

**JoVE: Science Education**  
**Decoding Auditory Imagery with Multivoxel Pattern Analysis**  
--Manuscript Draft--

<b>Manuscript Number:</b>	10267
<b>Full Title:</b>	Decoding Auditory Imagery with Multivoxel Pattern Analysis
<b>Article Type:</b>	Manuscript
<b>Section/Category:</b>	Manuscript Submission
<b>Corresponding Author:</b>	Jonas Kaplan UNITED STATES
<b>Corresponding Author Secondary Information:</b>	
<b>Corresponding Author's Institution:</b>	
<b>Corresponding Author's Secondary Institution:</b>	
<b>First Author:</b>	Jonas Kaplan
<b>First Author Secondary Information:</b>	
<b>Order of Authors:</b>	Jonas Kaplan
	Sarah Gimbel
<b>Order of Authors Secondary Information:</b>	

**PI:** Jonas T. Kaplan & Sarah I. Gimbel

**Psychology Education Title:** Decoding Auditory Imagery with Multivoxel Pattern Analysis

## Overview

Imagine the sound of a bell ringing. What is happening in the brain when we conjure up a sound like this in the “mind’s ear?” There is growing evidence that the brain uses the same mechanisms for imagination that it uses for perception<sup>1</sup>. For example, when imagining visual images, the visual cortex becomes activated, and when imagining sounds, the auditory cortex is engaged. However, to what extent are these activations of sensory cortices specific to the content of our imaginations?

One technique that can help to answer this question is multivoxel pattern analysis (MPVA), in which functional brain images are analyzed using machine learning techniques<sup>2-3</sup>. In an MPVA experiment, we train a machine learning algorithm to distinguish among the various patterns of activity evoked by different stimuli. For example, we might ask if imagining the sound of a bell produces different patterns of activity in auditory cortex compared with imagining the sound of a chainsaw, or the sound of a violin. If our classifier learns to tell apart the brain activity patterns produced by these three stimuli, then we can conclude that the auditory cortex is activated in a distinct way by each stimulus. One way to think of this kind of experiment is that instead of asking a question simply about the *activity* of a brain region, we ask a question about the *information content* of that region.

In this experiment, based on Meyer et al., 2010<sup>4</sup>, we will cue participants to imagine several sounds by presenting them with silent videos that are likely to evoke auditory imagery. Since we are interested in measuring the subtle patterns evoked by imagination in auditory cortex, it is preferable if the stimuli are presented in complete silence, without interference from the loud noises made by the fMRI scanner. To achieve this, we will use a special kind of functional MRI sequence known as sparse temporal sampling. In this approach, a single fMRI volume is acquired 4-5 seconds after each stimulus, timed to capture the peak of the hemodynamic response.

## Procedure

### 1. Participant Recruitment

#### 1.1. Recruit 20 participants.

- 1.1.1. Participants should be right-handed and have no history of neurological or psychological disorders.
- 1.1.2. Participants should have normal or corrected-to-normal vision to ensure that they will be able to see the visual cues properly.
- 1.1.3. Participants should not have metal in their body. This is an important safety requirement due to the high magnetic field involved in fMRI.

1.1.4. Participants should not suffer from claustrophobia, since the fMRI requires lying in the small space of the scanner bore.

## 2. Pre-scan Procedures

2.1. Fill out pre-scan paperwork.

2.2. When participants come in for their fMRI scan, instruct them to first fill out a metal screen form to make sure they have no counter-indications for MRI, an incidental-findings form giving consent for their scan to be looked at by a radiologist, and a consent form detailing the risks and benefits of the study.

2.3. Prepare participants to go in the scanner by removing all metal from their body, including belts, wallets, phones, hair clips, coins, and all jewelry.

## 3. Provide instructions for the participant.

3.1. Tell the participants that they will see a series of several short videos inside the scanner. This videos will be silent, but they may evoke a sound in their “mind’s ear.” Ask the participant to focus on and encourage this auditory imagery, to try to “hear” the sound as best they can.

3.2. Stress to the participant the importance of keeping their head still throughout the scan.

## 4. Put the participant in the scanner.

4.1. Give the participant ear plugs to protect their ears from the noise of the scanner and ear phones to wear so they can hear the experimenter during the scan, and have them lie down on the bed with their head in the coil.

4.2. Give the participant the emergency squeeze ball and instruct them to squeeze it in case of emergency during the scan.

4.3. Use foam pads to secure the participants head in the coil to avoid excess movement during the scan, and remind the participant that it is very important to stay as still as possible during the scan, as even the smallest movements blur the images.

