

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43

**Cooperating yet distinct brain networks engaged during naturalistic paradigms:  
A meta-analysis of functional MRI results**

**Running Title: Brain networks & naturalistic paradigms**

Katherine L. Bottenhorn<sup>1</sup>, Jessica S. Flannery<sup>1</sup>, Emily R. Boevig<sup>1</sup>, Michael C. Riedel<sup>2</sup>, Simon B. Eickhoff<sup>3,4</sup>,  
Matthew T. Sutherland<sup>1</sup>, Angela R. Laird<sup>2</sup>

<sup>1</sup>Department of Psychology, Florida International University, Miami, FL

<sup>2</sup>Department of Physics, Florida International University, Miami, FL

<sup>3</sup>Institute of Clinical Neuroscience and Medical Psychology, Heinrich-Heine University, Düsseldorf, Germany

<sup>4</sup>Institute of Neuroscience and Medicine, Research Center Jülich, Jülich, Germany

**Corresponding Author**

Dr. Angela R. Laird, Ph.D.  
Professor, Department of Physics  
Florida International University  
Modesto Maidique Campus  
11200 SW 8th Street  
Miami, FL 33199  
305.348.6737 (phone)  
305.348.6700 (fax)  
alaird@fiu.edu

## 1 **Abstract**

2 Cognitive processes do not occur by pure insertion and instead depend on the full complement of co-  
3 occurring mental processes, including perceptual and motor functions. As such, there is limited ecological  
4 validity to human neuroimaging experiments that use highly controlled tasks to isolate mental processes  
5 of interest. However, a growing literature shows how dynamic, interactive tasks have allowed researchers  
6 to study cognition as it more naturally occurs. Collective analysis across such neuroimaging experiments  
7 may answer broader questions regarding how naturalistic cognition is biologically distributed throughout  
8 the brain. We applied an unbiased, data-driven, meta-analytic approach that uses *k*-means clustering to  
9 identify core brain networks engaged across the naturalistic functional neuroimaging literature.  
10 Functional decoding allowed us to, then, delineate how information is distributed between these  
11 networks throughout the execution of dynamical cognition in realistic settings. This analysis revealed  
12 seven recurrent patterns of brain activation, representing sensory, domain-specific, and attentional  
13 neural networks that support the cognitive demands of naturalistic paradigms. Though gaps in the  
14 literature remain, these results suggest that naturalistic fMRI paradigms recruit a common set of networks  
15 that that allow both separate processing of different streams of information and integration of relevant  
16 information to enable flexible cognition and complex behavior.

## 1 Introduction

2 Across the life sciences, researchers often seek a balance between ecological validity and careful  
3 laboratory control when making experimental design decisions. This entails weighing the value of creating  
4 realistic stimuli representative of real-world, interactive experiences versus artificial, reductionist stimuli  
5 facilitating precise assessment of ‘isolated’ mental process of interest via cognitive subtraction. Cognitive  
6 subtraction assumes that a single added cognitive process does not alter the other, co-occurring  
7 processes, both neutrally and cognitively. As such, task-based fMRI has traditionally utilized precisely  
8 controlled tasks to study the neurobiological substrates of cognition. However, cognition does not occur  
9 by pure insertion; the functioning of any cognitive process is not wholly independent from other co-  
10 occurring processes (Friston et al., 1996). Instead, cognition is highly interactive, encompassing  
11 measurable changes in neural activity that are dependent on the full amalgamation of relevant social,  
12 cognitive, perceptual, and motor processes. Thus, it is perhaps unreasonable to expect findings from a  
13 highly restricted assessment of a psychological construct in the scanner to fully generalize to real-world  
14 behaviors and settings.

15 With advances in technology and a desire to study cognition with greater ecological validity, increasing  
16 numbers of studies are utilizing realistic, interactive, and rich stimuli in more ecologically valid  
17 experimental designs that fit within the scanner’s confines (Hasson & Honey, 2012; Maguire, 2012; Wang  
18 et al., 2016). Specifically, the use of video games, film clips, and virtual reality, among others, has brought  
19 a new dimension to cognitive neuroimaging experiments permitting researchers to study brain activity as  
20 participants engage in tasks that more closely represent real-life demands on attention and multimodal  
21 sensory integration. Appreciation of such attention and integration processes necessitates more complex  
22 stimuli than simple static images presented on a screen. Everyday activities, such as navigation or social  
23 observation, involve the integration of processes associated with object recognition, speech  
24 comprehension, motor control, and spatial orienting, which all require the interpretation of dynamic

1 signals often from more than one sensory modality (e.g. audiovisual film watching or visuotactile image  
2 tracing) and necessitate different attentional demands compared to the simplistic stimuli used in  
3 traditional fMRI experiments (Giard & Peronnet, 1999; McGurk & MacDonald, 1976; Sailer, Eggert,  
4 Ditterich, & Straube, 2000; Spence, 2010). Despite offering advantages, the growing body of naturalistic  
5 fMRI research has yet to be quantitatively assessed, and little is known of how the neural bases of these  
6 tasks support complex information processing and behavioral demands.

7 Here, we applied an unbiased, data-driven, meta-analytic approach to quantitatively explore and classify  
8 knowledge embedded in the naturalistic fMRI literature. Using an approach developed by Laird et al.  
9 (2015), we capitalized on the wealth of published naturalistic paradigms and investigated recurrent  
10 patterns of brain activation reported across a wide variety of tasks and behaviors of interest. This method  
11 is based on the premise that functionally similar tasks engage spatially similar patterns of brain activity  
12 and that, by clustering activation patterns from experimental contrasts, similar sets of experimental  
13 paradigms can be identified. We extracted relevant information about the stimuli and task demands of  
14 these paradigms and assessed motifs in the arrangement of this information, with respect the data-driven  
15 clustering analysis, to determine which paradigm aspects elicited activation patterns that subserve  
16 common and dissociable cognitive processes. Although naturalistic paradigms vary greatly and are  
17 designed to probe a wide range of psychological constructs and behaviors, we hypothesized that complex,  
18 multisensory processing are associated with a set of core neural networks engaged by similar content  
19 domains and task demands. The objectives of this study were to first elucidate core brain networks  
20 engaged across naturalistic fMRI paradigms and, then to characterize how information processing is  
21 potentially distributed between these networks to facilitate complex behaviors in realistic settings.

## 22 **Materials and Methods**

### 23 ***Naturalistic fMRI Paradigms***

1 Here, “naturalistic” paradigms were operationally defined as tasks employing any stimulus which  
2 demanded real-time integration of dynamic streams of information. As real-world behavior contextually  
3 involves all sensory modalities, we included naturalistic tasks in which such stimuli were presented via the  
4 visual, auditory, or tactile modalities or any combination thereof. Visual naturalistic tasks require either a  
5 real-time interaction with visual stimuli, in the case of video games and virtual reality, or the continuous  
6 integration of real-time information, such as during film viewing. Auditory tasks, including the perception  
7 of music and spoken stories, similarly require the continuous integration of, and often interaction with,  
8 real-time information. Our operational definition also included tactile naturalistic paradigms, which  
9 involve the manipulation and recognition of physical objects. During these tactile tasks, participants  
10 gather and integrate sensory information to create a mental representation of the object and, if  
11 necessary, form an appropriate behavioral response. Lastly, we note the inclusion of multisensory tasks.  
12 As in life, many naturalistic experiments simultaneously present auditory, visual, and tactile information,  
13 and such tasks demand the real-time integration of information from multiple sensory modalities.

#### 14 ***Literature Search, Filtering, and Annotation***

15 An extensive literature search was performed to amass a corpus of naturalistic fMRI studies that were  
16 published since the emergence of fMRI in 1992. To identify published naturalistic fMRI studies, PubMed  
17 searches were carried out by focusing on stimulus types common to naturalistic research (e.g., video  
18 games, film, virtual reality). The first search string included: (“naturalistic”[Title/Abstract] OR “real-  
19 world”[Title/Abstract] OR “ecologically valid”[Title/Abstract] OR “true-to-life”[Title/Abstract] OR  
20 “realistic”[Title/Abstract] OR “video game”[Title/Abstract] OR “film”[Title/Abstract] OR  
21 “movie”[Title/Abstract] OR “virtual reality”[Title/Abstract]) AND (“fMRI”[Title/Abstract] OR “functional  
22 magnetic resonance imaging”[Title/Abstract]) AND (“Humans”[MeSH])). This search yielded 679 studies  
23 (January 2016), some of which utilized stimulus types that we had not included in our initial query,  
24 including music, speech, and tactile objects. To identify any studies using these tasks that may not have

1 been returned by initial query, a second search was performed using the string [("music"[Title/Abstract]  
2 OR "speech"[Title/Abstract] OR "spoken"[Title/Abstract] OR "tactile object"[Title/Abstract]) AND  
3 ("naturalistic"[Title/Abstract] OR "real-world"[Title/Abstract] OR "ecologically valid"[Title/Abstract] OR  
4 "true-to-life"[Title/Abstract] OR "realistic"[Title/Abstract]) AND ("fMRI"[Title/Abstract] OR "functional  
5 magnetic resonance imaging"[Title/Abstract]) AND "Humans"[MeSH]]. This secondary search returned 48  
6 studies, some of which were included in the results of the first search. The two sets of search results were  
7 pooled, reviewed, and filtered to identify studies utilizing naturalistic paradigms as defined above, yielding  
8 a total of 230 candidate studies deemed potentially suitable for meta-analysis.

9 Each of the remaining 230 candidate studies were then reviewed according to the following exclusion  
10 criteria. Non-naturalistic tasks were excluded in which static, timed blocks of stimuli were presented with  
11 a well-defined window for participant response. We excluded studies that assessed training or learning  
12 across multiple trials or across some period of practice (e.g., pre vs. post contrasts), as our focus was on  
13 neural underpinnings of the tasks themselves and not training-induced changes thereof. Studies of  
14 participants under the age of 18 or of participants with any history of neurological or psychiatric diagnosis  
15 were excluded. After this study-level screening, we then inspected each reported contrast from the  
16 studies. Experimental contrasts from analyses that used an *a priori* region(s) of interest to investigate  
17 activation or functional connectivity were omitted permitting identification of whole-brain neural  
18 networks. We also excluded contrasts modeling ANOVA interaction-specific activations, due to the  
19 inherent complexity of such effects. Finally, any studies/contrasts that did not report the brain activation  
20 locations in a three-dimensional coordinate space were discarded.

