

Contextual Alignment of Cognitive and Neural Dynamics

Daniel L. Ames¹, Christopher J. Honey², Michael A. Chow¹,
Alexander Todorov¹, and Uri Hasson¹

Abstract

Effective real-world communication requires the alignment of multiple individuals to a common perspective or mental framework. To study how this alignment occurs at the level of the brain, we measured BOLD response during fMRI while participants ($n = 24$) listened to a series of vignettes either in the presence or absence of a valid contextual cue. The valid contextual cue was necessary to understand the information in each vignette. We then examined where and to what extent the shared valid context led to greater intersubject similarity

of neural processing. Regions of the default mode network including posterior cingulate cortex and medial pFC became more aligned when participants shared a valid contextual framework, whereas other regions, including primary sensory cortices, responded to the stimuli reliably regardless of contextual factors. Taken in conjunction with previous research, the present results suggest that default mode regions help the brain to organize incoming verbal information in the context of previous knowledge. ■

INTRODUCTION

One of the brain's primary functions is to organize perception and behavior in the context of previous information. Context occurs at multiple levels. At the perceptual level, preceding visual contexts make object recognition significantly more efficient (Bar, 2004). At a cognitive level, many studies on lexical priming demonstrate that the brain's response to any given word is highly dependent on the preceding linguistic context (Rissman, Eliassen, & Blumstein, 2003; Mummery, Shallice, & Price, 1999).

In recent years the neural processes underlying such contextual phenomena have become increasingly well defined. However, there are also broader, knowledge-based contextual factors that, although less well understood, are nonetheless critical to the brain's ability to organize information in an adaptive manner. Although "knowledge-based context" is somewhat difficult to define, examples are easily identified. For instance, in a typical conversation, the conversing parties call on contextual representations (e.g., shared knowledge of people, events and locations) as their discourse unfolds over time. Without the ability to integrate incoming information in light of existing contextual models, normal human discourse would not be possible.

Bransford and Johnson (1972) introduced a classic paradigm for studying such knowledge-based contextual effects. In this paradigm, participants read or hear paragraphs that are difficult to comprehend without knowledge of the paragraph's topic (e.g., doing laundry). These contextual cues are typically comprised of no more than a few

words or a briefly presented image; yet they reliably produce dramatic improvements in comprehension and memory (Johnson, Doll, Bransford, & Lapinski, 1974; Bransford & Johnson, 1972).

Although these behavioral effects are robust, their neural substrates remain poorly understood. Three previous experiments have attempted to address this gap by adapting Bransford and Johnson's paradigm for use with neuroimaging (Martín-Loeches, Casado, Hernández-Tamames, & Álvarez-Linera, 2008; Maguire, Frith, & Morris, 1999; St. George, Kutas, Martinez, & Sereno, 1999). In spite of their similar designs, these analyses identified areas of activation that were small, dispersed, and inconsistent across studies, with the only apparent overlap between any two studies (Martín-Loeches et al., 2008; Maguire et al., 1999) arising in a circumscribed region of posterior cingulate cortex (PCC).

One reason why these experiments may have produced inconsistent findings is that they relied on event-related averaging. Although this standard approach is extremely useful and has facilitated many important discoveries in cognitive neuroscience, it may not be appropriate for analyzing the knowledge-based parsing of naturalistic text (Ben-Yakov, Honey, Lerner, & Hasson, 2012; Smirnov et al., 2012). One reason for this is that modeling lengthy, naturalistic text fragments as 30–90 sec blocks fails to capture the dynamic changes in natural language that tend to occur within such periods of time. In contrast, the intersubject correlation (ISC) analysis method captures this variability by using each participant's brain responses as a model to predict other participants' brain responses to the same stimulus (thus circumventing the need to specify an a priori model; Hasson, Malach, & Heeger, 2010;

¹Princeton University, ²University of Toronto

Hasson, Nir, Levy, Fuhrmann, & Malach, 2004). Moreover, at a conceptual level, we suggest that context is fundamentally about aligning multiple individuals to a common perspective or mental model (i.e., “common ground” see e.g., Hagoort & van Berkum, 2007). The ISC method examines this alignment directly by indexing the extent to which individuals’ brains respond similarly to one another within a given context (see Smirnov et al., 2012, for similar approach).

