to mimic the task for which the model is being synthesized. For example, training soldiers would use virtual worlds that mimic certain aspects of the combat experience relevant to the objectives of the training and/or treatment. To measure the induced brain activation patterns, we will pursue the development of fMRI protocols that optimize the measurement process.
Aim 2: Machine Learning
Neurometric modeling will be based on a mapping of fMRI images to brain states. The performance of a single task of interest will be modeled by a sequence of such states, thus introducing the time dimension. An example task would be “searching for snipers”. The first stage in the process will be to use statistical machine learning algorithms to build brain-state recognizers. These are pattern recognition modules that process fMRI images and output an estimate of the degree to which an fMRI image sequence contains evidence for a particular brain state. The output of the machine learning algorithms will be organized into a multi-dimensional space where the dimensions of the space capture task-specific semantics. The algorithms will map each fMRI volume to a point in this space. The sequence of points will then be organized into a finite number of states labeled with task-specific semantics.
Aim 3: Computational Model Synthesis
Neurometric models will utilize a new variety of Finite State Machines. They will be synthesized algorithmically from the output of the brain state recognizers. Given a sequence of observations of stimulus/response, this model can compute the most probable sequence of states that will explain the observations. It will also provide an interpretation of the sequence using the semantics of the task being modeled.

## RPPR Final Report
as of 28-Feb-2019

**Accomplishments:** Executive summary

1) Used various forms of machine learning to ascertain an easily quantified but complex visual stimulus observable from a set of fMRI data. Created various schemes to optimize accuracy. Employed a novel form of sensitivity analysis to relate the ML results back to brain anatomy. Published results in Frontiers.

2) Used a similar virtual environment to evaluate an internal cognitive variable, task difficulty. Subjects now had to perform a task, weapon detection, while viewing a cluttered and complex visual scene. Task difficulty was varied randomly from scene-to-scene. Support vector methods were developed and optimized to predict task difficulty from fMRI data obtained during the task. We obtained excellent results using support-vector regression, which offered 80-90% accuracy. These results motivated additional subsequent funding from the Veterans Administration to use similar methods as a means to treat PTSD. A manuscript for this work is in preparation for submission to Neuroimage.

3) Developed a third virtual environment and task to evaluate the use of machine-learning models of state-space. The task featured three states: navigation, search, and evasion. Some fMRI data was obtained, but loss of ARO funding during the final year of the project greatly slowed deployment of this project and subsequent analysis of the data. Preliminary results were nevertheless encouraging.

For details, please see the attached PDF.

**Training Opportunities:** Most training concerned of one graduate student, Andrew Floren, who is still working his dissertation. Its chapters will largely follow the three Accomplishments summarized above. The unexpected halt of ARO support in 2014 was one factor that caused Mr. Floren to take a job with private industry, deferring completion of his Ph.D. Fortunately, he is still participating in our subsequent follow-up project with the VA, and working to finish his dissertation at the same time.

We also trained several undergraduate students in our lab, who participated as research assistants. Training included fMRI methods, neuroanatomy and brain segmentation, and general scientific research methods.

**Results Dissemination:** One journal publication was obtained, a second is under preparation.

Six conference proceeding were presented during the time-frame of this award. See "Products" section.

Andrew Floren will eventually include all products in his dissertation at UT Austin.

**Honors and Awards:** Nothing to Report

**Protocol Activity Status:**

**Technology Transfer:** Nothing to Report

**PARTICIPANTS:**

**Participant Type:** Faculty
**Participant:** Bruce Naylor
**Person Months Worked:** **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member:
Other Collaborators:

**Participant Type:** Faculty
**Participant:** David Ress
**Person Months Worked:** **Funding Support:**
Project Contribution:

International Collaboration:
International Travel:
National Academy Member:
Other Collaborators:

**Participant Type:** Faculty
**Participant:** Risto Miikkulainen
**Person Months Worked:** **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member:
Other Collaborators:

**Participant Type:** Graduate Student (research assistant)
**Participant:** Andrew Floren
**Person Months Worked:** **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member:
Other Collaborators:

**Participant Type:** Undergraduate Student
**Participant:** Elizabeth Halfen
**Person Months Worked:** 3.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

**Participant Type:** Undergraduate Student
**Participant:** Kevin Rullman
**Person Months Worked:** 1.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

**ARTICLES:**

**Publication Type:** Journal Article Peer Reviewed: Y **Publication Status:** 0-Other
**Journal:** NeuroImage
Publication Identifier Type: Publication Identifier:
Volume: 0 Issue: 0 First Page #: 0
Date Submitted: Date Published:
Publication Location:
**Article Title:** Decoding behavioral performance accurately from fMRI data obtained in a realistic virtual environment
**Authors:** Andrew Floren, Bruce Naylor, Risto Miikulainen, David Ress
**Keywords:** fMRI, MVPA, machine learning, virtual environments, support-vector machine, vision
**Abstract:** A complex internal cognitive variable, task difficulty, was decoded using machine-learning methods from functional magnetic resonance imaging (fMRI) data. Human subjects performed an object-detection task on a cluttered and complex scene while fMRI data was collected. Task difficulty was varied randomly during the task from trivially easy to very difficult. After substantial training, support vector regression was able to predict task difficulty from the fMRI data with excellent accuracy, mean value of 81% across the seven subjects. A novel sensitivity analysis was then devised to project the weighted support vectors onto the brains of each subject to evaluate the neuroanatomical correlates of the performance estimates. Results show diverse involvement of brain regions to the complex task, with a preponderance of visual cortices, but also including a variety of fronto-parietal networks.
**Distribution Statement:** 2-Distribution Limited to U.S. Government agencies only; report contains proprietary info
Acknowledged Federal Support: **Y**

**CONFERENCE PAPERS:**

**Publication Type:** Conference Paper or Presentation **Publication Status:** 1-Published
**Conference Name:** OHBM 2013
Date Received: 14-Jun-2016 Conference Date: 19-Jun-2013 Date Published:
Conference Location: Seattle, WA
**Paper Title:** Classifying brain state from complex visual stimuli
**Authors:** Andrew Floren, Bruce Naylor, Risto Miikkulainen, David Ress
Acknowledged Federal Support: **Y**

**Publication Type:** Conference Paper or Presentation **Publication Status:** 1-Published
**Conference Name:** OHBM 2014 Annual Meeting
Date Received: 14-Jun-2016 Conference Date: 08-Jun-2014 Date Published:
Conference Location: Hamburg, Germany
**Paper Title:** Identifying brain regions associated with task performance in a complex virtual environment
**Authors:** Andrew Floren, Bruce Naylor, Risto Miikkulainen, David Ress
Acknowledged Federal Support: **Y**