21 During inspection of each contrast, one study associate (KLB) manually annotated each experiment with  
22 a set of metadata terms that described the experimental design with respect to stimulus type utilized,  
23 sensory modality engaged, and the task nature. These terms described the salient aspects of the stimuli  
24 and behaviors associated with each of the naturalistic paradigms, annotating the tasks themselves and

1 not the intended psychological construct interrogated by the original report. These manual annotations  
2 were then independently reviewed and confirmed by a second study associate (JSF) to assure consistency  
3 and accuracy. Any disagreements or inconsistencies between KLB and JSF were resolved following a final  
4 conversation between the two associates.

## 5 ***Experimental Design and Statistical Analysis***

### 6 *Modeled Activation Maps*

7 Following the identification of relevant papers and experiments/contrasts, reported brain activation  
8 coordinates were extracted. All Talairach atlas-based coordinates (Talairach & Tournoux, 1988) were  
9 converted to Montreal Neurological Institute (MNI) space (Collins, Neelin, Peters, & Evans, 1994; Evans et  
10 al., 1993) using the tal2icbm transformation (Lancaster et al., 2007; Laird et al., 2010). Probabilistic  
11 modeled activation (MA) maps were created from the foci reported in each individual contrast by  
12 modeling a spherical Gaussian blur around each foci, representative of the uncertainty induced by the  
13 inherent variability from individual differences and between-lab differences and weighted by the number  
14 of subjects included in the experiment (Eickhoff et al., 2009). These MA maps were concatenated into an  
15 array of  $n$  experiments by  $p$  voxels, which was then analyzed for pairwise correlations that reflected the  
16 degree of spatial similarity between the MA maps from each of the  $n$  experiment and those of every other  
17 experiment. The resultant  $n \times n$  correlation matrix represented the similarity of spatial topography of MA  
18 maps between every possible pair of experiments.

### 19 *K-Means Clustering Analysis*

20 Individual naturalistic experiments ( $n$  MA maps) were then classified into  $K$  groups based on their spatial  
21 topography similarities. The  $k$ -means clustering procedure was performed in Matlab (Mathworks, R2013b  
22 for Linux), which grouped experiments by pairwise similarity, calculating correlation distance by one  
23 minus the correlation between MA maps (from the aforementioned correlation matrix) and finding the

1 “best” grouping by minimizing the sum of correlation distances within each cluster. This approach begins  
2 by choosing  $K$  arbitrary maps as representative centroids for each of  $K$  clusters and assigning experiments  
3 to each cluster based on the closest (most similar) centroid. This process continued iteratively until a  
4 stable solution was reached.

5 Solutions were investigated for a range of  $K = 2 - 10$  clusters. Once the clustering analysis was complete  
6 for all  $K$ , we compared each solution to the neighboring solutions and assessed for improvement across  
7 parcellation schemes using four metrics describing cluster separation and stability (Bzdok et al., 2015;  
8 Eickhoff, Laird, Fox, Bzdok, & Hensel, 2016). This allowed us to objectively select the number of clusters  
9 that most optimally divided the data set. The first metric, *average cluster silhouette* across clustering  
10 solutions, assessed the separation between clusters and described whether clusters were distinct or  
11 overlapping. A higher silhouette value indicates that greater separation is ideal and that each experiment  
12 fits well into its cluster, with lower misclassification likelihood of fringe experiments into neighboring  
13 clusters. Stability is indicated by minimal change in silhouette from one solution ( $K$ ) to the next ( $K + 1$ ).  
14 Second, we considered the *consistency of experiment assignment* by comparing the ratio of the minimum  
15 number of experiments consistently assigned to a cluster relative to the mean number of experiments  
16 consistently assigned to that cluster. In this case, only ratios above 0.5, in which at least half of the  
17 experiments were considered viable solutions. Third, the *variation of information* was quantified, which  
18 compared the entropy of clusters with the mutual information shared between them for each solution  $K$   
19 and its  $K - 1$  and  $K + 1$  neighbors. A significant decrease in variation of information from  $K - 1$  to  $K$  and  
20 increase from  $K$  to  $K + 1$  indicated a decrease in overlap between solutions and, thus, stability of solution  
21  $K$ . Finally, we computed a *hierarchy index* for each solution, which assessed how clusters split from the  $K$   
22  $- 1$  to  $K$  solution to form the additional cluster. A lower hierarchy index indicated that clusters present in  
23  $K$  stemmed from fewer of the clusters present in  $K - 1$ , another indication of stability in groupings. An  
24 optimal clustering solution is one that demonstrated minimal overlap between clusters (i.e., high



1 silhouette value), while exhibiting relative stability in comparison with the previous and next solutions  
2 (i.e., consistency > 0.5, a local minimum in variation of information, and lower hierarchy index than  
3 previous).

#### 4 *Meta-Analytic Groupings*

5 From the identified optimal clustering solution, we probed the underlying neural topography associated  
6 with each of the  $K$  groups of experiments (Laird et al., 2015). To this end, the ALE meta-analysis algorithm  
7 (Turkeltaub et al., 2002; Laird et al., 2005) was applied to generate a map of convergent activation for  
8 each grouping of experiments with similar topography. The ALE algorithm includes a weighting of the  
9 number of subjects when computing these maps of convergent activation and accounts for uncertainty  
10 associated with individual, template, and registration differences between and across experiments  
11 (Eickhoff et al., 2009; Turkeltaub et al., 2012). The union of these probability distributions was used to  
12 calculate ALE scores, a quantitative assessment of convergence between brain activation across different  
13 experiments, which was compared against 1000 permutations of a null distribution of random spatial  
14 arrangements (Eickhoff, Bzdok, Laird, Kurth, & Fox, 2012). These ALE values for each meta-analytic  
15 grouping of experiments were thresholded at  $P < 0.01$  (cluster-level corrected for family-wise error) with  
16 a voxel-level, cluster-forming threshold of  $P < 0.001$  (Eickhoff, Nichols, et al., 2016; Woo, Krishnan, &  
17 Wager, 2014). The resultant ALE maps thus reflected the convergent activation patterns within each of  
18 the  $K$  clusters. The experimental  $K$  clusters are hereafter referred to as meta-analytic groupings (MAGs),  
19 representing meta-analytic groups of experiments demonstrating similar activation patterns.

#### 20 *Functional Decoding*

21 Once we elucidated convergent activation patterns within MAGs, we sought to gain insight into what  
22 aspects of the naturalistic paradigms were most frequently associated with each MAG via functional  
23 decoding. Functional decoding is a quantitative, data-driven method by which researchers can infer which

1 mental processes are related to activation in a specific brain region (or set of brain regions) across  
2 published fMRI studies. We chose to use two complementary functional decoding approaches, one based  
3 on our study-specific, subjective manual annotations mentioned above, and another based on the  
4 objective, automated annotations provided by the Neurosynth database for over 11,000 functional  
5 neuroimaging studies (Yarkoni et al., 2011; Neurosynth.org). First, the manually annotated terms  
6 associated with each experiment were grouped into the MAGs identified above and were assessed by  
7 frequency of occurrence in each MAG. The distribution of stimulus modality, stimulus type, and salient  
8 terms across MAGs allowed us to evaluate the relationship between activation patterns and the aspects  
9 of naturalistic paradigms that elicited them. Second, we included an automated, data-driven annotation  
10 method using Neurosynth, which includes automatically extracted terms that occur at a high frequency  
11 in the abstract of each archived study. To functionally decode our MAGs, we compared the MAGs'  
12 activation patterns with those reported across published neuroimaging papers in the Neurosynth  
13 database. To this end, we uploaded each ALE map to NeuroVault, a web-based repository for 3D statistical  
14 neuroimaging maps that directly interfaces with Neurosynth (Gorgolewski et al., 2015; NeuroVault.org).  
15 NeuroVault enables “functional decoding” by correlating unthresholded uploaded maps with term-  
16 specific meta-analytic maps extracted from Neurosynth’s database of published functional neuroimaging  
17 studies. The Neurosynth functional decoding results were exported as a set of terms and correlation  
18 values representing how well the spatial distribution of activation associated with each term in the  
19 database matched the activation pattern of the uploaded map.

20 Both sets of terms (i.e., obtained via manual and automated approaches) were evaluated to assess the  
21 specific aspects of naturalistic paradigms associated with each MAG. The Neurosynth terms representing  
22 broad behavioral aspects across fMRI studies that elicit similar brain activation profiles provides both an  
23 unbiased description of the experiments engaging each MAG, as well as a comparison of our corpus of  
24 studies with the broader literature. On the other hand, manual annotation provides more concise,

1 accurate description of the paradigms, though it is predisposed to the subjective bias of human  
2 annotation. The results of this two-pronged functional decoding approach were designed to describe the  
3 processes that engage brain networks similar to each MAG and how these processes may be similar or  
4 different in naturalistic fMRI studies compared to the broader functional neuroimaging literature. The  
5 distribution of stimulus modalities and types across MAGs was assessed, too. Together, the functional  
6 decoding results and distributions of different stimuli were interpreted to provide insight into how  
7 information processing is functionally segregated across cooperating neural systems during naturalistic  
8 tasks.

## 9 **Results**

10 Of the 230 papers meeting inclusion, 120 were omitted based on further exclusion criteria, leaving 110  
11 studies that reported coordinates of brain activation from naturalistic fMRI tasks among healthy adults.  
12 The final data set included activation foci from 376 experimental contrasts ( $N = 1,817$  subjects) derived  
13 from tasks using a variety of stimulus types and sensory modalities. Across our corpus of naturalistic fMRI  
14 experiments, approximately 55% assessed a single stimulus modality, including 40% visual stimuli, 13%  
15 auditory, and 1% tactile.

<b>Stimulus Modality</b>	<b>Number of Experiments</b>
Auditory	50 (13%)
Audiovisual	154 (41%)
Visual	150 (40%)
Visual + tactile (pain)	9 (2%)
Visual + tactile	5 (1%)
Tactile	4 (1%)

16 **Table 1.** Distribution of stimulus modalities across the naturalistic corpus. Paradigms engaged auditory,  
17 visual, and tactile sensory modalities, both separately and in combination.

1 Conversely, 45% of experiments utilized multisensory stimuli, including 41% that employed audiovisual  
2 stimuli, 2% in which a visual stimulus was paired with painful, tactile stimuli, and 1% pairing visual and  
3 non-painful tactile stimuli (Table 1). Of the visual experiments, 69% involved a motor response, as did 25%  
4 of the audiovisual experiments, ranging from a button press to joystick and object manipulation. The  
5 stimulus types most frequently used across the included experiments were films (45%), virtual reality  
6 (32%), speech (9%), and music (6%) (Table 2).

