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[1], Jessica S. Flannery[1], Emily R. Boeving[1], Michael C. Riedel[2], Simon B. Eickhoff[3,4], Matthew T. Sutherland[1], Angela R. Laird[2]

[1]Department of Psychology, Florida International University, Miami, FL
[2]Department of Physics, Florida International University, Miami, FL
[3]Institute of Clinical Neuroscience and Medical Psychology, Heinrich-Heine University, Düsseldorf, Germany
[4]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

Abstract

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

Introduction

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

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

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

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

Materials and Methods

Naturalistic fMRI Paradigms

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

Literature Search, Filtering, and Annotation

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

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

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

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

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

Experimental Design and Statistical Analysis

Modeled Activation Maps

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

K-Means Clustering Analysis

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

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

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

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

Meta-Analytic Groupings

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

Functional Decoding

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

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

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

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

Results

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

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

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

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

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

K-Means Clustering Solutions

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

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

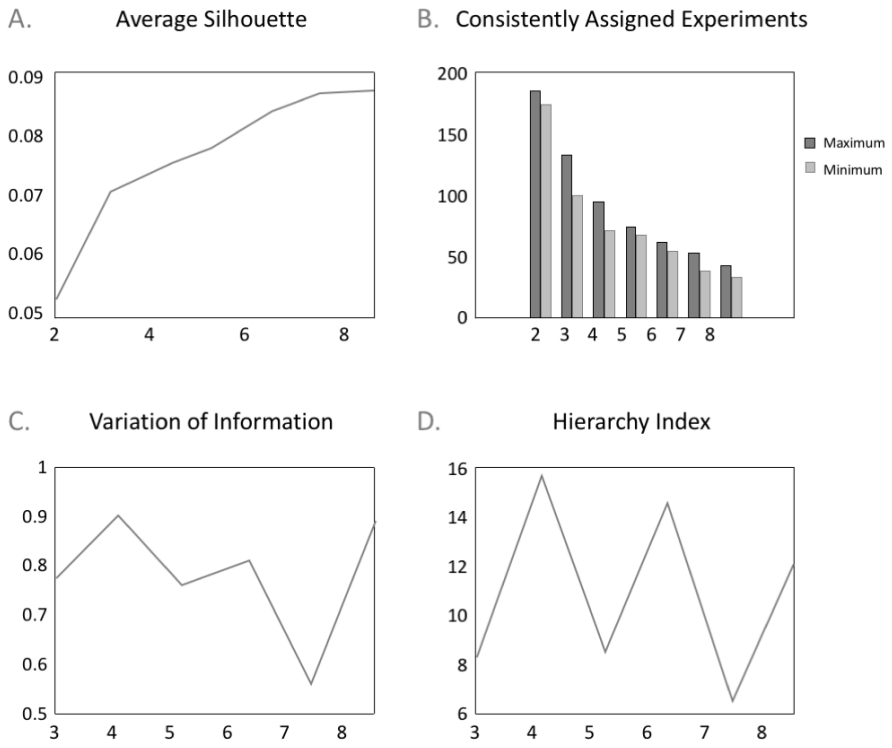

Figure 1. Metrics computed for K = 2 – 10 clustering solutions. (A) The average cluster silhouette for each solution K from 2 to 10 clusters. (B) Consistency in experiments assignment to clusters, plotting the minimum consistently assigned clusters next to the mean of consistently assigned clusters.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

stemmed from clusters in the *K* – 1 solution.