**Publication Type:** Conference Paper or Presentation **Publication Status:** 1-Published
**Conference Name:** Neuroscience 2013, SFN Annual Meeting
Date Received: 14-Jun-2016 Conference Date: 09-Nov-2013 Date Published:
Conference Location: San Diego, CA
**Paper Title:** Using sensitivity analysis to interpret machine learning applied to fMRI
**Authors:** Andrew Floren, Bruce Naylor, Risto Miikkulainen, David Ress
Acknowledged Federal Support: **Y**

**Publication Type:** Conference Paper or Presentation **Publication Status:** 1-Published
**Conference Name:** OHBM 2015
Date Received: 12-Jun-2016 Conference Date: 16-Jun-2015 Date Published: 14-Jun-2015
Conference Location: Honolulu, Hawaii
**Paper Title:** Decoding subject performance from patterns of neural activation in a complex virtual environment
**Authors:** Andrew Floren, David Ress, Bruce Naylor, Risto Miikkulainen
Acknowledged Federal Support: **Y**

**Publication Type:** Conference Paper or Presentation **Publication Status:** 1-Published
**Conference Name:** Neuroscience 2015
Date Received: 14-Jun-2016 Conference Date: 21-Oct-2015 Date Published: 17-Oct-2015
Conference Location: Chicago, IL
**Paper Title:** Decoding cognitive states with a hidden Markov model
**Authors:** Andrew Floren, Bruce Naylor, Risto Miikkulainen, and David Ress
Acknowledged Federal Support: **Y**

**Publication Type:** Conference Paper or Presentation **Publication Status:** 1-Published
**Conference Name:** Neuroscience 2014
Date Received: 07-Sep-2016 Conference Date: 15-Nov-2014 Date Published: 15-Nov-2014
Conference Location: Washington, DC
**Paper Title:** Decoding external and internal cognitive variables during natural tasks in realistic virtual environments
**Authors:** Andrew Floren , Bruce Naylor, Risto Miikkulainen, David Ress
Acknowledged Federal Support: **Y**

# Neurometrics

***Executive Summary***


The goal of this research project was to push forward the known and demonstrated capacity for using Virtual Environments as the primary stimulus for neuroscience research that utilized functional MRI, with emphasis on its utility for training and support of our military.

1) We created a virtual environment that simulated a town in the Middle East. Subjects viewed scenes within the town featuring various numbers of virtual combatants. We evaluated using various forms of machine learning to ascertain the number of combatants in each scene, which was a complex visual stimulus, from a set of fMRI data collected while a subject was watching of the stimulus. We developed and evaluated various schemes to optimize accuracy of the machine learning to identify how many combatants. 1-6, were being shown based solely on the fMRI data. We invented a novel form of sensitivity analysis to relate the machine learning results back to brain anatomy. This work was published in Frontiers of Neuroscience.

2) We used a similar virtual environment to evaluate task difficulty by measuring an internal cognitive variable. In this experiment, subjects had to perform a very simple task, 'weapon detection", while viewing a cluttered and complex visual scene. Task difficulty was varied randomly from scene-to-scene. Support vector methods were developed and optimized to predict task difficulty from fMRI data measured during the task. We obtained excellent results using support-vector regression, which offered 80-90% accuracy. These results motivated additional subsequent funding from the Veterans Administration to use similar methods as a means to treat PTSD. A manuscript for this work is in preparation for submission to Neuroimage.

3) For a third experiment, we developed a third virtual environment and task to evaluate the use of machine-learning of models of "cognitive state-space". The task featured three states taken from video games: *navigation*, *search*, and *evasion*. While some fMRI data was obtained, we ran out of funding before we were able to complete this phase of the project. Nonetheless, preliminary results were encouraging.

Below is a summary of what was learned from the first experiment, followed by a draft manuscript prepared for the second experiment. The third experiment was executed, but has not yet been written up as part of graduate student's Andrew Floren's Ph.D. dissertation

## Experiment 1 Summary

### Methods and Results

The overarching goal of this project was to push the frontier in the use of 3D Virtual Environments as stimuli for functional MRI based research. There were two related motivations for this. Firstly, we wanted to enable understanding brain functioning under somewhat more realistic circumstances. Such knowledge would compliment the standard approach of highly controlled and more artificial stimuli that is commonly used in fMRI research. For example, most stimuli are often static images, whereas we maximized motion, both of the virtual camera and the virtual characters within a virtual environment to better mimic natural experiences. Secondly, we wanted to lay the groundwork for individualized therapies and training, and so we took an approach more characteristic of clinical research. In particular, we developed computational methods targeting individuals separately, as opposed to aggregating data from groups of subjects to test hypotheses about human brains in general, as is most often the case.

A central component of the methods we developed for this entailed utilizing machine learning for the analysis. This was in contrast to the most common analysis technique at the time, viz. the General Linear Model (GLM), which treats every voxel independently to establish the degree to which each voxel responds to the stimulus. Instead, machine learning finds patterns of response within collections of voxels that can be distributed throughout the brain (often called Multi-Voxel Pattern Analysis). Indeed, the progression of fMRI research over the last decade has made increasingly clear that the brain's large scale functional organization entails many distributed voxels activating in a highly coordinated manner. Machine learning is an important tool for discovering this synchronized response. Our applying machine learning to virtual worlds induced brain response patterns constituted another pioneering aspect of our work.

Machine learning algorithms require "training" in order for them to "learn" what it is we want them to "know". This often is used to create what are called "classifiers" that identify which one of a small set of classes (e.g. houses, faces, etc.) the current input most likely belongs to. We used both Support Vector Machines and Artificial Neural Networks for this. The test for success is the accuracy with which the machine learning can classify a new brain pattern that is a response to stimuli from the same set of classes used for training but has not actually been used in the training. We tested our methods with two similar but distinct experiments in which an overriding concern was achieving as high an accuracy as possible in recognizing the right class.

The first experiment used passive viewing. It had as it's setting a virtual version of a small Middle Eastern village. We wanted to see how well the classifiers could determine how many characters were present in the stimulus at any given time; i.e. could the machine learning alone identify the number of characters in the stimulus that caused the measured brain activation pattern. We used virtual models of both U.S. soldiers as well as insurgents. Their number varied from 1-6, and we trained the classifiers to distinguish between only the number of characters. The camera would travel throughout the town until a group of combatants was encountered. The camera would remain there, while still moving slightly, for 12 seconds, and then move on to another location over a 4 sec. interval. At each location, the composition of the set of characters varied.

This may seem like a simple enough objective, but given the poor signal to noise ratio of fMRI, it was actually quite challenging. We spent well over a year trying to understand what techniques could be utilized to get the highest classification accuracy (also called prediction accuracy). What we arrived at was a multi-stage processing pipeline. The key components of this pipeline were 1) dimensionality reduction using ANOVA, a single voxel statistical correlation test 2) multi-voxel pattern analysis using machine learning 3) an output averaging technique that we showed was superior to standard input averaging. In the end, the prediction accuracy varied quite a bit between different subjects, of which there were 8, and between sessions, of which there were two per subject. The highest accuracy was > 90% where are the lowest was >60%, after dropping two sessions that a quantitative measure of the quality of the data, based on cross-validation, indicated were of poor quality.