Prior work suggests that core elements of the default mode network (DMN; Raichle et al., 2001; Shulman et al., 1997) are well suited to integrating information in relation to prior knowledge-based context (Lerner, Honey, Silbert, & Hasson, 2011; Hasson et al., 2009). In these studies, we measured the influence of past information on the moment-to-moment processes in each brain area. Specifically, we parametrically varied the temporal structure of real-life audio stories and audio-visual movies by breaking them into smaller and smaller temporal segments (e.g., paragraphs, sentences, words) and then scrambling the order of the segments. Next, we asked whether BOLD responses to each segment (i.e., each paragraph, sentence, or word) changed as a function of prior events. ISC analyses revealed that areas with short processing timescales (e.g., auditory and visual cortices) responded in the same way to each event regardless of prior context. In contrast, areas with long processing timescales, including medial pFC (mPFC), PCC, and bilateral TPJ, modulated their responses to a given event as a function of earlier information. In other words, what happened in a story 1 min ago (i.e., prior context) reliably influences how the DMN responds to incoming information, whereas primary sensory cortices show no such sensitivity.

The ability to retain and use prior knowledge to process incoming information is critical to the context-based processes being studied here (i.e., using information from the initial visual cue to help guide processing of the subsequently spoken text). In addition, because the application of knowledge-based context requires combining previous information with present input (e.g., Ericsson & Kintsch, 1995), one would expect that regions subserving such integration would have strong anatomical and functional connections with medial-temporal memory systems (Ferstl, Neumann, Bogler, & Von Cramon, 2008). Such connectivity has been observed for core nodes of the DMN (Ranganath & Ritchey, 2012; Buckner, Andrews-Hanna, & Schacter, 2008). We therefore predicted that the response time courses in PCC, mPFC, and TPJ would be more reliably aligned across participants when they processed information following a valid, rather than invalid, contextual cue.

METHODS

Participants

Twenty-four right-handed participants, ages 18–32 years, listened to 12 vignettes during fMRI scanning. All participants

had normal or corrected-to-normal vision. Participants provided informed consent. All experimental procedures were approved by the Princeton University internal review panel.

Experimental Design and Stimuli

All participants listened to 12 vignettes ($M = 85.5$ sec, $SD = 11.3$ sec), which were difficult to understand in the absence of contextual background.

Below we provide an example of one such vignette:

Kicking or stomping is usually required, both at the beginning and at the end. You should know relatively quickly if you need to do it, as things will feel all tilted and sluggish. Millions of people get professional help with this every year-but most people could deal with it on their own. If you have this problem, you almost certainly have everything you need to fix it. You need to get the right amount of elevation: too little and you won't be able to get the job done, but too much elevation can be dangerous, as everything can come crashing down if you're not careful. Be sure to loosen it evenly from different sides. It can be a little tricky to get everything aligned appropriately when you go to make the replacement. It's easier to complete the procedure in some places than in others, but the fact of the matter is that you don't usually have too much choice about where you do it. Be extra careful to make sure that everything is tight when you finish-the consequences of not doing so could be disastrous.

Pretesting¹ confirmed that few people were able to make sense of this text on its own; however, when the vignette was preceded by a picture of a person changing a tire, its meaning became clear. Participants viewed one picture immediately before each vignette. Half of these pictures provided a valid contextual cue that framed the vignette. In this valid context condition, comprehension required the online interpretation of present information in light of knowledge-based contextual information provided by the preceding visual cue. The remaining pictures were stylistically similar non sequiturs, which were presented only to control for the presence of an image. Thus, in this invalid context condition, participants were left to interpret each vignette without the benefit of informative context. Participants did not know in advance which of the pictures would be relevant or irrelevant to the forthcoming vignette; however, participants were informed that half of the pictures would be irrelevant, and the stimuli were designed such that the relevance or irrelevance of each picture became immediately obvious once the vignette began. Any given vignette was always paired with the same non sequitur image for one half of the participants and with the same valid image for the remaining half

of participants—that is, the consistency of image–vignette pairings was equivalent across conditions.

Each picture was presented for 4 sec, beginning 10 sec before onset of speech. A fixation cross was presented during the intervening 6 sec and for an additional 6 sec after each vignette (Figure 1A). Assignment of each of the 12 vignettes to the valid context or invalid context condition was counterbalanced across participants. Thus, all vignettes (and all participants) contributed equally to the two conditions. To maximize comparability across conditions, vignette order was held constant. Listeners were instructed to simply listen to the vignettes and made no behavioral responses during scanning.