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

Meta-Analytic Groupings

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

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

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

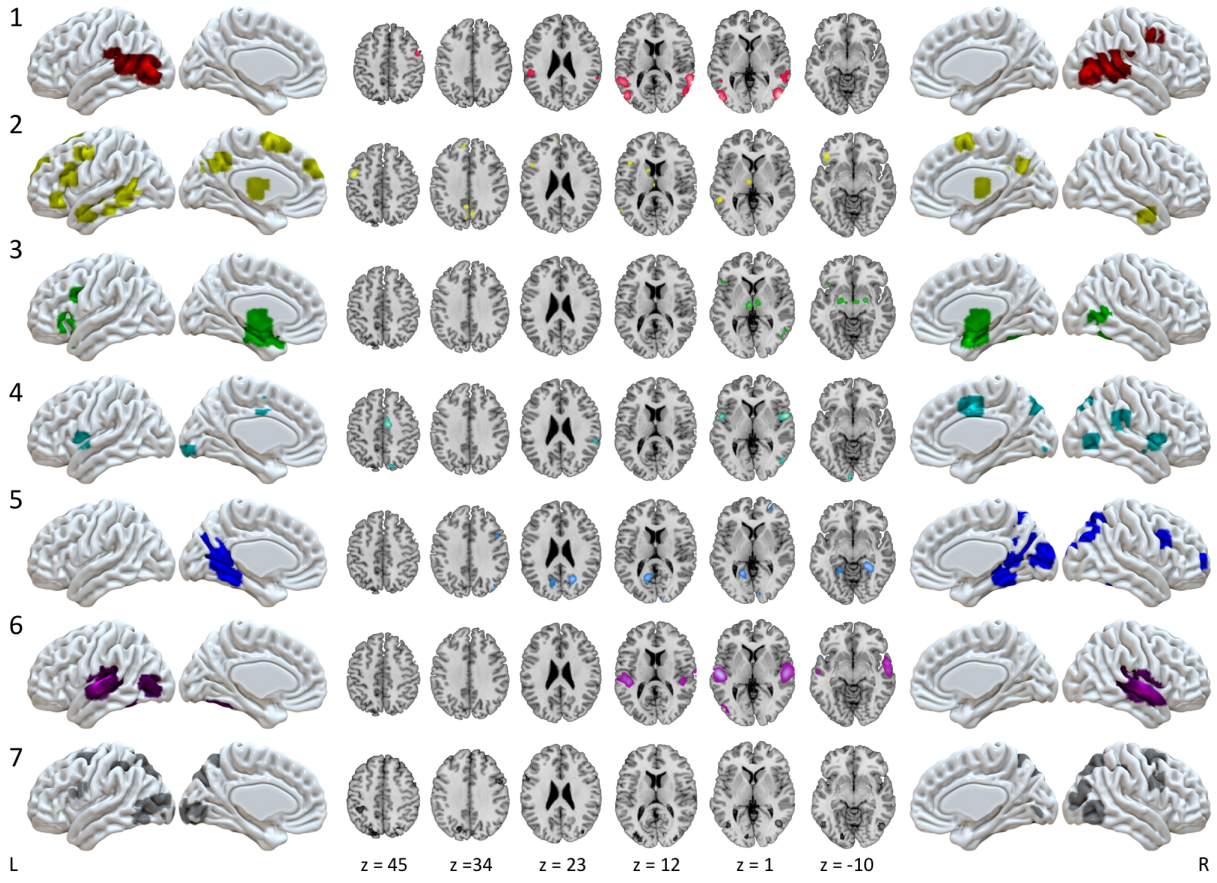

Figure 2. Convergent activation patterns of MAGs from the naturalistic corpus. ALE meta-analysis of experiments in each MAG yielded seven patterns of convergent activation.