As a final step in the processing, we introduced a novel approach to using artificial neural networks (NN) to create "spots on the brain" pictures. Such pictures have been the standard approach to interpreting the analysis provided by GLM. But nothing analogous had been developed when using NN's. For this we used "sensitivity analysis", which is a general technique that can be applied to many modeling problems, include NN's. In our case, we in effect inverted the NN (which cannot be done analytically) to numerically determine for each voxel the degree to which that voxel contributed to discrimination between the set of classes (how many combatants were being seen). This provided a new and different way to analyze brain activity using brain maps: one that reflected how the voxels collectively participated in "solving" the classification problem.

The second experiment was designed to move beyond classification based on obvious sensory properties of the stimulus, such as the number of combatants. Instead, the objective was to measure a more complex internal state. For this, we devised a stimulus based on *task difficulty*; a classic type of experiment, and unlike the first experiment, one that relied on user input. Once again, we used a virtual town, although one considerably more complex than that of the first stimulus. We used the same set of virtual combatants, but now they were controlled by a simple type of Artificial Intelligence often used for simulating simple group behaviors in video games. In our case, we had 20 combatants milling around an intersection of the town. At unpredictable moments, one would pull out a weapon, and hold it at chest height for 3 seconds. The task for the subject was to press a button when they saw a weapon. The difficulty was determined by where in the crowd the combatant with the weapon was located. The closer the weapon was to the virtual location of the subject, the easier the task, and the further away, the more difficult the task. The reason for this was simple: the greater the distance from the subject, 1) the smaller the size of the image of the character due to perspective projection 2) the greater the occlusion of the weapon and the character pulling out the weapon due to closer characters 3) the more the motion of other closer characters distracted the subject from noticing the weapon being pulled out.

For purposes of training the machine learning, we divided the distance into three classes: near, medium and far.

Compared to the first experiment, we expected decoding task difficulty to be significantly more complex and would therefore require an improved processing pipeline to maximize classification accuracy. We had already found that reducing the dimensionality of the data was a key processing step for maximizing classification accuracy, but the ANOVA correlation test we used in the first experiment requires an independent measure of task activation which in turn requires the acquisition of a large amount of control data. This control data cannot easily be acquired for all experimental designs and is unsuitable for training the classifier. We therefore developed a general technique that exploits the structure of the brain to reduce the dimensionality of the data while at the same time improving the signal to noise ratio. First, cortical surface meshes are constructed using high resolution anatomical scans, and the meshes for each hemisphere are aligned to a spherical template. Then, functional data from each session is projected onto this surface after alignment with the subject's anatomical scan. This results in an immediate reduction in dimension as only data on the cortical surface is retained. The data is then smoothed along the surface and resampled at a lower resolution. Smoothing along the surface improves signal to noise with minimal introduction of error compared to smoothing within the volume; in which data is smoothed across both gyral and sulcal boundaries, while resampling is at a lower resolution further and reduces the dimensionality of the data even further with minimal loss of information. We explored a wide parameter space of smoothing parameters, sampling rates, and ANOVA dimensionality reduction to determine the optimal settings for maximizing classification accuracy.

We were successfully able to decode the subjects performance at a task between the three states poor, average, and good with accuracy varying from 70% to 91% (chance decoding accuracy was 33%) depending on the particular session. We were only able to achieve this high decoding accuracy by combining the cognitive state-space filtering techniques developed in our first experiment with the improved preprocessing techniques developed in this experiment. While the ANOVA dimensionality reduction was still useful, we found that the structural dimensionality reduction technique was more than sufficient to achieve strong classification accuracy. In future experiments, this means more of each scanning session can be devoted to collecting data useful for training the classifier instead of control data. Furthermore, the spherical template alignment simplified aggregating training data across both sessions and subjects which yielded yet further improved classification accuracy.

# [1]Decoding behavioral performance accurately from fMRI data obtained in a realistic virtual environment


## Abstract

**Commented [AF1]:** Finish abstract

Abstract text


## Introduction

Decoding cognitive-state variables (Akama, Murphy, Na, Shimizu, & Poesio, 2012; Floren, Naylor, Miikkulainen, & Ress, 2015; Spiers & Maguire, 2007) from fMRI opens up exciting new applications in training and therapy. However, many hurdles remain to be cleared before real-world applications are possible: training or therapy must be performed within the scanner, decoding techniques must be able to handle less structured and controlled stimuli to accommodate training and therapy programs, we must be able to decode across different sessions to reduce total time in the scanner, and we must be able to decode variables that are inherent to the subject rather than the stimuli.

The requirement that the subject must remain almost completely motionless over the course of the session severely limits the types of programs where fMRI decoding is realistically applicable. A potential solution to this limitation is to use virtual reality environments for the stimuli. While motion and other somatosensory inputs are still restricted, virtual environments provide a more immersive experience that should cause the subject's neural response to be closer to that evoked by the real world. Subjects can still attain the experience of controlling motion using various input devices as is done in video games. Encouragingly, some therapies already utilize virtual reality such as PTSD extinction therapy (Gonçalves, Pedrozo, Coutinho, Figueira, & Ventura, 2012).

Most cognitive neuroscience experiments that have used machine learning focused on explanatory power at the expense of classification accuracy and stimuli complexity. For example, in the work of (Goesaert & Op de Beeck, 2013) faces are presented to the viewer and an SVM classifier is trained to determine small differences in the presented faces. The faces are highly controlled for contrast and small variations in the structure of the face. This is ideal for interrogating the neural representation of the perception of faces, but not a strong indicator of achievable classification accuracy in a more realistic setting. Conversely, in (Kauppi et al., 2011) faces are presented in a realistic stimulus, but classification is performed using ordinary least squares regression. The statistical properties of the regression make hypothesis testing easier, but decoding accuracy suffers. These applications of

---

1

machine learning make sense when the objective is to answer a basic research question. However, our goal is to apply fMRI cognitive state decoding during training and therapy, where classification accuracy is extremely important, and realistic stimuli are essential. For such applications, we have investigated the use of more complex multivariate decoding techniques including feed-forward neural-networks and time-dependent filtering techniques. While these are more difficult to interpret using the usual linear correlations between stimuli and response, they provide significantly improved decoding accuracy.