## 5. Data Collection

5.1. Collect high resolution anatomical scan.

5.2. Begin functional scanning.

5.2.1. Synchronize the start of stimulus presentation with the start of the scanner.

- 5.2.2. To achieve the sparse temporal sampling, set the acquisition time (TR) of an MRI volume to 2 seconds, with a 9 second delay between volume acquisitions.
- 5.2.3. Present the silent videos via a laptop connected to a projector. The participant has a mirror above their eyes, reflecting a screen at the back of the scanner bore.
- 5.2.4. Synchronize the start of each 5-second video clip to begin 4 seconds after the previous MRI acquisition starts. This will ensure that the next MRI volume is acquired 7 seconds after the start of the video clip, to capture the hemodynamic activity that corresponds to the middle of the movie.
- 5.2.5. Present three different silent videos that evoke vivid auditory imagery: a bell swinging back and forth, a chainsaw cutting through a tree, and a person playing a violin.
- 5.2.6. In each functional scan, present each video 10 times, in random order. With each trial lasting 11-seconds, this will result in a scan 330 seconds (5.5 minutes) long.
- 5.2.7. Perform 4 functional scans.

## 6. Data Analysis

### 6.1. Define a Region of Interest

- 6.1.1. Use the high resolution anatomical scan of each participant to trace the voxels that correspond to the early auditory cortex. This corresponds to the surface of the temporal lobe, called the planum temporale. Use the anatomical features of each person's brain to create a mask specific to their auditory cortex.

### 6.2. Preprocess the data.

- 6.2.1. Perform motion correction to reduce motion artifacts.
- 6.2.2. Perform temporal filtering to remove signal drifts.

### 6.3. Train and test the classifier algorithm

- 6.3.1. Divide the data into training and testing sets. Training data will be used to train the classifier, and the left-out testing data will be used to assess what it has learned. To maximize the independence of training and testing data, leave out data from one functional scan as the testing set.

6.3.2. Train a Support Vector Machine algorithm on the labeled training data from auditory cortex in each subject. Test the classifier's ability to correctly guess the identity of the unlabeled testing set, and record the classifier's accuracy.

6.3.3. Repeat this procedure 4 times, leaving out each scan as testing data each time. This type of procedure, in which each section of the data is left out once, is called cross-validation.

6.3.4. Combine classifier accuracies across the 4 cross-validation folds by averaging.

#### 6.4. Statistical testing

6.4.1. To determine if the classifier is performing better than chance (33%), we can compare results at the group level to chance. To do this, gather the accuracies for each subject, and test that the distribution is different from chance using a nonparametric Wilcoxon Signed-Rank test.

6.4.2. We can also ask whether the classifier is performing better than chance for each individual. To determine the probability of a given accuracy level in chance data, create a null distribution by training and testing the MVPA algorithm on data whose labels have been randomly shuffled. Permute the labels 10,000 times to create a null distribution of accuracy values, and then compare the actual accuracy value to this distribution.

6.4.3. To demonstrate the specificity of the information within the auditory cortex, we can train and test the classifier on voxels from a different location in the brain. Here, we will use a mask of the frontal pole, taken from a probabilistic atlas and warped to fit each subject's individual brain.

### Representative Results

The average classifier accuracy in planum temporale across all 20 participants was 59%. According to the Wilcoxon Signed-Rank test, this is significantly different from chance level of 33%. The mean performance in the frontal pole mask was 32.5% which is not greater than chance. ( **Figure 2** )

The permutation test found that only 5% of the permutations achieved accuracy greater than 42%, thus our statistical threshold for individual subjects is 42%. 19 of the 20 subjects had classifier performance significantly greater than chance using voxels from the planum temporale, while none had performance greater than chance using voxels from the frontal pole.

Thus, we are able to successfully predict from patterns of activity in auditory cortex which of the three sounds the participant was imagining. We were not able to

make this prediction based on activity patterns from the frontal pole, suggesting that the information is not global throughout the brain.

## Applications

MVPA is a useful tool for understanding how the brain represents information. Instead of considering the timecourse of each voxel separately as in a traditional activation analysis, this technique considers patterns across many voxels at once, offering increased sensitivity compared with univariate techniques. Often a multivariate analysis uncovers differences where a univariate technique is not able to. In this case, we learned something about the mechanisms of mental imagery by probing the information content in a specific area of the brain, the auditory cortex. The content-specific nature of these activation patterns would be difficult to test with univariate approaches.

There are additional benefits that come from direction of inference in this kind of analysis. In MVPA we start with patterns of brain activity and attempt to infer something about the mental state of the participant. This kind of “brain reading” approach can lead to the development of brain-computer interfaces, and may allow new opportunities for communication with those with impaired speech or movement.