Stimulus Distribution Across MAGs

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

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

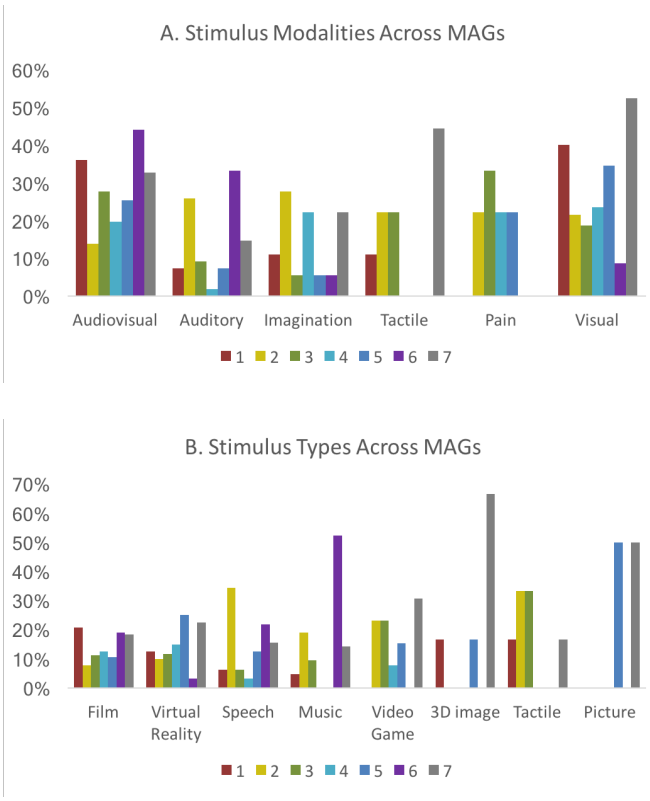

Figure 3. Distribution of stimulus modalities and types across MAGs. (A) The presence of each sensory modality across the corpus that is associated with each MAG. (B) The proportion of each stimulus type present within the corpus that is associated with each MAG. These percentages represent the proportion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

Functional Decoding

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

Manual Annotations

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

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

Automated Neurosynth Annotations

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

Term	Total		Frequency per MAG													
			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%)

Table 4. Manual functional decoding results across meta-analytic groupings. The relative contributions of each manually-derived metadata term (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 count across the corpus. Base-rates are provided as the total count for each term.

Functional Interpretation of MAGs

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

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

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

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

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

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

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

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

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

Discussion

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

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

Distributed Processing for Complex Functions

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

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

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

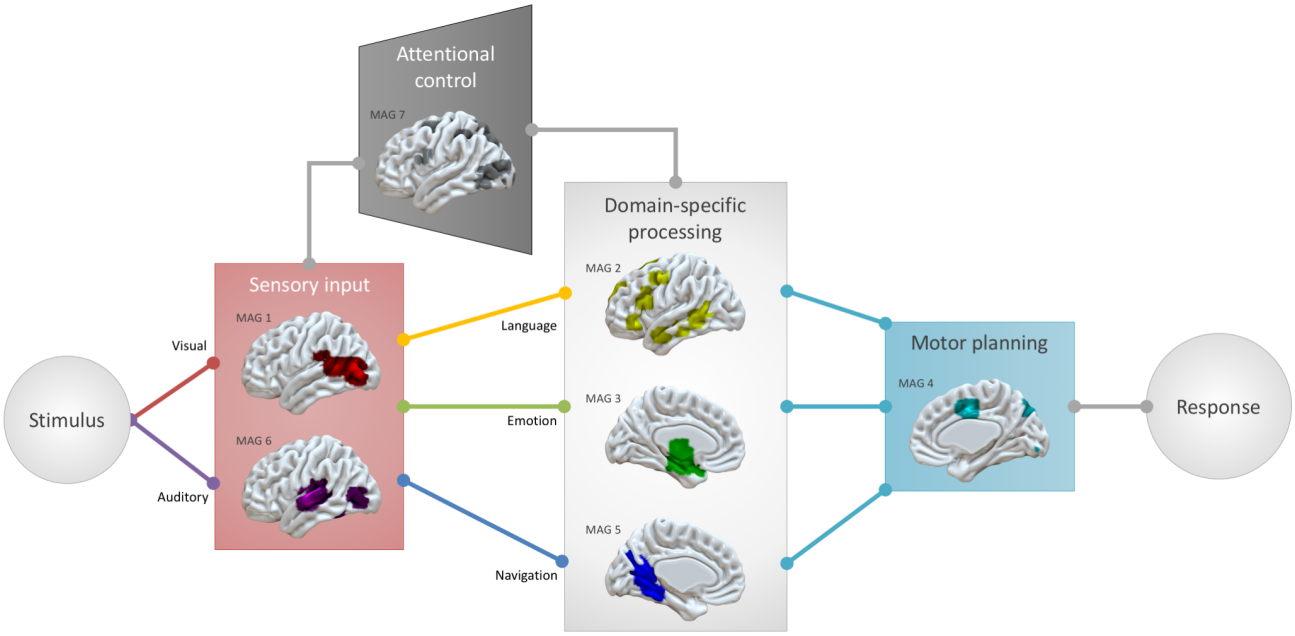

Figure 4. Complex systems for dynamical information processing. The identified MAGs present a framework of component systems that interact to enable complex information processing needed for naturalistic behavior, including necessary input and output systems, as well as systems for attentional gating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