Unfortunately, fMRI data is relatively ill-posed for advanced machine learning methods. The dimensionality of whole brain data is large compared to the number of samples that can be collected during a typical hour-long scanning session (e.g., 256,000 voxels vs. ~ 1000 time slices). Furthermore, because of the hemodynamic character of fMRI, the samples are correlated in both space and time (Buxton, Uludağ, Dubowitz, & Liu, 2004; Glover, 1999). Many data-agnostic approaches have been used to attempt to solve this problem such as principal component analysis (PCA; Hotelling, 1933) and independent component analysis (ICA; Comon, 1994). However, we want to use the structure of fMRI data and more specifically the cortical structure of the brain to reduce the dimensionality of the data intelligently.

For training and therapy applications, it will be helpful to be able to train the classifier in one session, and then use that trained classifier in subsequent sessions. This will reduce the necessary scanning time per subject because the classifier will only need to be trained once rather than every session. It also opens up the possibility of a decoding algorithm that gets better with each successive scanning session. Ideally, we would also like to be able to user decoders trained on a large number of different subjects. The main difficulty to solving both of these problems is accurately registering the volumes between different sessions and subjects. We have found that spherical registration (Yeo et al., 2008) yields the best results for both cross-session and cross-subject registration in terms of decoding accuracy.

For the desired applications, it is important to decode cognitive variables that are inherent to the subject rather than a reflection of the stimuli. We already know what stimuli we are presenting; we want to learn more about the subject. In this paper, we decode the subject's performance at a difficult task. While we control the difficulty of the task which is correlated with their performance, we are really interested in how their attention and strategy for completing the task influence performance. Furthermore, predicting a subject's performance has obvious applications for training and therapy such as modulating difficulty to keep predicted performance on a specific trajectory.

In this manuscript, we present a unified framework for training and evaluating machine learning models across sessions and subjects while leveraging the underlying structure of the brain and vasculature to minimize the dimensionality of the data. We show that within this framework it is possible to predict, with significant accuracy, subject performance at a complex task within a virtual environment (similar to that of existing PTSD extinction therapy treatments). And we present sensitivity maps of the trained models which show that the drivers of these predictions, while correlated with functional regions, utilize diverse networks across the brain and vary significantly between individuals.

## Methods

### Subjects

Six adult males, ages 24-57, with normal or corrected-to-normal vision, participated in the experiments. All subjects participated in two fMRI sessions and a third session to acquire a high-resolution structural anatomy. Informed consent was obtained from all subjects under a protocol approved by the University of Texas at Austin Institutional Review Board.

### Stimulus

Given our long-term interest in PTSD, we created a virtual environment very similar to those used in current PTSD extinction therapy programs (Figure 1. Virtual reality environment and search task.). Subjects perform a weapon-detection task at six different locations in this virtual town. Between detection tasks, the screen smoothly fades from the current location to a new location over five seconds. Next, a cue is presented to the subject that indicates the difficulty of the upcoming block for five seconds. Then the subject performs the task at that location for 45 seconds. During the task, 20 characters move about within the subject's view. A randomly chosen character will pull out a weapon every two to four seconds. Subjects press a button to indicate their detection of the weapon which was previously not visible. In response to their button press, the character puts the weapon away, and the trial is flagged as successful. If after a certain amount of time the subject fails to detect the weapon, the character puts the weapon away, and the task is flagged as unsuccessful. After the task, we have a control period in which weapons are replaced by non-weapon objects. The control period begins with another 5-second cue followed by a 15-second period during which the characters move in the same fashion, but pull out flashlights rather than weapons. The subject is instructed to gaze at the horizon during this period, and to not respond with button presses. This process is repeated at six different locations within the virtual town, and the entire stimuli is repeated six times during each fMRI scan session.

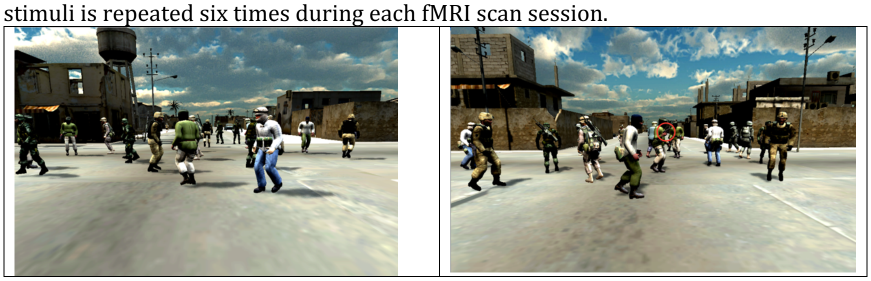


*Figure 1. Virtual reality environment and search task. A) An example frame from an intersection in the virtual reality environment. B) A frame in which a weapon is visible circled in red. C) A frame during the cue presentation. D) A frame during the control period when a flashlight is visible circled in red.*

Task difficulty was carefully adjusted for each of these periods to vary between easy, medium, and hard. Task difficulty was primarily controlled by varying the distance between the character selected to hold the weapon and the camera. Characters far away from the camera where both visually smaller and more likely to be partially or fully occluded by other characters. However, the actual projected size of the weapon and level of occlusion could vary significantly for the same target depth due the random nature of the characters' movement. To account for this variability in difficulty, the duration that the weapon is held by the character is modeled in real-time to achieve a roughly constant quantity of visibility. We calculate the total area of visible weapon in pixels and multiply this by the frame rate of the stimuli to arrive at the weapon's total visibility in pixel-seconds. The character holds the weapon until the accumulated pixel-seconds exceed the threshold for the current difficulty setting. To keep the expected duration of presentation constant across different distances equal, the pixel-second threshold was divided by the square of the distance between the viewer and the character.

During training periods outside of the scanner, the subject received audio feedback for successfully finding the weapon, missing a weapon, and indicating that they have seen a weapon while one is not present. However, during the scanning sessions there was no performance feedback.

The expected difficulty of the stimuli is adjusted each time the subject moves to a new location. However, the difficulty settings and locations were balanced so that there would be no correlation between location and task difficulty. Furthermore, low-level contrast was held constant in real-time using a custom GPU shader.

## Psychophysics

Performance data was collected for each subject outside of the scanner to collect performance data and to ensure that the subjects understood and correctly responded to the stimuli cues. Task performance was estimated as the average

performance of the subject during a 30-second block while the distance of the weapon from the viewer and its total visibility were used to estimate difficulty. However, due to the semi-random movements of the characters, effective distance and visibility could vary considerably. Therefore, we collected extensive logging information during the stimuli that included the exact timing, positions, and visibility of all characters and weapons as well as the responses of the subject.
In an early prototype of the game, weapons were presented for fixed durations at specified distances to control task difficulty (Fig. #a). However, we found that the correlation between distance and performance was much better when the total visibility (the number of pixels containing the weapon in each frame multiplied by the duration of the frame summed over all frames that the weapon was presented) was held constant rather than the duration (Fig. #b). This helped control the difficulty of the task when the weapon was occluded due to the random motion of the characters. For a particular expected duration, the total visibility is tightly correlated with distance, that is, the closer the weapon is to the screen the more pixel-seconds it accrues every frame. To limit this interaction, the total visibility for a particular presentation was linearly related to the distance rather than being a constant value. We found that a slope of 1000 and an intercept of -65 gave a reasonable duration for the presentation while still creating a strong performance curve with distance for all subjects. Note that this relationship is dependent on the resolution of the screen displaying the stimuli. Expected performance versus weapon distance curves show similar shapes for each subject, but there is clear variability in task performance between subjects.