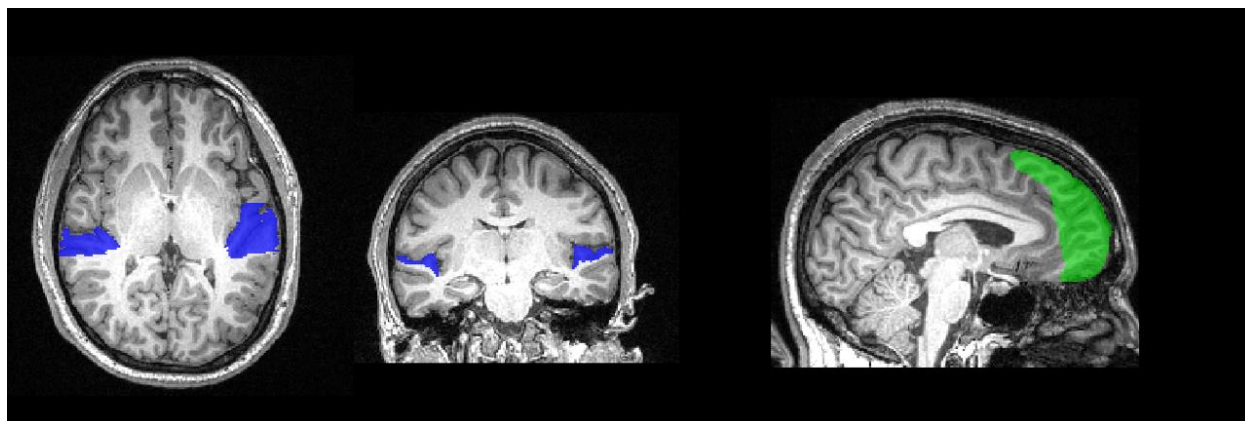
## Legend

**Figure 1: Region of interest tracing.** The surface of the planum temporale has been traced on this participant’s high resolution anatomical image. These voxels will be used for MVPA analysis.

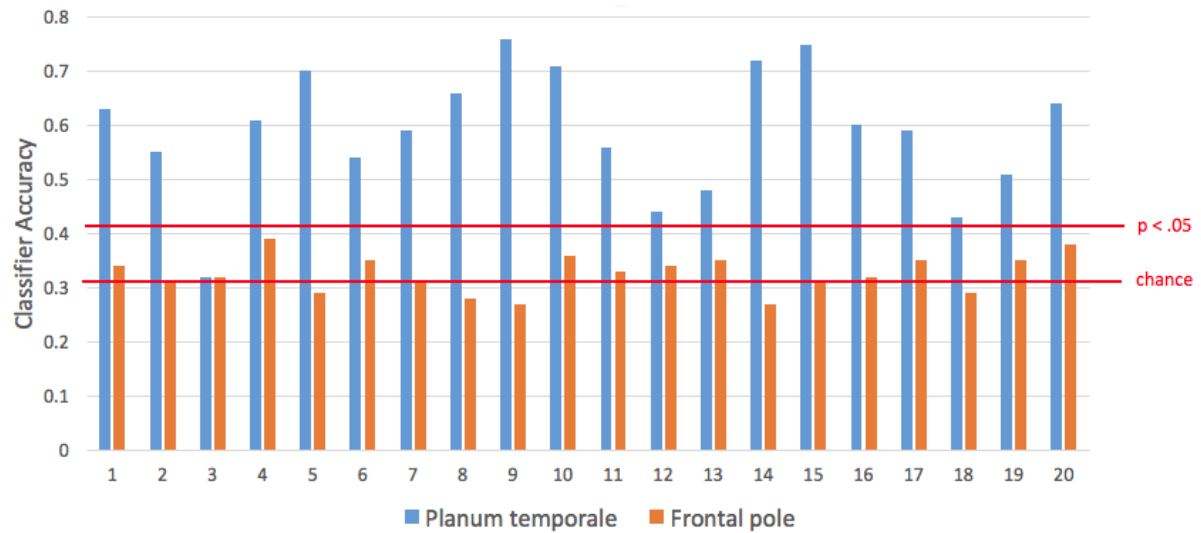
**Figure 2. Classification performance in each participant.** For three-way classification, chance performance is 33%. According to a permutation test, the alpha level of  $p < .05$  corresponds to 42%.

## References

1. Kosslyn, S.M., Ganis, G. & Thompson, W.L. Neural foundations of imagery. *Nat Rev Neurosci* **2**, 635-642 (2001).
2. Haynes, J.D. & Rees, G. Decoding mental states from brain activity in humans. *Nat Rev Neurosci* **7**, 523-534 (2006).
3. Norman, K.A., Polyn, S.M., Detre, G.J. & Haxby, J.V. Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends Cogn Sci* **10**, 424-430 (2006).
4. Meyer, K., *et al.* Predicting visual stimuli on the basis of activity in auditory cortices. *Nat Neurosci* **13**, 667-668 (2010).



**Figure 1: Region of interest tracing.** The surface of the planum temporale has been traced on this participant's high resolution anatomical image, and is shown here in blue. In green is the control mask of the frontal pole. These voxels will be used for MVPA analysis.



**Figure 2. Classification performance in each participant.** For three-way classification, chance performance is 33%. According to a permutation test, the alpha level of  $p < .05$  corresponds to 42%.