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

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

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

Limitations

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

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

Summary and Future Work

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

ACKNOWLEDGMENTS

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

References

AbdulSabur, N. Y., Xu, Y., Liu, S., Chow, H. M., Baxter, M., Carson, J., & Braun, A. R. (2014). Neural correlates and network connectivity underlying narrative production and comprehension: A combined fMRI and PET study. *Cortex*, *57*, 107–127. http://doi.org/10.1016/j.cortex.2014.01.017

Apps, M. A. J., Lockwood, P. L., & Balsters, J. H. (2013). The role of the midcingulate cortex in monitoring others' decisions. *Frontiers in Neuroscience*, *7*, 251. http://doi.org/10.3389/fnins.2013.00251

Barrett, L. F., & Satpute, A. B. (2013). Large-scale brain networks in affective and social neuroscience: towards an integrative functional architecture of the brain. *Current Opinion in Neurobiology*, *23*(3), 361–72. http://doi.org/10.1016/j.conb.2012.12.012

Bird, C. M., & Burgess, N. (2008). The hippocampus and memory: insights from spatial processing. *Nature Reviews. Neuroscience*, *9*(3), 182–194. http://doi.org/10.1038/nrn2335

Braga, R. M., Fu, R. Z., Seemungal, B. M., Wise, R. J. S., & Leech, R. (2016). Eye Movements during Auditory Attention Predict Individual Differences in Dorsal Attention Network Activity. *Frontiers in Human Neuroscience*, *10*, 164. http://doi.org/10.3389/fnhum.2016.00164

Burgess, N., Maguire, E. A., & O'Keefe, J. (2002). The human hippocampus and spatial and episodic memory. *Neuron*. http://doi.org/10.1016/S0896-6273(02)00830-9

Bzdok, D., Heeger, A., Langner, R., Laird, A. R., Fox, P. T., Palomero-Gallagher, N., … Eickhoff, S. B. (2015). Subspecialization in the human posterior medial cortex. *NeuroImage*, *106*, 55–71. http://doi.org/10.1016/j.neuroimage.2014.11.009

Cohen Kadosh, K., Henson, R. N. a, Cohen Kadosh, R., Johnson, M. H., & Dick, F. (2010). Task-dependent activation of face-sensitive cortex: an fMRI adaptation study. *Journal of Cognitive Neuroscience*, *22*(5), 903–917. http://doi.org/10.1162/jocn.2009.21224

Collins, D. L., Neelin, P., Peters, T. M., & Evans, A. C. (1994). Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. *Journal of Computer Assisted Tomography*, *18*(2), 192–205. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8126267

Eickhoff, S. B., Bzdok, D., Laird, A. R., Kurth, F., & Fox, P. T. (2012). Activation likelihood estimation meta-analysis revisited. *NeuroImage*, *59*(3), 2349–2361. http://doi.org/10.1016/j.neuroimage.2011.09.017

Eickhoff, S. B., Laird, A. R., Fox, P. T., Bzdok, D., & Hensel, L. (2016). Functional Segregation of the Human Dorsomedial Prefrontal Cortex. *Cerebral Cortex*, *26*(1), 304–321. http://doi.org/10.1093/cercor/bhu250

Eickhoff, S. B., Laird, A. R., Grefkes, C., Wang, L. E., Zilles, K., & Fox, P. T. (2009). Coordinate-based activation likelihood estimation meta-analysis of neuroimaging data: A random-effects approach based on empirical estimates of spatial uncertainty. *Human Brain Mapping*, *30*(9), 2907–2926. http://doi.org/10.1002/hbm.20718

Eickhoff, S. B., Nichols, T. E., Laird, A. R., Hoffstaedter, F., Amunts, K., Fox, P. T., … Eickhoff, C. R. (2016). Behavior, sensitivity, and power of activation likelihood estimation characterized by massive empirical simulation. *NeuroImage*, *137*, 70–85. http://doi.org/10.1016/j.neuroimage.2016.04.072

Epstein, R. A. (2008). Parahippocampal and retrosplenial contributions to human spatial navigation. *Trends in Cognitive Sciences*. http://doi.org/10.1016/j.tics.2008.07.004