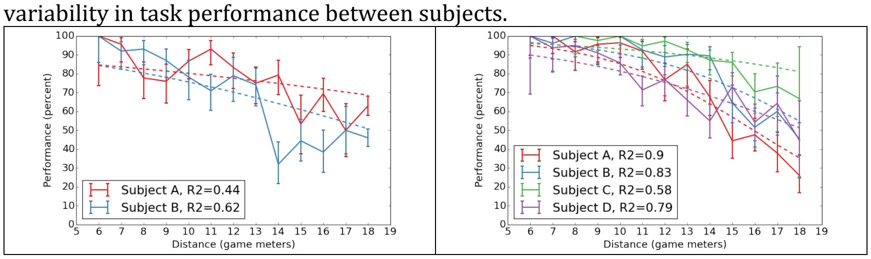



*Figure 2. Psychophysics results*

*We fit a logistic regression to the subject's performance given the in-game distance of the presented character with fixed duration (a) and fixed total visibility (b). The $R^2$ value of the logistic regression predictions and the actual subject performance for each difficulty setting are presented in the figures.*

## MRI Protocols

Imaging was performed on a Siemens Skyra 3T scanner using the product 32-channel head coil. Structural reference volumes were T1-weighted with good gray-white contrast and acquired using a MP-RAGE sequence (minimum TE, TR = 2600 ms, TI = 900 ms, 9° flip angle, isometric voxel size of 0.7 mm, 2 excitations, ~28-min duration). fMRI scans were collected using a whole-brain GRAPPA EPI sequence with g-factor = 4, TE = 25 ms, TR = 2 s, and 2-mm cubic voxels across a 200-mm field-of-view. The slice prescription included 60 slices oriented along the AC-PC axis. A high-order shim was performed before the start of the functional imaging to improve field homogeneity. A set of T1-weighted structural images was obtained on the same prescription as the functional acquisition runs in the same session directly before the functional scans were collected using a three-dimensional (3D) FLASH sequence (minimum TE and TR, ~1-mm pixel size, 15° flip angle). These anatomical images had good gray-white contrast and were used to align the functional data to the structural reference volume.

### Preprocessing

Automatic cortical segmentation and surface extraction was performed on the structural reference volume using FreeSurfer (ref). The cortical surfaces for each hemisphere were inflated into a sphere while minimizing metric distortion. These spherical surfaces are then registered to FreeSurfer's spherical atlas first by coarsely aligning on large-scale folding patterns and then fine-tuned using small-scale curvature patterns (freesurfer sphere ref).
The inplane anatomical volumes were skull-stripped and normalized in the same fashion as the first stage of the automatic cortical segmentation and surface extraction. The processed inplane volumes were then affinely registered to the structural reference volumes using a method based on robust statistics to detect outliers and remove them from the registration (ref: Highly Accurate Inverse Consistent Registration: A Robust Approach).
Simultaneous slice-timing and motion correction was performed on the functional scans (ref nipy.SpaceTimeRealign). Then, a rigid-body registration was performed between scans to align each frame to the first volume, i.e., the frame closest in time to the structural inplane (ref Improved Optimisation for the Robust and Accurate Linear Registration and Motion Correction of Brain Images).
The functional data was then approximately aligned to the structural reference volume using the previously calculated registration with the inplane anatomical data. A boundary based registration technique was then used to fine-tune the registration of the functional data to the structural reference volume (freesurfer bbregister).
In the next step, the functional data was projected onto the extracted cortical surfaces of the left and right hemispheres by averaging between white and pial surfaces along the surface normal (mri_vol2surf). To minimize partial volume effects, values were only averaged between 20% and 80% of the distance between the white and pial surfaces along the normal.
The functional data was smoothed along the surface with a Gaussian filter and then projected on to an icosahedron with uniform spacing of vertices in the spherical template space (freesurfer mri_surf2surf). We experimented with several different smoothing values and icosahedron order numbers (which determines the density of

vertices on the sphere, i.e., the resolution of the data) to determine parameters that sufficiently reduce the dimensionality of the data while still retaining as much information as possible. The number of vertices on the template sphere is:

$$v = 10 \cdot (2^N)^2 + 2$$

These vertices are approximately evenly spaced along the entire surface of the sphere (Yeo et al., 2008). Note that it is impossible subdivide a sphere with perfectly even spacing for more than 20 vertices (Tegmark, 1996).

Finally, linear detrending was applied independently to the time series at each vertex for all scans to mitigate the effects of scanner drift (Tanabe, Miller, Tregellas, Freedman, & Meyer, 2002). That is, a line was fit to the functional data at each vertex for each scan. This line is then subtracted from the data used for the fit. The resulting residuals are then used as the detrended functional data in subsequent classification and regression stages.

## Feature Selection

The number of remaining dimensions – i.e., vertices on the template sphere – is dependent on the icosahedron order number. For order numbers greater than four, the dimensionality of the data is still too high to get reasonable decoding accuracy. Based on our previous work [ref], we used an ANOVA univariate feature-selection (ANOVA ref and sk-learn ref) method to reduce the dimensionality down to 3000. These voxels are selected by comparing the distribution of the data during task periods and control periods. The 3000 voxels whose distributions show the most significant mean difference between these periods are selected. For smaller icosahedron order numbers, this feature-selection step is not necessary as the number of input dimensions is already on the order of 3000 or smaller. Note that feature-selection is not based on the subject performance labels, but rather an independent label that indicates whether the subject is performing the task or not.

## Classification and Regression

Classification of subject performance was performed using a linear support-vector machine (SVM) (SVM ref) with $C = 1$. Continuous performance values were binned into three equal categories: poor, average, and good. Thresholds for labeling were chosen such that the number of samples in each category were equal. Specifically, the worst third of blocks were labeled poor, the middle third were labeled average, and the top third were labeled good. Block filtering (neurometrics1 ref) was performed on the output of the SVM.

Regression of block averaged subject performance was performed using a support-vector regression (SVR) (SVR ref) with $C = 1$ and the rotational basis function as the kernel. No smoothing was applied to the regression. Classifier and regression performance metrics were both estimated using a six-fold cross-validation procedure (cv ref). The folds were selected based on the six scanning runs during the fMRI sessions. This minimizes the confounding effects of temporal correlation between training and test examples that can result in overly optimistic accuracy estimates [ref].