Stimulus Type	Number of Experiments
Film	169 (45%)
Virtual Reality	121 (32%)
Speech	32 (9%)
Music	21 (6%)
Video Game	13 (4%)
3D image	6 (2%)
Tactile	6 (2%)
Picture	4 (1%)
Sounds	1 (<1%)

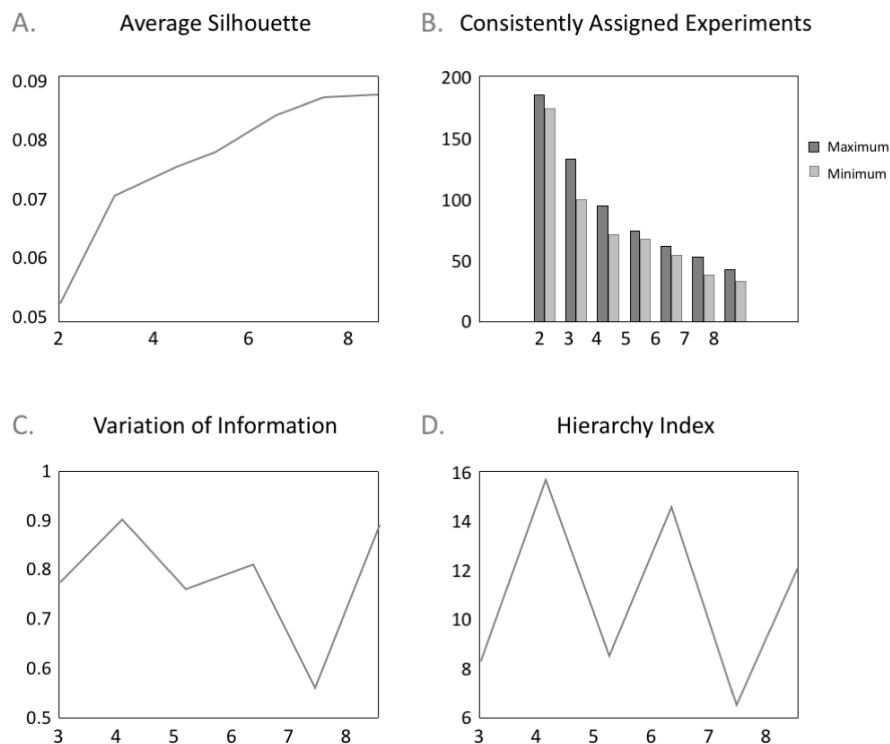
7 **Table 2.** Distribution of stimulus types across the naturalistic corpus. Within each stimulus modality,  
8 multiple types of experimental stimuli were included across the data set.

9

## 10 ***K-Means Clustering Solutions***

11 MA maps were created for each contrast, and then clustered to identify groups with similar activation  
12 topographies. For completeness, the *k*-means clustering solutions for  $K = 2 - 10$  clusters were  
13 quantitatively evaluated across four metrics to identify an optimal solution (Figure 1). When considering  
14 the average silhouette metric (Fig. 1A), values generally increased as *K* increased and the *smallest increase*  
15 was observed between  $K = 7$  to  $K = 8$ , indicating little additional separation between clusters gained by  
16 moving from 7 to 8 clusters. With respect to the consistency of assigned experiments metric (Fig. 1B),

1 each of the solutions  $K = 2 - 10$  met the stability requirement whereby the minimum number of  
2 experiments included in any iteration of the solution was at least 50% of the mean number of experiments  
3 included across iterations. The variation of information metric (Fig. 1C), suggested the stability of both 5-  
4 and 7-cluster solutions as parameter value *decreases* were observed when moving from  $K = 4$  to  $K = 5$  and  
5 from  $K = 6$  to  $K = 7$ , which also combined with parameter *increases* when moving from  $K = 5$  to  $K = 6$  and  
6  $K = 7$  to  $K = 8$ . The hierarchy index metric (Fig.1D) further corroborated a 5- and 7-cluster solution where  
7 local minima were observed for these two solutions. While both the 5- and 7-cluster solutions appeared  
8 stable (Fig. 1C, 1D) and consisted of groupings with a satisfactory amount of separation (Fig 1A), based on  
9 the smaller increase in average cluster silhouette from  $K = 7$  to  $K = 8$  (Fig. 1A), we proceeded with the  $K =$   
10 7 solution in subsequent analyses.



11

12 **Figure 1. Metrics computed for  $K = 2 - 10$  clustering solutions.** (A) The average cluster silhouette for each  
13 solution  $K$  from 2 to 10 clusters. (B) Consistency in experiments assignment to clusters, plotting the  
14 minimum consistently assigned clusters next to the mean of consistently assigned clusters. (C) The change  
15 in variation of information, a distance metric, from the  $K - 1$  to  $K$  and from  $K$  to  $K + 1$ . (D) The hierarchy  
16 index for each of  $K$  clustering solutions, which provides information about how clusters in the  $K$  solution

1 stemmed from clusters in the  $K - 1$  solution.

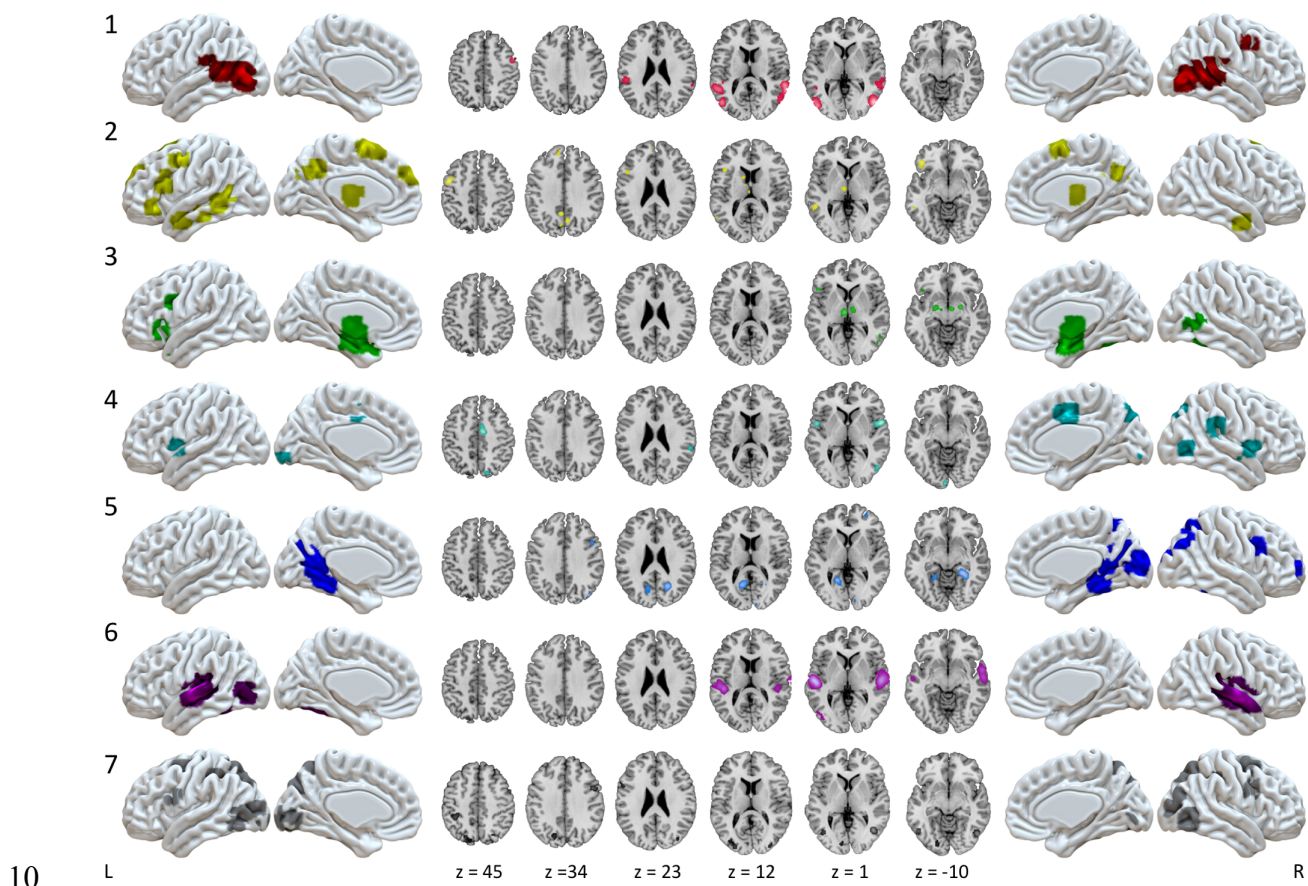
2 In this  $K = 7$  solution, four experiments were not grouped into any cluster, due to the dissimilarity of their  
3 functional activation patterns with those across the rest of the data set, yielding a total set of 372  
4 experiments across seven clusters.

### 5 **Meta-Analytic Groupings**

6 The optimal clustering solution yielded seven meta-analytic groupings (MAGs) of experiments in our  
7 corpus, suggesting similarities in brain activation across this sample of the naturalistic literature coalesce  
8 into seven distinct patterns. The number of experiments that were clustered into each MAG were  
9 uniformly distributed and ranged from 41 to 77 experiments (*mean* = 53.14; *SD* = 12.46; Table 3). ALE maps  
10 of the seven MAGs were generated and demonstrated little overlap in activation patterns, suggesting  
11 distinct patterns of recurrent activation across our set of naturalistic experiments (Figure 2, Supplemental  
12 Table 1). Whereas some of the MAGs exhibited focal patterns of convergent activation, restricted to a  
13 single or neighboring gyri (e.g., MAG 1 and 6), others presented with distributed convergence across  
14 multiple lobes (e.g., MAG 2 and 5). Most of the resulting MAGs were restricted to cortical activation  
15 patterns, although MAGs 2 and 3 exhibited convergent activation in subcortical and brainstem regions.