Evans, A. C., Collins, D. L., Mills, S. R., Brown, E. D., Kelly, R. L., & Peters, T. M. (1993). 3D statistical neuroanatomical models from 305 MRI volumes. In *1993 IEEE Conference Record Nuclear Science Symposium and Medical Imaging Conference* (pp. 1813–1817). IEEE. http://doi.org/10.1109/NSSMIC.1993.373602

Fox, P. T., Laird, A. R., Fox, S. P., Fox, P. M., Uecker, A. M., Crank, M., … Lancaster, J. L. (2005). BrainMap taxonomy of experimental design: description and evaluation. *Human Brain Mapping*, *25*(1), 185–198. http://doi.org/10.1002/hbm.20141

Friederici, A. D., & Gierhan, S. M. (2013). The language network. *Current Opinion in Neurobiology*, *23*(2), 250–254. http://doi.org/10.1016/j.conb.2012.10.002

Friston, K. J., Price, C. J., Fletcher, P., Moore, C., Frackowiak, R. S., & Dolan, R. J. (1996). The trouble with cognitive subtraction. *NeuroImage*, *4*(2), 97–104. http://doi.org/10.1006/nimg.1996.0033

Gazzola, V., & Keysers, C. (2009). The observation and execution of actions share motor and somatosensory voxels in all tested subjects: single-subject analyses of unsmoothed fMRI data. *Cerebral Cortex (New York, N.Y. : 1991)*, *19*(6), 1239–55. http://doi.org/10.1093/cercor/bhn181

Giard, M. H., & Peronnet, F. (1999). Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. *J Cogn Neurosci*, *11*(5), 473–490. http://doi.org/10.1162/089892999563544

Gorgolewski, K. J., Varoquaux, G., Rivera, G., Schwarz, Y., Ghosh, S. S., Maumet, C., … Margulies, D. S. (2015). NeuroVault.org: a web-based repository for collecting and sharing unthresholded statistical maps of the human brain. *Frontiers in Neuroinformatics*, *9*, 8. http://doi.org/10.3389/fninf.2015.00008

Gray, H., Carver, H. V, & Standring, S. (2009). Gray's Anatomy 40th edition. *Churchill Livingstone*. http://doi.org/10.1308/003588406X116873

Hasson, U., & Honey, C. J. (2012). Future trends in Neuroimaging: Neural processes as expressed within real-life contexts. *NeuroImage*, *62*(2), 1272–1278. http://doi.org/10.1016/j.neuroimage.2012.02.004

Heim, S., Opitz, B., Müller, K., & Friederici, A. D. (2003). Phonological processing during language production: fMRI evidence for a shared production-comprehension network. *Brain Research. Cognitive Brain Research*, *16*(2), 285–96. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12668238

Kalpouzos, G., Eriksson, J., Sjölie, D., Molin, J., & Nyberg, L. (2010). Neurocognitive Systems Related to Real-World Prospective Memory. *PLoS ONE*, *5*(10), e13304. http://doi.org/10.1371/journal.pone.0013304

Kanungo, T., Mount, D. M., Netanyahu, N. S., Piatko, C. D., Silverman, R., & Wu, A. Y. (2004). A local search approximation algorithm for k-means clustering. *Computational Geometry*, *28*(2–3), 89–112. http://doi.org/10.1016/j.comgeo.2004.03.003

Laird, A. R., Fox, P. M., Price, C. J., Glahn, D. C., Uecker, A. M., Lancaster, J. L., … Fox, P. T. (2005). ALE meta-analysis: Controlling the false discovery rate and performing statistical contrasts. *Human Brain Mapping*, *25*(1), 155–164. http://doi.org/10.1002/hbm.20136

Laird, A. R., Riedel, M. C., Sutherland, M. T., Eickhoff, S. B., Ray, K. L., Uecker, A. M., … Fox, P. T. (2015). Neural architecture underlying classification of face perception paradigms. *NeuroImage*, *119*, 70–80. http://doi.org/10.1016/j.neuroimage.2015.06.044

Laird, A. R., Robinson, J. L., McMillan, K. M., Tordesillas-Gutiérrez, D., Moran, S. T., Gonzales, S. M., … Lancaster, J. L. (2010). Comparison of the disparity between Talairach and MNI coordinates in functional neuroimaging data: validation of the Lancaster transform. *NeuroImage*, *51*(2), 677–83. http://doi.org/10.1016/j.neuroimage.2010.02.048