## Performance Analysis

To confirm the efficacy of our preprocessing methods as well as determine the cumulative impact of each step, we estimated decoding accuracy within each session after projecting onto the surface, smoothing along the surface, and finally down-sampling and projecting onto the template sphere. We also estimated decoding accuracy on the full voxel data with only motion and slice-timing correction as well as with simple volumetric smoothing for comparison purposes.

We also estimated decoding accuracy for different smoothing kernel parameters and icosahedron order numbers. Larger icosahedron order numbers correspond to higher effective sampling rates. We calculate the effective sampling density as the surface area of the brain divided by the total number of vertices on the icosahedron. We performed a gridded search for these parameters on the range of 2—8-mm FWHM smoothing, and icosahedron order numbers 2—8. Unfortunately, one cannot select arbitrary sampling densities but are constrained to integer icosahedron order numbers.

We know that the problem is severely under-determined even for low sampling densities, that is, we still have many more dimensions than input examples. To better understand how this affects our decoding accuracy we performed a learning curve analysis [ref], where we artificially reduced our training examples and estimated decoding performance on the reduced set, plotting decoding performance versus the number of training examples. Samples were reduced based on scanning runs, that is, performance was estimated using the first run, then the first and second runs combined, and so on. This limits potential issues with temporal correlations in the training and test sets. We then combined data across sessions for each subject to determine if this could improve decoding accuracy and repeated the learning curve analysis.

## Sensitivity Analysis

We performed sensitivity mapping [ref] of the trained SVR models for each subject. Sensitivity analysis is a technique to determine the impact a perturbation on an input variable would have on the output of a classifier or regressor. It is calculated very directly as the average matrix of all first-order partial derivatives of the trained classifier or regressor with respect to the input variables for all training examples. For a regressor with a single output, this reduces to a vector rather than a matrix. We then project these partial derivatives onto slightly inflated cortical surfaces for display and analysis. It is interesting to note that for a simple linear regression, this technique corresponds to traditional beta maps, but it is applicable to all differentiable models regardless of linearity.

We only present the SVR maps as they are significantly easier to interpret. The problem with sensitivity mapping with classifiers is that the result is a vector field. In our previous work [ref], we reported the magnitude of this vector which indicates how strongly a region of the brain contributes to classification but does not indicate how it contributes to that classification. For a regression, the calculation results in a

single value for each vertex/voxel which can be directly mapped onto the cortical surface.
For a trained SVR model with $L$ support vectors, dual coefficients $w_i$, and a kernel function $k(\cdot,\cdot)$, the derivative with respect to an example vector $\boldsymbol{x}$ can be derived as follows.

$$f(\boldsymbol{x}) = \sum_{i=1}^{L} w_i \cdot k(\boldsymbol{x}_i, \boldsymbol{x}) + b, \qquad w_i = \alpha_i - \alpha_i^*$$

$$\frac{\delta f(\boldsymbol{x})}{\delta \boldsymbol{x}} = \sum_{i=1}^{L} w_i \cdot \frac{\delta k(\boldsymbol{x}_i, \boldsymbol{x})}{\delta \boldsymbol{x}}$$

For the rotational basis function kernel:

$$k(\boldsymbol{x}_i, \boldsymbol{x}) = \exp(-\gamma \|\boldsymbol{x}_i - \boldsymbol{x}\|^2)$$

$$\frac{\delta k(\boldsymbol{x}_i, \boldsymbol{x})}{\delta \boldsymbol{x}} = \exp(-\gamma \|\boldsymbol{x}_i - \boldsymbol{x}\|^2) \cdot 2\gamma \|\boldsymbol{x}_i - \boldsymbol{x}\|$$

$$\frac{\delta f(\boldsymbol{x})}{\delta \boldsymbol{x}} = \sum_{i=1}^{L} w_i \cdot \exp(-\gamma \|\boldsymbol{x}_i - \boldsymbol{x}\|^2) \cdot 2\gamma(\boldsymbol{x}_i - \boldsymbol{x})$$

This result can be interpreted as a weighted sum of differences between the support vectors and the example vector $\boldsymbol{x}$.

## Results

We evaluated the average accuracy improvement across sessions after each stage in our preprocessing pipeline. There is significant variability in performance between subjects and sessions, but the relative improvement was much more consistent. Volumetric smoothing is not normally used in our pipeline, but was tested for comparison purposes because it is the most common approach to improving signal to noise in fMRI data. Here we can see that it does in fact improve decoding accuracy but much less than surface-based approaches. The surface projection, surface smoothing, and spherical down-sampling steps are all cumulative, that is, surface smoothing included surface projection, and spherical down-sampling includes both surface projection and smoothing.

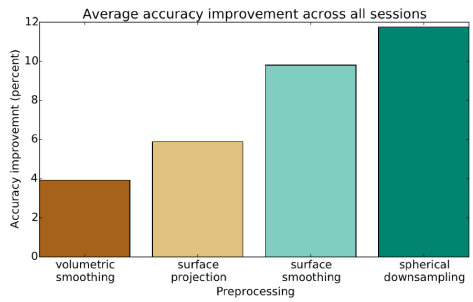



*Figure 3. Accuracy improvement at each stage of preprocessing averaged across all sessions.*

We plotted average decoding accuracy across all sessions for different surface smoothing and spherical down-sampling parameters (Fig. #). Best accuracy occurs at a FWHM of 2mm for the surface smoothing kernel and an icosahedron order of 4, which corresponds to a sampling density of ~5.1 samples/mm. Additionally, as the size of the smoothing kernel increases the optimal corresponding sampling density decreases.

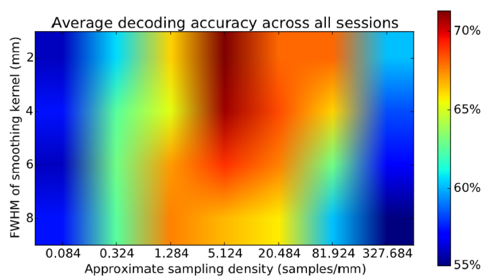



*Figure 4. Decoding accuracy averaged across all sessions for different surface smoothing and down-sampling parameters.*

Learning curve analysis allows us to plot decoding accuracy against varying numbers of training examples by artificially reducing the total number of examples. From the single session learning curves (Fig. ?) we can see a strong correlation between number of training examples and decoding accuracy. For the sessions with the highest decoding accuracy, the performance does appear to approach a asymptote, but for the less accurate sessions we expect that more training examples would continue to improve decoding accuracy.
By combining data across sessions for individual subjects we have more effective training examples available. Learning curve analysis on this data shows that the conclusion suggested by the single session learning curves is correct: more training

examples does in fact lead to greater decoding accuracy even for the subject whose individual sessions did not appear to be approaching an asymptote.