16 **MAG 1** included convergent activation in the bilateral posterior temporal areas, including portions of the  
17 inferior, middle, and superior temporal gyri, extending into the inferior parietal lobule on the right and  
18 into the middle occipital gyrus on the left, as well as in the left supramarginal gyrus and right precentral  
19 and middle frontal gyri. **MAG 2** exhibited convergence in left inferior frontal gyrus, left precentral gyrus,  
20 anterior and posterior aspects of the middle temporal gyrus, precuneus, thalamus, and caudate nucleus,  
21 in addition to both the left and right superior frontal gyri. **MAG 3** demonstrated a largely symmetric  
22 convergence pattern across multiple subcortical structures including bilateral amygdalae, thalamus,  
23 parahippocampal gyrus, and periaqueductal gray, with cortical clusters observed in the left inferior frontal

1 sulcus and inferior frontal gyrus. **MAG 4** included clusters of convergence in right supramarginal gyrus and  
2 right midcingulate cortex as well as in the left lingual gyrus, right anterior middle occipital gyrus, right  
3 precuneus, and bilateral inferior frontal gyri. **MAG 5** exhibited convergent activation in bilateral medial  
4 temporal lobe and hippocampal regions, bilateral precuneus, retrosplenial posterior cingulate cortex, and  
5 occipital regions including the lingual gyrus, right calcarine sulcus, and cuneus. **MAG 6** showed  
6 convergence in the bilateral superior temporal gyri and sulci and the left temporo-occipital junction. **MAG**  
7 **7** demonstrated convergence in the bilateral superior frontal gyri, intraparietal sulci, and inferior parietal  
8 lobules as well as convergence in higher-order visual processing areas in the middle occipital and lingual  
9 gyri.



11 **Figure 2. Convergent activation patterns of MAGs from the naturalistic corpus.** ALE meta-analysis of  
12 experiments in each MAG yielded seven patterns of convergent activation.

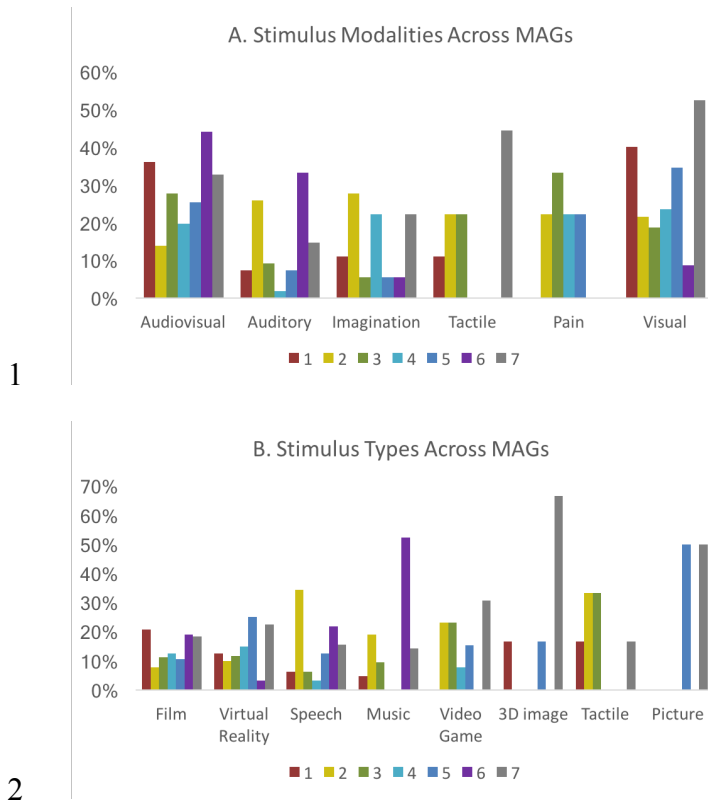
## 1 ***Stimulus Distribution Across MAGs***

2 Each stimulus modality was represented in multiple MAGs, but modalities were not evenly distributed  
3 across MAGs (Figure 3A). Experiments utilizing audiovisual tasks were somewhat uniformly distributed  
4 across the MAGs, with a slightly higher proportion of audiovisual tasks in MAGs 1, 3, and 6. In contrast,  
5 more than half of the experiments using auditory tasks were grouped into MAGs 2 and 6. Notably, more  
6 experiments based on auditory and audiovisual stimuli were clustered into MAG 6 than any other MAG.  
7 Experiments in which participants experienced physical pain were not present in MAGs 1, 6, and 7, but  
8 distributed nearly evenly among MAGs 2 through 5, with a slightly higher portion in MAG 3. Experiments  
9 that used tactile stimuli were grouped into MAG 7 twice as often as any other MAG (MAGs1-3). Visual  
10 experiments were more evenly distributed across clusters, though there was a markedly smaller  
11 proportion in MAG 6 than any other MAG. The complete distribution of stimulus modalities across MAGs  
12 is provided in Supplemental Table 2.

13 As with stimulus modality, most stimulus types showed unequal, but not necessarily selective, distribution  
14 across MAGs (Figure 3B). Film-based experiments and virtual reality tasks were uniformly distributed  
15 across MAGs; and tasks utilizing spoken stimuli were more frequently grouped into MAGs 2 and 6. Again,  
16 auditory stimuli were highly associated with MAG 6, as ~50% of music experiments and ~20% of speech  
17 experiments were clustered into MAG 6. Experiments that required subjects to play video games were  
18 most often grouped into MAGs 2, 3, and 7. Experimental contrasts which included a condition in which  
19 participants received tactile stimulation, were most prevalent in MAGs 3 and 7. A detailed distribution of  
20 stimulus types across MAGs is shown in Supplemental Table 3.

21





**Figure 3. Distribution of stimulus modalities and types across MAGs.** (A) The presence of each sensory modality across the corpus that is associated with each MAG. (B) The proportion of each stimulus type present within the corpus that is associated with each MAG. These percentages represent the proportion modality or stimulus type present in each MAG, compared to the total count of that modality or stimulus type across all MAGs.

## 8 **Functional Decoding**

9 Two approaches for functionally decoding each MAG, manual and automated annotations, were  
 10 performed to develop a functional interpretation of each MAGs' association with aspects of naturalistic  
 11 paradigms.

## 12 *Manual Annotations*

13 Our manual annotations utilized a list of 26 corpus-specific metadata terms, which captured salient  
 14 features of the naturalistic design, rather than the psychological constructs assumed to be involved. Table  
 15 4 displays the frequency distribution of these terms (Column = "Total"), highlighting which terms  
 16 described the largest number of experiments in our corpus (e.g., "navigation", "visual features",

1 “emotional film”, “attention”), as well as those that accounted for a minimal number of experiments in  
2 our corpus (e.g., “violence”, “tactile”, “pain”). These values indicate the percent of experiments labeled  
3 with each term, or the base-rate of each term throughout the data set, keeping in mind that each  
4 experiment was labeled with only one or two terms. Once the experiments were clustered into seven  
5 MAGs, we evaluated the relative contributions of each term per MAG, controlling for base-rate by dividing  
6 each term’s per-MAG count by that term’s total count across the corpus (Table 4). These outcomes  
7 provide the association of each term with each MAG (Table 4). Some of the terms in the manual  
8 annotation analysis corresponded to stimulus types in Figure 3B (e.g., per-MAG distribution for “music”  
9 and “video game”). However, many of the manually derived terms highlighted experimental aspects that  
10 reflect the unique and salient features of the naturalistic corpus (e.g., “anthropomorphic”, “violence”) and  
11 are not included in standard neuroimaging paradigm ontologies such as BrainMap (Fox et al., 2005) or  
12 CogPO (Turner & Laird, 2012).

### 13 *Automated Neurosynth Annotations*

14 To complement the manual annotation analysis, we used Neurosynth’s automated annotations, which  
15 describes experiments that engage each MAG based on published neuroimaging data, allowing  
16 comparison of our corpus with the broader literature. MAG results were decoded in Neurosynth, yielding  
17 correlation values indicating the similarity of the input map (i.e., each MAG’s ALE map) and maps  
18 associated with each term from the Neurosynth database. To facilitate interpretation, the top ten terms  
19 with the highest correlation values for each MAG are presented (Table 5). Terms that were near-duplicates  
20 of terms already included in the list were removed, such as “emotion” and “emotions” if “emotional” was  
21 higher on the list. Non-content terms (e.g. “abstract”, “reliable”) and terms that described brain regions,  
22 such as “insula” or “mt”, were also excluded

Term	Frequency per MAG														
	Total	MAG 1		MAG 2		MAG 3		MAG 4		MAG 5		MAG 6		MAG 7	
Anthropomorphic	21 (6%)	6 (25%)	0 (0%)	1 (4%)	7 (29%)	3 (13%)	1 (17%)	3 (13%)							
Attention	50 (14%)	17 (30%)	4 (7%)	1 (2%)	2 (4%)	8 (12%)	11 (32%)	8 (14%)							
Auditory features	17 (5%)	3 (16%)	0 (0%)	0 (0%)	1 (5%)	2 (11%)	11 (68%)	0 (0%)							
Congruence	2 (6%)	3 (12%)	4 (15%)	0 (0%)	1 (4%)	4 (15%)	3 (27%)	7 (27%)							
Emotional film	59 (16%)	14 (23%)	3 (5%)	15 (25%)	6 (10%)	4 (5%)	11 (23%)	6 (10%)							
Encoding	24 (6%)	3 (10%)	4 (14%)	1 (3%)	2 (7%)	5 (17%)	1 (21%)	8 (28%)							
Erotic	15 (4%)	2 (12%)	1 (6%)	3 (18%)	4 (24%)	2 (12%)	0 (12%)	3 (18%)							
Faces	21 (6%)	7 (30%)	0 (0%)	2 (9%)	1 (4%)	2 (9%)	7 (39%)	2 (9%)							
Imagination	23 (6%)	3 (12%)	5 (19%)	1 (4%)	5 (19%)	3 (12%)	1 (15%)	5 (19%)							
Inference	11 (3%)	0 (0%)	6 (50%)	1 (8%)	2 (17%)	1 (8%)	0 (8%)	1 (8%)							
Language	47 (13%)	4 (7%)	10 (19%)	2 (4%)	2 (4%)	7 (13%)	12 (35%)	10 (19%)							
Movement	14 (4%)	1 (6%)	0 (0%)	1 (6%)	5 (31%)	2 (13%)	1 (19%)	4 (25%)							
Music	21 (6%)	1 (5%)	4 (19%)	2 (10%)	0 (0%)	0 (0%)	11 (52%)	3 (14%)							
Narrative	30 (8%)	3 (9%)	3 (9%)	2 (6%)	2 (6%)	5 (14%)	10 (43%)	5 (14%)							
Navigation	81 (22%)	5 (5%)	8 (8%)	11 (10%)	12 (11%)	24 (23%)	2 (3%)	19 (18%)							
Negative valence	24 (6%)	4 (17%)	1 (4%)	8 (33%)	2 (8%)	0 (0%)	4 (17%)	5 (21%)							
Pain	9 (2%)	0 (0%)	2 (18%)	3 (27%)	2 (18%)	2 (18%)	0 (18%)	0 (0%)							
Positive valence	11 (3%)	3 (23%)	1 (8%)	2 (15%)	0 (0%)	2 (15%)	1 (23%)	2 (15%)							
Recognition	12 (3%)	1 (8%)	3 (23%)	1 (8%)	2 (15%)	1 (8%)	1 (15%)	3 (23%)							
Retrieval	23 (6%)	1 (4%)	6 (23%)	4 (15%)	5 (19%)	3 (12%)	1 (15%)	3 (12%)							
Social	26 (7%)	3 (10%)	6 (21%)	4 (14%)	5 (17%)	3 (10%)	1 (14%)	4 (14%)							
Tactile	9 (2%)	1 (11%)	2 (22%)	2 (22%)	0 (0%)	0 (0%)	0 (0%)	4 (44%)							
Video game	15 (4%)	0 (0%)	3 (17%)	3 (17%)	1 (6%)	3 (17%)	0 (17%)	5 (28%)							
Violence	8 (2%)	0 (0%)	3 (33%)	3 (33%)	0 (0%)	1 (11%)	0 (11%)	1 (11%)							
Visual features	65 (17%)	18 (24%)	4 (5%)	2 (3%)	2 (3%)	10 (14%)	7 (22%)	22 (29%)							