Lancaster, J. L., Tordesillas-Gutiérrez, D., Martinez, M., Salinas, F., Evans, A., Zilles, K., … Fox, P. T. (2007). Bias between MNI and Talairach coordinates analyzed using the ICBM-152 brain template. *Human Brain Mapping*, *28*(11), 1194–205. http://doi.org/10.1002/hbm.20345

Lee, A. C. H., Buckley, M. J., Pegman, S. J., Spiers, H., Scahill, V. L., Gaffan, D., … Graham, K. S. (2005). Specialization in the medial temporal lobe for processing of objects and scenes. *Hippocampus*, *15*(6), 782–797. http://doi.org/10.1002/hipo.20101

Lindquist, K. A., Wager, T. D., Kober, H., Bliss-Moreau, E., & Barrett, L. F. (2012). The brain basis of emotion: A meta-analytic review. *Behavioral and Brain Sciences*, *35*(3), 121–143. http://doi.org/10.1017/S0140525X11000446

Maguire, E. A. (2012). Studying the freely-behaving brain with fMRI. *NeuroImage*, *62*(2), 1170–1176. http://doi.org/10.1016/j.neuroimage.2012.01.009

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*, 691–811. http://doi.org/10.1038/264746a0

Mišić, B., & Sporns, O. (2016). From regions to connections and networks: new bridges between brain and behavior. *Current Opinion in Neurobiology*, *40*, 1–7. http://doi.org/10.1016/j.conb.2016.05.003

Molenberghs, P., Cunnington, R., & Mattingley, J. B. (2009). Is the mirror neuron system involved in imitation? A short review and meta-analysis. *Neuroscience & Biobehavioral Reviews*, *33*(7), 975–980. http://doi.org/10.1016/j.neubiorev.2009.03.010

Pelphrey, K. A., Morris, J. P., & McCarthy, G. (2004). Grasping the intentions of others: the perceived intentionality of an action influences activity in the superior temporal sulcus during social perception. *Journal of Cognitive Neuroscience*, *16*(10), 1706–1716. http://doi.org/10.1162/0898929042947900

Pelphrey, K. A., Morris, J. P., Michelich, C. R., Allison, T., & McCarthy, G. (2005). Functional anatomy of biological motion perception in posterior temporal cortex: An fMRI study of eye, mouth and hand movements. *Cerebral Cortex*, *15*(12), 1866–1876. http://doi.org/10.1093/cercor/bhi064

Price, C. J. (2010). The anatomy of language: a review of 100 fMRI studies published in 2009. *Annals of the New York Academy of Sciences*, *1191*(1), 62–88. http://doi.org/10.1111/j.1749-6632.2010.05444.x

Puce, a, Allison, T., Bentin, S., Gore, J. C., & McCarthy, G. (1998). Temporal cortex activation in humans viewing eye and mouth movements. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *18*(6), 2188–2199.

Puschmann, S., Huster, R. J., & Thiel, C. M. (2016). Mapping the spatiotemporal dynamics of processing task-relevant and task-irrelevant sound feature changes using concurrent EEG-fMRI. *Human Brain Mapping*. http://doi.org/10.1002/hbm.23248

Sailer, U., Eggert, T., Ditterich, J., & Straube, A. (2000). Spatial and temporal aspects of eye-hand coordination across different tasks. *Experimental Brain Research*, *134*(2), 163–173. http://doi.org/10.1007/s002210000457

Saur, D., Schelter, B., Schnell, S., Kratochvil, D., Küpper, H., Kellmeyer, P., … Weiller, C. (2010). Combining functional and anatomical connectivity reveals brain networks for auditory language comprehension. *NeuroImage*, *49*(4), 3187–3197. http://doi.org/10.1016/j.neuroimage.2009.11.009

Spence, C. (2010). Crossmodal spatial attention. *Annals of the New York Academy of Sciences*. http://doi.org/10.1111/j.1749-6632.2010.05440.x

Spreng, R. N., Sepulcre, J., Turner, G. R., Stevens, W. D., & Schacter, D. L. (2013). Intrinsic architecture underlying the relations among the default, dorsal attention, and frontoparietal control networks of the human brain. *Journal of Cognitive Neuroscience*, *25*(1), 74–86. http://doi.org/10.1162/jocn_a_00281