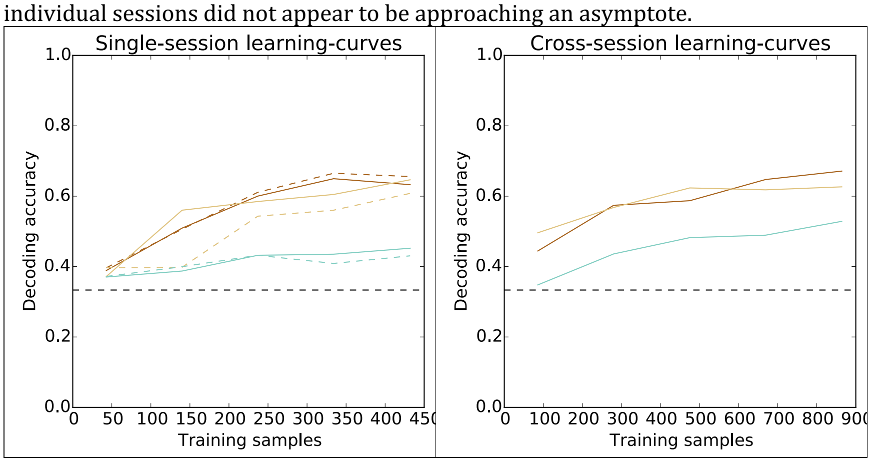



*Figure 5. A) Within-session learning curve analysis. Each plot represents the learning curve for a different session. B) Across-session learning curve analysis. Each plot represents the learning curve for a different subject.*

Regression analysis was performed at the same optimal smoothing value of 2mm full-width half-max and icosahedron order of 4 on the combined data for all subjects. The average $R^2$ value was {average r2 value} and the average RMSE was {average rmse value} across subjects. An example time series is depicted in Figure 6.

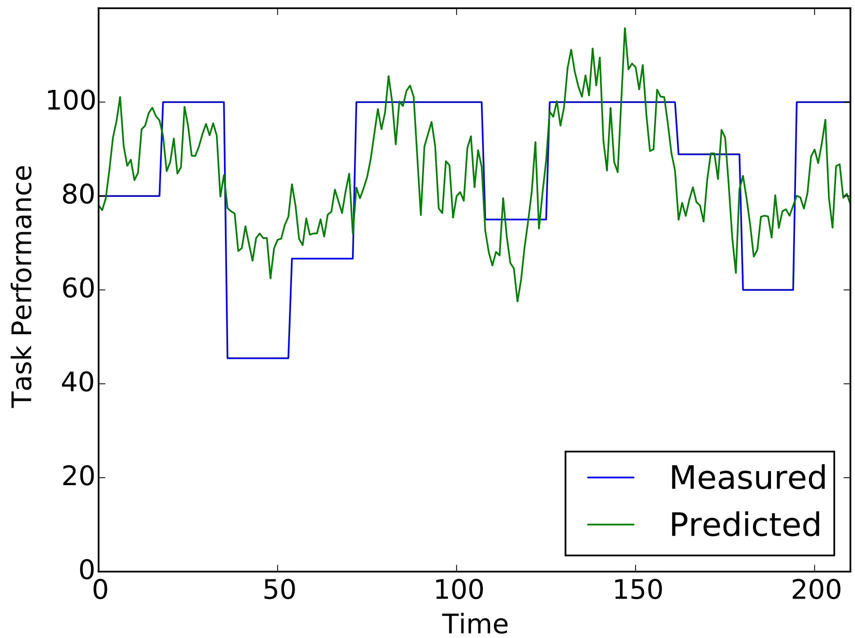
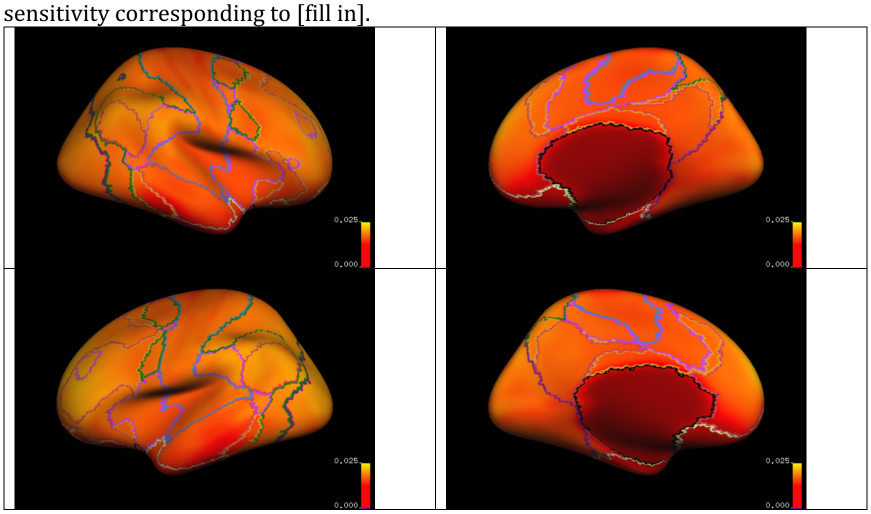



*Figure 6*

Sensitivity mapping in individual subjects does show some variability between subjects, but broadly speaking they cover similar brain regions. We see areas of sensitivity corresponding to [fill in].



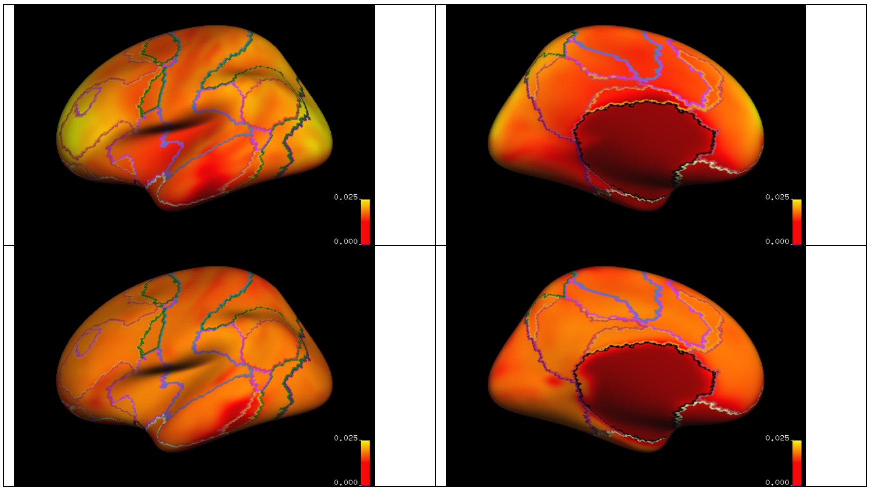



*Figure 7. Individual sensitivity maps for decoding subject performance*

We see this same pattern repeated again in the sensitivity map constructed on the combined cross-subject data.