15 **Table 4.** Manual functional decoding results across meta-analytic groupings. The relative contributions of each manually-derived metadata term  
16 (e.g., term frequencies) were computed for all MAGs, controlling for the base-rate by dividing each term's per-MAG count by that term's total  
17 count across the corpus. Base-rates are provided as the total count for each term.

## 1 *Functional Interpretation of MAGs*

2 The combined knowledge gained from the MAGs topography, as well as the manual and automated  
3 metadata decoding analyses provided insight into the functional interpretation of the clustering results.  
4 Overall, the terms yielded by Neurosynth decoding generally agreed with the manual annotation terms in  
5 characterizing the MAGs. Below is a summary of the seven MAGs; note that reported labels do not refer  
6 to the definitive function of these regions, but rather indicate how each MAG reflects differential network  
7 contributions during naturalistic fMRI paradigms.

8 Manual annotations indicated that **MAG 1** experiments involved attention and the processing of dynamic  
9 visual features, in addition to visually-presented anthropomorphic forms and faces. Most of the stimuli in  
10 these experiments were films (Fig 3B), especially affective films. Neurosynth results largely converged  
11 with these manual annotations, as terms including “*videos*”, “*body*”, “*observation*”, and “*visual motion*”  
12 (Table 5) were associated with activations in MAG 1 regions. These annotations, together with the  
13 presence of convergent activation across regions commonly associated with higher-level visual  
14 processing, suggest that MAG 1 was associated with the ***Observation of Body and Biological Motion***  
15 (Figure 2.1).

16 Manual annotations indicated that **MAG 2** experiments involved language processing and  
17 comprehension, retrieval, encoding, inference, and judgements about congruence. This MAG included  
18 relatively large proportions of the experiments using speech, video games, and tactile stimulation (Fig 3B).  
19 Neurosynth results supported the manual annotations’ indication that this MAG was associated with  
20 language processing and comprehension, as terms such as “*sentence*”, “*comprehension*”, “*semantic*”, and  
21 “*mentalizing*” (Table 5) were returned. These annotations and the presence of convergent activation in  
22 predominately left lateralized regions typically associated with higher-order cognition and language  
23 suggest that MAG 2 related to ***Language Processing*** (Figure 2.2).

24 Manual annotations indicated that **MAG 3** experiments involved human interactions or affective displays,  
25 including emotional films with a negative valence. Films were the predominantly used stimuli across these  
26 experiments, while video games, virtual reality, and painful stimuli made up a smaller proportion of  
27 paradigms (Fig 3B). Neurosynth results corroborated these manual annotation interpretations regarding  
28 affective and social processing, with terms such as “*emotional*”, “*faces*”, “*fearful*”, and “*valence*” (Table  
29 5). Together, these annotations and a convergent activation pattern involving bilateral amygdalae suggest  
30 that MAG 3 was associated with ***Emotional Processing*** (Figure 2.3).

31 Manual annotations indicated that more than half of the experiments in **MAG 4** required participants to  
32 respond to stimuli with an executed or imagined motor movement. Over 90% of the paradigms utilized  
33 by experiments in this MAG were based on films and virtual reality (Extended Data Figure 3-2), the former  
34 of which displayed human social and erotic interactions, emotional displays, and anthropomorphic figures  
35 (Table 4). Neurosynth results were notably absent of social terms, but corroborated the manual  
36 annotation interpretation of motor execution, as terms associated with activation in these regions include  
37 “*execution*”, “*visual motion*”, and “*action*” (Table 5). These results, together with convergent activation  
38 in, suggest that MAG 4 was linked with ***Motor Planning*** (Figure 2.4).

MAG 1		MAG 2		MAG 3		MAG 4		MAG 5		MAG 6		MAG 7	
NS term	corr.	NS term	corr.	NS term	corr.	NS term	corr.	NS term	corr.	NS term	corr.	NS term	corr.
motion	0.418	sentence	0.334	emotional	0.380	painful	0.208	navigation	0.29	auditory	0.638	visual	0.441
static	0.323	comprehension	0.328	expressions	0.375	execution	0.168	virtual	0.286	sounds	0.628	motion	0.329
videos	0.315	language	0.287	neutral	0.375	motion	0.138	scene	0.269	listening	0.595	spatial	0.322
body	0.312	semantic	0.271	facial	0.366	hand	0.115	memory	0.22	acoustic	0.588	action	0.315
moving	0.285	linguistic	0.252	fearful	0.355	video	0.105	episodic	0.217	sound	0.588	eye movements	0.284
perception	0.279	verb	0.241	affective	0.349	action	0.098	retrieval	0.202	pitch	0.582	actions	0.278
visual	0.273	syntactic	0.237	happy	0.338	visual motion	0.095	place	0.195	speech	0.559	execution	0.276
observation	0.271	mind	0.233	facial expressions	0.337	eye movements	0.092	visual	0.185	tones	0.526	observation	0.264
visual motion	0.267	theory mind	0.231	fear	0.323	observer	0.085	remembering	0.149	music	0.521	hand	0.247
viewed	0.259	mentalizing	0.219	valence	0.317	noxious	0.083	encoding	0.139	spectral	0.52	attention	0.245

1

2 **Table 5.** Automated functional decoding results from Neurosynth. The top ten Neurosynth (NS) terms are provided for each MAG, along with the  
3 corresponding Pearson’s correlation coefficient (“corr”) that indicates the strength of similarity between Neurosynth maps and each MAG

1  
2 Manual annotations indicated that **MAG 5** heavily represented experiments involving navigation through  
3 virtual reality environments, with both attentional and memory demands related to encoding unfamiliar  
4 virtual landscapes for future use. A few of these experiments required language processing, as well, and  
5 half of the experiments that used 3D images were grouped into MAG 5 (indicate table or fig where this  
6 info comes from). The manual annotations were reflected in the Neurosynth results, as similar patterns  
7 of activation have been associated with “*navigation*”, “*scene*”, “*memory*”, and “*place*”. Additional related  
8 terms added depth to our characterization, expanding on the memory demands with “*encoding*”,  
9 “*episodic*”, “*remembering*”, and “*retrieval*” (Table 5). Overall, these experimental characteristics and  
10 convergent activation in medial temporal regions and along the visual processing stream suggest that  
11 MAG 5 was associated with **Navigation and Spatial Memory** (Figure 2.5).

12 Manual annotations showed that **MAG 6** experiments primarily involved either film or music stimuli  
13 (Figure 3B) and engaged either audiovisual or purely auditory processing (Fig. 3A). More than half of the  
14 included experiments that used music or spoken narratives as stimuli were grouped into this MAG (Fig.  
15 2B), with some stimuli involving an emotional quality (Table 4). Neurosynth corroborated these  
16 interpretations returning terms such as “*auditory*”, “*sounds*”, “*listening*”, and “*pitch*” associated with  
17 activation of the regions in this MAG. These metadata descriptions combined with convergent activation  
18 in superior temporal regions suggest this MAG’s association with **Auditory Processing** (Figure 2.6).

19 Manual annotations of **MAG 7** experiments implicated tasks involving visual attentional demands and the  
20 processing of visual features, as participants engaged in both film-watching and virtual reality navigation  
21 (Figure 3B, Table 4). Some experiments involved judgements of stimuli congruency, memory encoding,  
22 and language processing. Neurosynth supported this characterization returning terms including “*visual*,”  
23 “*attention*”, “*eye movements*”, and “*observation*” associated with activation of the regions in this MAG  
24 (Table 5). These annotations and convergent activation in regions resembling the dorsal attention network  
25 and areas of higher level visual processing (e.g., superior frontal and parietal regions, extrastriate cortex)  
26 suggest this MAG’s association with **Visual Attention** (Figure 2.7).