Squire, L. R., Stark, C. E. L., & Clark, R. E. (2004). The medial temporal lobe. *Annual Review of Neuroscience*, *27*, 279–306. http://doi.org/10.1146/annurev.neuro.27.070203.144130

Summerfield, J. J., Hassabis, D., & Maguire, E. A. (2010). Differential engagement of brain regions within a “core” network during scene construction. *Neuropsychologia*, *48*(5), 1501–9. http://doi.org/10.1016/j.neuropsychologia.2010.01.022

Talairach, J., & Tournoux, P. (1988). *Co-Planar Stereotaxic Atlas of the Human Brain: 3-D Proportional System: An Approach to Cerebral Imaging* (1st Editio). Stuttgart ; New York: Thieme. Retrieved from http://www.amazon.com/Co-Planar-Stereotaxic-Atlas-Human-Brain/dp/0865772932

Türe, U., Yaşargil, D. C. H., Al-Mefty, O., & Yaşargil, M. G. (1999). Topographic anatomy of the insular region. *Journal of Neurosurgery*, *90*(4), 720–733. http://doi.org/10.3171/jns.1999.90.4.0720

Turkeltaub, P. E., Eden, G. F., Jones, K. M., & Zeffiro, T. A. (2002). Meta-Analysis of the Functional Neuroanatomy of Single-Word Reading: Method and Validation. *NeuroImage*, *16*(3, Part A), 765–780. http://doi.org/10.1006/nimg.2002.1131

Turkeltaub, P. E., Eickhoff, S. B., Laird, A. R., Fox, M., Wiener, M., & Fox, P. (2012). Minimizing within-experiment and within-group effects in activation likelihood estimation meta-analyses. *Human Brain Mapping*, *33*(1), 1–13. http://doi.org/10.1002/hbm.21186

Turner, J. A., & Laird, A. R. (2012). The cognitive paradigm ontology: design and application. *Neuroinformatics*, *10*(1), 57–66. http://doi.org/10.1007/s12021-011-9126-x

Wallentin, M., Nielsen, A. H., Vuust, P., Dohn, A., Roepstorff, A., & Lund, T. E. (2011). Amygdala and heart rate variability responses from listening to emotionally intense parts of a story. *NeuroImage*, *58*(3), 963–973. http://doi.org/10.1016/j.neuroimage.2011.06.077

Wang, J., Ren, Y., Hu, X., Nguyen, V. T., Guo, L., Han, J., & Guo, C. C. (2016). Test-retest reliability of functional connectivity networks during naturalistic fMRI paradigms. *bioRxiv*.

Wheaton, K. J., Thompson, J. C., Syngeniotis, A., Abbott, D. F., & Puce, A. (2004). Viewing the motion of human body parts activates different regions of premotor, temporal, and parietal cortex. *NeuroImage*, *22*(1), 277–288. http://doi.org/10.1016/j.neuroimage.2003.12.043

Wolbers, T., Weiller, C., & Büchel, C. (2004). Neural foundations of emerging route knowledge in complex spatial environments. *Cognitive Brain Research*, *21*(3), 401–411. http://doi.org/10.1016/j.cogbrainres.2004.06.013

Woo, C.-W., Krishnan, A., & Wager, T. D. (2014). Cluster-extent based thresholding in fMRI analyses: pitfalls and recommendations. *NeuroImage*, *91*, 412–9. http://doi.org/10.1016/j.neuroimage.2013.12.058

Xu, J., Kemeny, S., Park, G., Frattali, C., & Braun, A. (2005a). Language in context: Emergent features of word, sentence, and narrative comprehension. *NeuroImage*, *25*(3), 1002–1015. http://doi.org/10.1016/j.neuroimage.2004.12.013

Xu, J., Kemeny, S., Park, G., Frattali, C., & Braun, A. (2005b). Language in context: emergent features of word, sentence, and narrative comprehension. *NeuroImage*, *25*(3), 1002–15. http://doi.org/10.1016/j.neuroimage.2004.12.013

Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., & Wager, T. D. (2011). Large-scale automated synthesis of human functional neuroimaging data. *Nature Methods*, *8*(8), 665–70. http://doi.org/10.1038/nmeth.1635