*Figure 8. Combined sensitivity map for decoding subject performance*

## Discussion

We created our experiment with the goal of decoding subject performance from fMRI data in a realistic environment and developing methods to decode as accurately as possible. The stimulus was designed to be visually complex and realistic while at the same time minimize low-level visual correlations with task difficulty. The careful control of low-level visual correlations makes the task of decoding more difficult, but at the same time more meaningful. Our results show that not only is it possible to decode subject performance over time, but that it can be decoded with very high accuracy with the right preprocessing stages (put specific numbers here).

Decoding internal or non-sensory variables is significantly more difficult than external or sensory variables. First, the ground truth is easy to measure for sensory variables, since we are in control of the stimuli as presented to the subject. On the other hand, for an internal variable we must have some other means of measuring the response to determine ground truth. In our case, we measure subject responses over time in order to estimate their performance. Second, sensory information is easier to decode because it has a spatially large and over-complete representation in early processing areas which makes it easier to measure with the resolutions available to fMRI. The neural representation is large and over-complete because the brain has not yet decorrelated the sensory information in order to more efficiently encode it in the brain.

Decoding external variables helps us understand how the brain encodes sensory information, but decoding internal variables can give us a window into how a subject is behaving or feeling. This window can have many applications from new cognitive experimental designs to training and therapy programs. For example, decoding subject performance could be used to adjust the difficulty of a task in real-time to maintain a specific performance level. In our experiment, we had to keep task difficulty constant over long blocks to estimate performance and establish ground truth, but with a trained and accurate decoder this restriction could be lifted.

Why realism is important

The results show that a preprocessing pipeline that uses the structure of the brain to reduce noise and dimensionality of the data significantly improves decoding accuracy over standard volumetric smoothing. Projecting the data onto the cortical surface reduces the dimensionality by ignoring spurious information not located in the cortex. The surface is also a more natural representation of cortical function where distance along the surface is more likely to reflect functional organization than distance in the volume. Smoothing on the surface avoids introducing noise by not blurring surface vasculature and CSF signals into the data, as well as avoiding blurring data across sulci and gyri. Down-sampling is a natural next step after smoothing to reduce the dimensionality of the data with minimal information loss if the smoothing and down-sampling parameters are chosen appropriately. The various elements of the pipeline are not new to the neuroimaging community but they are not yet standard practice. The pipeline is largely automatic and robust and

**Commented [RD2]:** These statements are interesting, but border on being speculative. They need to be much better supported from the literature if they are to remain here.

**Commented [RD3]:** How?

**Commented [RD4]:** Why is this better than just self-report?

the only parameters to be selected are the smoothing kernel size and sampling density. We hope that by making our automated pipeline freely available (link) more researchers will consider using these excellent tools developed by the neuroimaging community. Our experiments indicate good values to select for these parameters, though we expect the optimal parameters are dependent on the scanning protocol. For example, if the scanning resolution is reduced then the optimal smoothing kernel is likely larger.

Another advantage of the pipeline is that resulting data is embedded in FreeSurfer's spherical template space (ref). This surface based representation is not only a more natural fit to the functional organization of the cortex, but it is also a convenient way to accurately and automatically co-register data across both sessions and subjects. Although cross-subject decoding accuracy was still below within-subject decoding accuracy, the cross-subject decoding accuracy when using spherical registration was significantly better than when using non-linear volumetric registration to Talairach coordinates – the most common template space for combining data from multiple subjects. At the same time, the data takes up significantly less hard drive space in the spherical coordinates without losing any information relevant for decoding which makes it a better choice for large databases of fMRI data.

Learning curves imply that more data is better

Can be used to determine appropriate number of samples

Indicates that registering across sessions is very important

Sensitivity maps indicate … That is consistent with …

## Bibliography


Akama, H., Murphy, B., Na, L., Shimizu, Y., & Poesio, M. (2012). Decoding semantics across fMRI sessions with different stimulus modalities: a practical MVPA study. *Frontiers in Neuroinformatics*, *6*(August), 24. https://doi.org/10.3389/fninf.2012.00024

Buxton, R. B., Uludağ, K., Dubowitz, D. J., & Liu, T. T. (2004). Modeling the hemodynamic response to brain activation. *NeuroImage*, *23 Suppl 1*, S220-33. https://doi.org/10.1016/j.neuroimage.2004.07.013

Comon, P. (1994). Independent component analysis, A new concept? *Signal Processing*, *36*(3), 287–314. https://doi.org/https://doi.org/10.1016/0165-1684(94)90029-9

Floren, A., Naylor, B., Miikkulainen, R., & Ress, D. (2015). Accurately decoding visual information from fMRI data obtained in a realistic virtual environment. *Frontiers in Human Neuroscience*, *9*(June), 1–13. https://doi.org/10.3389/fnhum.2015.00327

Glover, G. H. (1999). Deconvolution of impulse response in event-related BOLD fMRI. *NeuroImage*, *9*(4), 416–29. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10191170

Goesaert, E., & Op de Beeck, H. P. (2013). Representations of facial identity information in the ventral visual stream investigated with multivoxel pattern

analyses. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *33*(19), 8549–58. https://doi.org/10.1523/JNEUROSCI.1829-12.2013
Gonçalves, R., Pedrozo, A. L., Coutinho, E. S. F., Figueira, I., & Ventura, P. (2012). Efficacy of virtual reality exposure therapy in the treatment of PTSD: a systematic review. *PloS One*, *7*(12), e48469. https://doi.org/10.1371/journal.pone.0048469
Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, *24*(6), 417–441. https://doi.org/10.1037/h0071325
Kauppi, J. P., Huttunen, H., Korkala, H., Jääskeläinen, I. P., Sams, M., & Tohka, J. (2011). Face prediction from fMRI data during movie stimulus: Strategies for feature selection. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *6792 LNCS*, 189–196. https://doi.org/10.1007/978-3-642-21738-8_25
Spiers, H. J., & Maguire, E. a. (2007). Decoding human brain activity during real-world experiences. *Trends in Cognitive Sciences*, *11*(8), 356–65. https://doi.org/10.1016/j.tics.2007.06.002
Tanabe, J., Miller, D., Tregellas, J., Freedman, R., & Meyer, F. G. (2002). Comparison of detrending methods for optimal fMRI preprocessing. *NeuroImage*, *15*(4), 902–907. https://doi.org/10.1006/nimg.2002.1053
Tegmark, M. (1996). An icosahedron-based method for pixelizing the celestial sphere. https://doi.org/10.1086/310310
Yeo, B. T. T., Sabuncu, M., Vercauteren, T., Ayache, N., Fischl, B., & Golland, P. (2008). Spherical demons: fast surface registration. *Medical Image Computing and Computer-Assisted Intervention : MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention*, *11*(Pt 1), 745–53. https://doi.org/10.1016/j.jmb.2008.10.054.The