## 27 **Discussion**

28 To characterize a core set of brain networks engaged in more ecologically valid neuroimaging designs, we  
29 employed a data-driven approach that meta-analytically grouped published naturalistic fMRI results  
30 according to their spatial topographies. Objective metrics suggested that a solution of  $K = 7$  clusters  
31 provided the most stable and disparate grouping of experiments across the naturalistic fMRI literature,  
32 and ALE meta-analysis delineated convergent activation across spatially distinct brain regions for each  
33 meta-analytic grouping (MAG) of experiments. We then considered how such networks subdivide  
34 information processing by assessing the characteristics of the constituent experiments from each MAG.  
35 Utilizing both manual and automated functional decoding approaches, enhanced interpretations of the  
36 mental processes associated with specific constellations of brain regions were gleaned such that the

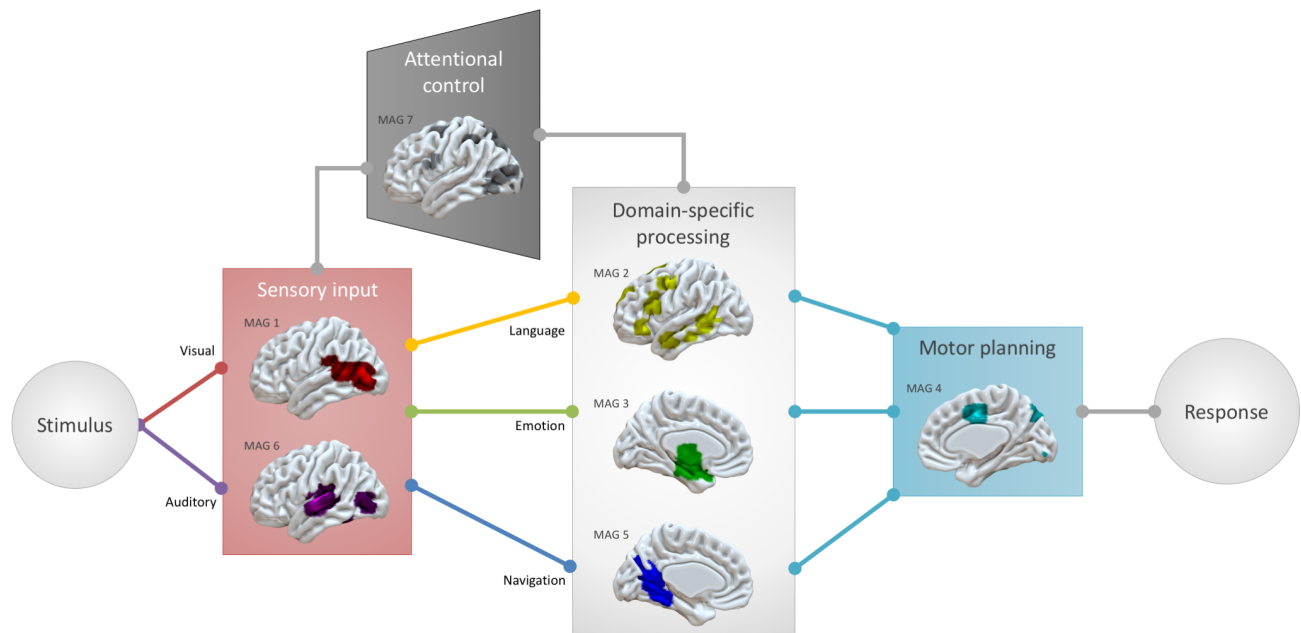
1 outcomes of the two approaches generally agreed, with differences highlighting domain-specific and  
2 domain-general processes associated with naturalistic paradigms.

### 3 ***Distributed Processing for Complex Functions***

4 Though the seven identified MAGs are spatially distinct and appear to correspond with dissociable mental  
5 processes, most of the included naturalistic tasks recruited more than one MAG (71 of 110). This is  
6 consistent with functional segregation and the flexible nature of the naturalistic design, demonstrating  
7 that the manipulation of different contrasts can identify distinct networks that likely cooperate to  
8 successfully perform a complex task. Further indicative of coordinated interactions and distributed  
9 processing, each MAG included experiments that utilized different task modalities and task types.  
10 Overwhelmingly, the identified MAGs and the functional characterizations thereof support the notion that  
11 complex behaviors are facilitated by coordinated interactions between several large-scale sensory, motor,  
12 attentional, and domain-specific networks, a position increasingly endorsed in neuroimaging endeavors  
13 (Barrett & Satpute, 2013; Lindquist, Wager, Kober, Bliss-Moreau, & Barrett, 2012; Mišić & Sporns, 2016;  
14 Spreng, Sepulcre, Turner, Stevens, & Schacter, 2013). The characterization of identified MAGs from  
15 aspects of the naturalistic paradigms that elicit them suggest an information processing model of  
16 cooperating systems (Figure 4) for sensory input (MAGs 1 and 6), attentional control (MAG 7), domain-  
17 specific processing (MAGs 2, 3, and 5), and motor planning (MAG 4), into and from which information is  
18 segregated and integrated to enable complex behaviors (e.g., language, emotion, spatial navigation).

19 MAGs 1 and 6 primarily represent the perceptual processing streams of incoming auditory and visual  
20 information, and likely cooperate to process audiovisual information. Functional decoding suggests that  
21 MAG 1 is involved in viewing faces and anthropomorphic figures, which is consistent with previous  
22 research showing that posterior temporal and temporo-occipital regions corresponding with area V5/MT  
23 are associated with the perception of movement, specifically biological movement (Cohen Kadosh,

1 Henson, Cohen Kadosh, Johnson, & Dick, 2010; Pelphrey, Morris, & McCarthy, 2004; Pelphrey, Morris,  
2 Michelich, Allison, & McCarthy, 2005; Puce, Allison, Bentin, Gore, & McCarthy, 1998; Wheaton,  
3 Thompson, Syngeniotis, Abbott, & Puce, 2004). Similarly, MAG 6 is associated with listening to music and  
4 speech, as well as perceiving pitch and tone, stretching across primary auditory cortex and into regions of  
5 higher auditory processing (Gray, Carver, & Standring, 2009; Türe, Yaşargil, Al-Mefty, & Yaşargil, 1999).  
6 Per functional decoding of MAG 7 of both manual and automated annotations, MAG 7 is associated with  
7 visual attention. This functional characterization is also supported by the corresponding fronto-parietal  
8 activations that are often associated with attending to both auditory and visual stimuli (Braga, Fu,  
9 Seemungal, Wise, & Leech, 2016; Puschmann, Huster, & Thiel, 2016), a necessary process for successful  
10 language processing. MAGs 1 and 6 represent the perceptual processing streams of audiovisual  
11 information.



12

13 **Figure 4. Complex systems for dynamical information processing.** The identified MAGs present a  
14 framework of component systems that interact to enable complex information processing needed for  
15 naturalistic behavior, including necessary input and output systems, as well as systems for attentional  
16 gating of irrelevant information and domain-specific processing for language-, emotion-, and navigation-  
17 related tasks



1 Information processing depends on input from perceptual systems, filtered by attentional gating, by  
2 proceeds in a functionally-segregated manner, seen in domain-specific MAGs for linguistic, emotional,  
3 and spatial processing. When considering language processing, there is necessary input from primary  
4 auditory areas (MAG 6) and attentional selection (MAG 7) that enables speech perception and  
5 comprehension (MAG 2). More than a third of contrasts from experiments that utilized speech-based  
6 paradigms contributed to the convergent activation pattern of MAG 2, which was linked by both  
7 functional decoding techniques to language-related processes. Furthermore, the regions of MAG 2  
8 resembles a neural “language network” (Friederici & Gierhan, 2013; Heim, Opitz, Müller, & Friederici,  
9 2003; Price, 2010; Saur et al., 2010), including some regions associated with orofacial articulation (lip,  
10 tongue, and jaw movements) and motor planning (SMA, pre-SMA) that allow the motor components of  
11 speech. By presenting language in a context that is more representative of how we process language in  
12 everyday life, such as through the use of spoken fictional narratives (AbdulSabur et al., 2014; Wallentin et  
13 al., 2011; Xu, Kemeny, Park, Frattali, & Braun, 2005a) or scene descriptions (Summerfield, Hassabis, &  
14 Maguire, 2010), naturalistic fMRI paradigms allow researchers to explore the multiple neural networks at  
15 work in performing the cooperating processes that facilitate language processing. Similarly, emotional  
16 processing (MAG 3) often necessitates audiovisual input (MAGs 1 and 6) and observing human forms  
17 (MAGs 1, 4). Emotional films recruited regions across these four MAGs, suggesting a similarly diverse  
18 group of coordinated neural systems are engaged when observing affective displays. Additionally,  
19 navigation (Burgess, Maguire, & O’Keefe, 2002; Kalpouzos, Eriksson, Sjölie, Molin, & Nyberg, 2010;  
20 Wolbers, Weiller, & Büchel, 2004) depends on effective attentional selection (MAG 7), as well as spatial  
21 memory and processing (MAG 5). The functional characterization of MAG 5 from manual and Neurosynth  
22 decoding highlights its involvement in navigation and spatial memory, supported by studies of rats and  
23 humans with brain lesions that indicate the importance of medial temporal, hippocampal, and precuneus  
24 regions in processing visual scenes and spatial information (Bird & Burgess, 2008; Epstein, 2008; Lee et

1 al., 2005; Sailer et al., 2000; Squire, Stark, & Clark, 2004; Summerfield et al., 2010; Xu, Kemeny, Park,  
2 Frattali, & Braun, 2005b).

3 Finally, interaction with dynamic stimuli of any type requires a system that facilitates integration of  
4 relevant information for behavioral selection (MAG 4). Both manual and Neurosynth decoding indicate a  
5 role for this MAG in action execution, as well as in observing, imagining, and executing motor movements.  
6 Across the corpus, experiments predicated on navigating virtual environments, viewing emotional or  
7 erotic films, playing video games, and listening to speech all engaged MAG 4 and paradigms that required  
8 a motor response disproportionately engaged MAGs 4 and 7 above the other MAGs. Regions that  
9 comprise MAG 4 have been associated with the mirror neuron system, response outcome monitoring,  
10 and higher-level visual processing, supporting an integrative system for motor planning (Apps, Lockwood,  
11 & Balsters, 2013; Gazzola & Keysers, 2009; Molenberghs, Cunnington, & Mattingley, 2009).

## 12 **Limitations**

13 The present results may be limited by the *k*-means clustering method, which requires that parameters  
14 including distance metric, number of iterations, and number of clusters are specified beforehand.  
15 Experiments in our corpus were grouped using the *kmeans++* algorithm for each of  $K = 2$  through  $K = 20$   
16 solutions, repeated 1000 times to ensure that each solution minimized the point-to-centroid distance,  
17 indicative of optimal clustering (Kanungo et al., 2004). Pearson's correlation was selected as the distance  
18 metric, as recommended by Laird et al. (2015). The  $K = 7$  solution was designated as an optimal candidate  
19 solution before assessing the convergent activation patterns of each MAG, based on the aforementioned  
20 metrics, yielding a data-driven result. Although the functional decoding based manual annotations relied  
21 on a subjective process, the results were largely confirmed by comparison with the wider body of  
22 functional neuroimaging literature facilitated by Neurosynth's automated functional decoding. It is worth  
23 noting that the naturalistic literature is somewhat limited, with an emphasis on navigation and affective

1 processing, and continued research and expansion of this corpus will facilitate development of a more  
2 comprehensive model of the neural networks that support realistic behavior.

### 3 ***Summary and Future Work***

4 In summary, this meta-analysis of naturalistic fMRI studies that apply dynamic, lifelike tasks to explore the  
5 neural correlates of behavior has shown that these paradigms engage a set of core neural networks,  
6 supporting both separate processing of different streams of information and the integration of related  
7 information to enable flexible cognition and complex behavior. We identified seven patterns of consistent  
8 activation that correspond with neural networks that are involved in sensory input, top-down attentional  
9 control, domain-specific processing, and motor planning, representing the set of behavioral processes  
10 elicited by naturalistic paradigms in our corpus. Across the corpus, tasks provided mainly visual and  
11 auditory sensory input which engaged regions across MAGs 1 and 6, while MAG 7 appeared to contribute  
12 to top-down attentional control to filter out nonessential information from those sensory streams. Salient  
13 information can be processed by the relevant domain-specific networks, shown in MAGs 2 (language), 3  
14 (emotion), and 5 (navigation and spatial memory), and the appropriate motor response, coordinated by  
15 regions in MAG 4. The majority of tasks engaged multiple networks to process the relevant information  
16 from a stimulus and generate an appropriate response. A greater literature of fMRI experiments utilizing  
17 naturalistic paradigms would benefit the field and complement the existing literature of precisely  
18 controlled tasks, as naturalistic stimuli more closely approximate the processing necessary for realistic  
19 behavior. Exploring how multifaceted processes interact and, ultimately, contribute to behavior will allow  
20 us to better understand the brain and human behavior in the real world. In the future, studies of this sort  
21 would greatly benefit from an automated annotation process for an objective functional decoding of  
22 included papers, instead of subjective manual annotation.

23

1 **ACKNOWLEDGMENTS**

2 This study was supported by awards from the National Institute of Drug Abuse (U01-DA041156, K01-  
3 DA037819, U24-DA039832, R01DA041353), the National Institute of Mental Health (R56-MH097870), and  
4 the National Science Foundation (1631325 and REAL DRL-1420627). The authors declare no competing  
5 financial interests.

6

1

## References

- 2 AbdulSabur, N. Y., Xu, Y., Liu, S., Chow, H. M., Baxter, M., Carson, J., & Braun, A. R. (2014). Neural  
3 correlates and network connectivity underlying narrative production and comprehension: A  
4 combined fMRI and PET study. *Cortex*, *57*, 107–127. <http://doi.org/10.1016/j.cortex.2014.01.017>
- 5 Apps, M. A. J., Lockwood, P. L., & Balsters, J. H. (2013). The role of the midcingulate cortex in monitoring  
6 others' decisions. *Frontiers in Neuroscience*, *7*, 251. <http://doi.org/10.3389/fnins.2013.00251>
- 7 Barrett, L. F., & Satpute, A. B. (2013). Large-scale brain networks in affective and social neuroscience:  
8 towards an integrative functional architecture of the brain. *Current Opinion in Neurobiology*, *23*(3),  
9 361–72. <http://doi.org/10.1016/j.conb.2012.12.012>
- 10 Bird, C. M., & Burgess, N. (2008). The hippocampus and memory: insights from spatial processing.  
11 *Nature Reviews. Neuroscience*, *9*(3), 182–194. <http://doi.org/10.1038/nrn2335>
- 12 Braga, R. M., Fu, R. Z., Seemungal, B. M., Wise, R. J. S., & Leech, R. (2016). Eye Movements during  
13 Auditory Attention Predict Individual Differences in Dorsal Attention Network Activity. *Frontiers in*  
14 *Human Neuroscience*, *10*, 164. <http://doi.org/10.3389/fnhum.2016.00164>
- 15 Burgess, N., Maguire, E. A., & O'Keefe, J. (2002). The human hippocampus and spatial and episodic  
16 memory. *Neuron*. [http://doi.org/10.1016/S0896-6273\(02\)00830-9](http://doi.org/10.1016/S0896-6273(02)00830-9)
- 17 Bzdok, D., Heeger, A., Langner, R., Laird, A. R., Fox, P. T., Palomero-Gallagher, N., ... Eickhoff, S. B. (2015).  
18 Subspecialization in the human posterior medial cortex. *NeuroImage*, *106*, 55–71.  
19 <http://doi.org/10.1016/j.neuroimage.2014.11.009>
- 20 Cohen Kadosh, K., Henson, R. N. a, Cohen Kadosh, R., Johnson, M. H., & Dick, F. (2010). Task-dependent  
21 activation of face-sensitive cortex: an fMRI adaptation study. *Journal of Cognitive Neuroscience*,  
22 *22*(5), 903–917. <http://doi.org/10.1162/jocn.2009.21224>
- 23 Collins, D. L., Neelin, P., Peters, T. M., & Evans, A. C. (1994). Automatic 3D intersubject registration of MR  
24 volumetric data in standardized Talairach space. *Journal of Computer Assisted Tomography*, *18*(2),  
25 192–205. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8126267>
- 26 Eickhoff, S. B., Bzdok, D., Laird, A. R., Kurth, F., & Fox, P. T. (2012). Activation likelihood estimation meta-  
27 analysis revisited. *NeuroImage*, *59*(3), 2349–2361.  
28 <http://doi.org/10.1016/j.neuroimage.2011.09.017>
- 29 Eickhoff, S. B., Laird, A. R., Fox, P. T., Bzdok, D., & Hensel, L. (2016). Functional Segregation of the Human  
30 Dorsomedial Prefrontal Cortex. *Cerebral Cortex*, *26*(1), 304–321.  
31 <http://doi.org/10.1093/cercor/bhu250>
- 32 Eickhoff, S. B., Laird, A. R., Grefkes, C., Wang, L. E., Zilles, K., & Fox, P. T. (2009). Coordinate-based  
33 activation likelihood estimation meta-analysis of neuroimaging data: A random-effects approach  
34 based on empirical estimates of spatial uncertainty. *Human Brain Mapping*, *30*(9), 2907–2926.  
35 <http://doi.org/10.1002/hbm.20718>
- 36 Eickhoff, S. B., Nichols, T. E., Laird, A. R., Hoffstaedter, F., Amunts, K., Fox, P. T., ... Eickhoff, C. R. (2016).  
37 Behavior, sensitivity, and power of activation likelihood estimation characterized by massive  
38 empirical simulation. *NeuroImage*, *137*, 70–85. <http://doi.org/10.1016/j.neuroimage.2016.04.072>
- 39 Epstein, R. A. (2008). Parahippocampal and retrosplenial contributions to human spatial navigation.  
40 *Trends in Cognitive Sciences*. <http://doi.org/10.1016/j.tics.2008.07.004>

- 1 Evans, A. C., Collins, D. L., Mills, S. R., Brown, E. D., Kelly, R. L., & Peters, T. M. (1993). 3D statistical  
2 neuroanatomical models from 305 MRI volumes. In *1993 IEEE Conference Record Nuclear Science  
3 Symposium and Medical Imaging Conference* (pp. 1813–1817). IEEE.  
4 <http://doi.org/10.1109/NSSMIC.1993.373602>
- 5 Fox, P. T., Laird, A. R., Fox, S. P., Fox, P. M., Uecker, A. M., Crank, M., ... Lancaster, J. L. (2005). BrainMap  
6 taxonomy of experimental design: description and evaluation. *Human Brain Mapping*, *25*(1), 185–  
7 198. <http://doi.org/10.1002/hbm.20141>
- 8 Friederici, A. D., & Gierhan, S. M. (2013). The language network. *Current Opinion in Neurobiology*, *23*(2),  
9 250–254. <http://doi.org/10.1016/j.conb.2012.10.002>
- 10 Friston, K. J., Price, C. J., Fletcher, P., Moore, C., Frackowiak, R. S., & Dolan, R. J. (1996). The trouble with  
11 cognitive subtraction. *NeuroImage*, *4*(2), 97–104. <http://doi.org/10.1006/nimg.1996.0033>
- 12 Gazzola, V., & Keysers, C. (2009). The observation and execution of actions share motor and  
13 somatosensory voxels in all tested subjects: single-subject analyses of unsmoothed fMRI data.  
14 *Cerebral Cortex (New York, N.Y. : 1991)*, *19*(6), 1239–55. <http://doi.org/10.1093/cercor/bhn181>
- 15 Giard, M. H., & Peronnet, F. (1999). Auditory-visual integration during multimodal object recognition in  
16 humans: a behavioral and electrophysiological study. *J Cogn Neurosci*, *11*(5), 473–490.  
17 <http://doi.org/10.1162/089892999563544>
- 18 Gorgolewski, K. J., Varoquaux, G., Rivera, G., Schwarz, Y., Ghosh, S. S., Maumet, C., ... Margulies, D. S.  
19 (2015). NeuroVault.org: a web-based repository for collecting and sharing unthresholded statistical  
20 maps of the human brain. *Frontiers in Neuroinformatics*, *9*, 8.  
21 <http://doi.org/10.3389/fninf.2015.00008>
- 22 Gray, H., Carver, H. V., & Standring, S. (2009). Gray's Anatomy 40th edition. *Churchill Livingstone*.  
23 <http://doi.org/10.1308/003588406X116873>
- 24 Hasson, U., & Honey, C. J. (2012). Future trends in Neuroimaging: Neural processes as expressed within  
25 real-life contexts. *NeuroImage*, *62*(2), 1272–1278.  
26 <http://doi.org/10.1016/j.neuroimage.2012.02.004>
- 27 Heim, S., Opitz, B., Müller, K., & Friederici, A. D. (2003). Phonological processing during language  
28 production: fMRI evidence for a shared production-comprehension network. *Brain Research.  
29 Cognitive Brain Research*, *16*(2), 285–96. Retrieved from  
30 <http://www.ncbi.nlm.nih.gov/pubmed/12668238>
- 31 Kalpouzos, G., Eriksson, J., Sjölie, D., Molin, J., & Nyberg, L. (2010). Neurocognitive Systems Related to  
32 Real-World Prospective Memory. *PLoS ONE*, *5*(10), e13304.  
33 <http://doi.org/10.1371/journal.pone.0013304>
- 34 Kanungo, T., Mount, D. M., Netanyahu, N. S., Piatko, C. D., Silverman, R., & Wu, A. Y. (2004). A local  
35 search approximation algorithm for k-means clustering. *Computational Geometry*, *28*(2–3), 89–  
36 112. <http://doi.org/10.1016/j.comgeo.2004.03.003>
- 37 Laird, A. R., Fox, P. M., Price, C. J., Glahn, D. C., Uecker, A. M., Lancaster, J. L., ... Fox, P. T. (2005). ALE  
38 meta-analysis: Controlling the false discovery rate and performing statistical contrasts. *Human  
39 Brain Mapping*, *25*(1), 155–164. <http://doi.org/10.1002/hbm.20136>
- 40 Laird, A. R., Riedel, M. C., Sutherland, M. T., Eickhoff, S. B., Ray, K. L., Uecker, A. M., ... Fox, P. T. (2015).  
41 Neural architecture underlying classification of face perception paradigms. *NeuroImage*, *119*, 70–  
42 80. <http://doi.org/10.1016/j.neuroimage.2015.06.044>

- 1 Laird, A. R., Robinson, J. L., McMillan, K. M., Tordesillas-Gutiérrez, D., Moran, S. T., Gonzales, S. M., ...  
2 Lancaster, J. L. (2010). Comparison of the disparity between Talairach and MNI coordinates in  
3 functional neuroimaging data: validation of the Lancaster transform. *NeuroImage*, *51*(2), 677–83.  
4 <http://doi.org/10.1016/j.neuroimage.2010.02.048>
- 5 Lancaster, J. L., Tordesillas-Gutiérrez, D., Martinez, M., Salinas, F., Evans, A., Zilles, K., ... Fox, P. T. (2007).  
6 Bias between MNI and Talairach coordinates analyzed using the ICBM-152 brain template. *Human*  
7 *Brain Mapping*, *28*(11), 1194–205. <http://doi.org/10.1002/hbm.20345>
- 8 Lee, A. C. H., Buckley, M. J., Pegman, S. J., Spiers, H., Scahill, V. L., Gaffan, D., ... Graham, K. S. (2005).  
9 Specialization in the medial temporal lobe for processing of objects and scenes. *Hippocampus*,  
10 *15*(6), 782–797. <http://doi.org/10.1002/hipo.20101>
- 11 Lindquist, K. A., Wager, T. D., Kober, H., Bliss-Moreau, E., & Barrett, L. F. (2012). The brain basis of  
12 emotion: A meta-analytic review. *Behavioral and Brain Sciences*, *35*(3), 121–143.  
13 <http://doi.org/10.1017/S0140525X11000446>
- 14 Maguire, E. A. (2012). Studying the freely-behaving brain with fMRI. *NeuroImage*, *62*(2), 1170–1176.  
15 <http://doi.org/10.1016/j.neuroimage.2012.01.009>
- 16 McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*, 691–811.  
17 <http://doi.org/10.1038/264746a0>
- 18 Mišić, B., & Sporns, O. (2016). From regions to connections and networks: new bridges between brain  
19 and behavior. *Current Opinion in Neurobiology*, *40*, 1–7.  
20 <http://doi.org/10.1016/j.conb.2016.05.003>
- 21 Molenberghs, P., Cunnington, R., & Mattingley, J. B. (2009). Is the mirror neuron system involved in  
22 imitation? A short review and meta-analysis. *Neuroscience & Biobehavioral Reviews*, *33*(7), 975–  
23 980. <http://doi.org/10.1016/j.neubiorev.2009.03.010>
- 24 Pelphrey, K. A., Morris, J. P., & McCarthy, G. (2004). Grasping the intentions of others: the perceived  
25 intentionality of an action influences activity in the superior temporal sulcus during social  
26 perception. *Journal of Cognitive Neuroscience*, *16*(10), 1706–1716.  
27 <http://doi.org/10.1162/0898929042947900>
- 28 Pelphrey, K. A., Morris, J. P., Michelich, C. R., Allison, T., & McCarthy, G. (2005). Functional anatomy of  
29 biological motion perception in posterior temporal cortex: An fMRI study of eye, mouth and hand  
30 movements. *Cerebral Cortex*, *15*(12), 1866–1876. <http://doi.org/10.1093/cercor/bhi064>
- 31 Price, C. J. (2010). The anatomy of language: a review of 100 fMRI studies published in 2009. *Annals of*  
32 *the New York Academy of Sciences*, *1191*(1), 62–88. [http://doi.org/10.1111/j.1749-](http://doi.org/10.1111/j.1749-6632.2010.05444.x)  
33 [6632.2010.05444.x](http://doi.org/10.1111/j.1749-6632.2010.05444.x)
- 34 Puce, a, Allison, T., Bentin, S., Gore, J. C., & McCarthy, G. (1998). Temporal cortex activation in humans  
35 viewing eye and mouth movements. *The Journal of Neuroscience : The Official Journal of the*  
36 *Society for Neuroscience*, *18*(6), 2188–2199.
- 37 Puschmann, S., Huster, R. J., & Thiel, C. M. (2016). Mapping the spatiotemporal dynamics of processing  
38 task-relevant and task-irrelevant sound feature changes using concurrent EEG-fMRI. *Human Brain*  
39 *Mapping*. <http://doi.org/10.1002/hbm.23248>
- 40 Sailer, U., Eggert, T., Ditterich, J., & Straube, A. (2000). Spatial and temporal aspects of eye-hand  
41 coordination across different tasks. *Experimental Brain Research*, *134*(2), 163–173.  
42 <http://doi.org/10.1007/s002210000457>

- 1 Saur, D., Schelter, B., Schnell, S., Kratochvil, D., Küpper, H., Kellmeyer, P., ... Weiller, C. (2010). Combining  
2 functional and anatomical connectivity reveals brain networks for auditory language  
3 comprehension. *NeuroImage*, 49(4), 3187–3197.  
4 <http://doi.org/10.1016/j.neuroimage.2009.11.009>
- 5 Spence, C. (2010). Crossmodal spatial attention. *Annals of the New York Academy of Sciences*.  
6 <http://doi.org/10.1111/j.1749-6632.2010.05440.x>
- 7 Spreng, R. N., Sepulcre, J., Turner, G. R., Stevens, W. D., & Schacter, D. L. (2013). Intrinsic architecture  
8 underlying the relations among the default, dorsal attention, and frontoparietal control networks  
9 of the human brain. *Journal of Cognitive Neuroscience*, 25(1), 74–86.  
10 [http://doi.org/10.1162/jocn\\_a\\_00281](http://doi.org/10.1162/jocn_a_00281)
- 11 Squire, L. R., Stark, C. E. L., & Clark, R. E. (2004). The medial temporal lobe. *Annual Review of*  
12 *Neuroscience*, 27, 279–306. <http://doi.org/10.1146/annurev.neuro.27.070203.144130>
- 13 Summerfield, J. J., Hassabis, D., & Maguire, E. A. (2010). Differential engagement of brain regions within  
14 a “core” network during scene construction. *Neuropsychologia*, 48(5), 1501–9.  
15 <http://doi.org/10.1016/j.neuropsychologia.2010.01.022>
- 16 Talairach, J., & Tournoux, P. (1988). *Co-Planar Stereotaxic Atlas of the Human Brain: 3-D Proportional*  
17 *System: An Approach to Cerebral Imaging* (1st Editio). Stuttgart ; New York: Thieme. Retrieved  
18 from <http://www.amazon.com/Co-Planar-Stereotaxic-Atlas-Human-Brain/dp/0865772932>
- 19 Türe, U., Yaşargil, D. C. H., Al-Mefty, O., & Yaşargil, M. G. (1999). Topographic anatomy of the insular  
20 region. *Journal of Neurosurgery*, 90(4), 720–733. <http://doi.org/10.3171/jns.1999.90.4.0720>
- 21 Turkeltaub, P. E., Eden, G. F., Jones, K. M., & Zeffiro, T. A. (2002). Meta-Analysis of the Functional  
22 Neuroanatomy of Single-Word Reading: Method and Validation. *NeuroImage*, 16(3, Part A), 765–  
23 780. <http://doi.org/10.1006/nimg.2002.1131>
- 24 Turkeltaub, P. E., Eickhoff, S. B., Laird, A. R., Fox, M., Wiener, M., & Fox, P. (2012). Minimizing within-  
25 experiment and within-group effects in activation likelihood estimation meta-analyses. *Human*  
26 *Brain Mapping*, 33(1), 1–13. <http://doi.org/10.1002/hbm.21186>
- 27 Turner, J. A., & Laird, A. R. (2012). The cognitive paradigm ontology: design and application.  
28 *Neuroinformatics*, 10(1), 57–66. <http://doi.org/10.1007/s12021-011-9126-x>
- 29 Wallentin, M., Nielsen, A. H., Vuust, P., Dohn, A., Roepstorff, A., & Lund, T. E. (2011). Amygdala and  
30 heart rate variability responses from listening to emotionally intense parts of a story. *NeuroImage*,  
31 58(3), 963–973. <http://doi.org/10.1016/j.neuroimage.2011.06.077>
- 32 Wang, J., Ren, Y., Hu, X., Nguyen, V. T., Guo, L., Han, J., & Guo, C. C. (2016). Test-retest reliability of  
33 functional connectivity networks during naturalistic fMRI paradigms. *bioRxiv*.
- 34 Wheaton, K. J., Thompson, J. C., Syngeniotis, A., Abbott, D. F., & Puce, A. (2004). Viewing the motion of  
35 human body parts activates different regions of premotor, temporal, and parietal cortex.  
36 *NeuroImage*, 22(1), 277–288. <http://doi.org/10.1016/j.neuroimage.2003.12.043>
- 37 Wolbers, T., Weiller, C., & Büchel, C. (2004). Neural foundations of emerging route knowledge in  
38 complex spatial environments. *Cognitive Brain Research*, 21(3), 401–411.  
39 <http://doi.org/10.1016/j.cogbrainres.2004.06.013>
- 40 Woo, C.-W., Krishnan, A., & Wager, T. D. (2014). Cluster-extent based thresholding in fMRI analyses:  
41 pitfalls and recommendations. *NeuroImage*, 91, 412–9.  
42 <http://doi.org/10.1016/j.neuroimage.2013.12.058>



- 1 Xu, J., Kemeny, S., Park, G., Frattali, C., & Braun, A. (2005a). Language in context: Emergent features of  
2 word, sentence, and narrative comprehension. *NeuroImage*, 25(3), 1002–1015.  
3 <http://doi.org/10.1016/j.neuroimage.2004.12.013>
- 4 Xu, J., Kemeny, S., Park, G., Frattali, C., & Braun, A. (2005b). Language in context: emergent features of  
5 word, sentence, and narrative comprehension. *NeuroImage*, 25(3), 1002–15.  
6 <http://doi.org/10.1016/j.neuroimage.2004.12.013>
- 7 Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., & Wager, T. D. (2011). Large-scale automated  
8 synthesis of human functional neuroimaging data. *Nature Methods*, 8(8), 665–70.  
9 <http://doi.org/10.1038/nmeth.1635>
- 10  
11