Behavioral Tests

After scanning, participants completed two comprehensibility ratings for each vignette (Bransford & Johnson, 1972). The first rating asked participants to indicate “how well [they were] able to understand the topic and the statements made in each scenario.” Participants indicated their responses by making a mark on a 10.8-cm line, anchored with “I was totally confused” and “it was all totally clear.” For the second comprehensibility item, participants were asked to guess the topic of the vignette and indicate their confidence in this guess (again using

a 10.8-cm line, anchored with “I’m guessing randomly” and “I’m totally certain”). We then measured the distance of each mark from the line’s zero point (leftmost end) to calculate participant comprehension. The two measures were highly correlated (average $r = .90$); we therefore collapsed them into a single mean comprehension index (Figure 1B).

Data Acquisition

Participants were scanned using a 3T head-only MRI scanner (Allegra; Siemens). A custom radiofrequency head coil was used to achieve high-resolution structural scans (NM-011 transmit head coil; Nova Medical). Eight hundred seventy functional volumes were acquired across four runs using a T2*-weighted EPI pulse sequence (repetition time = 1500 msec, echo time = 30 msec, flip angle = 76°; field of view = 192 mm², right-to-left phase encoding). Each volume included 25 interleaved slices of 3 mm thickness (1 mm gap; in-plane resolution: 3 × 3 mm²) for near-whole-brain coverage. An anatomical scan was then acquired using a T1-weighted high-resolution (1 mm³) MP-RAGE pulse sequence. Stimuli were presented using MATLAB (MathWorks; Natick, MA) and Psychophysics Toolbox and were delivered via a scanner-safe projector and high-fidelity MR-compatible headphones (MR Confon;

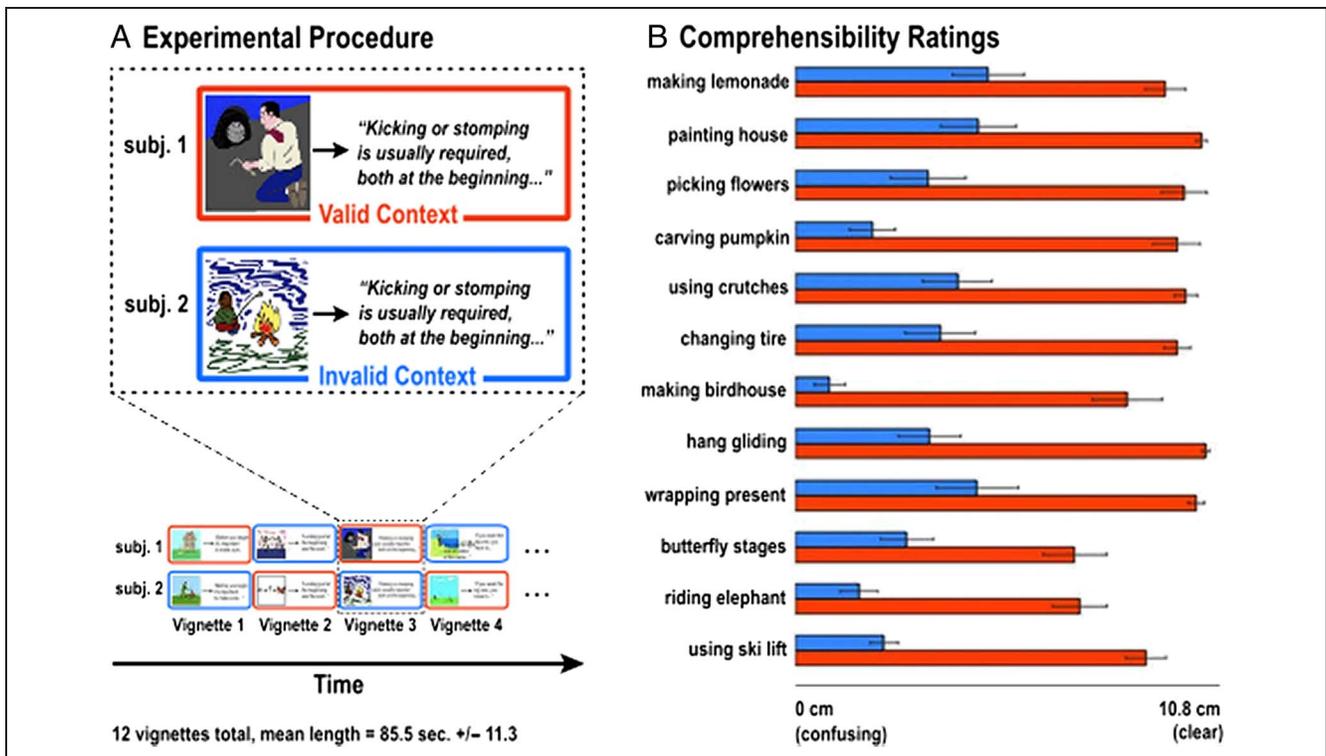
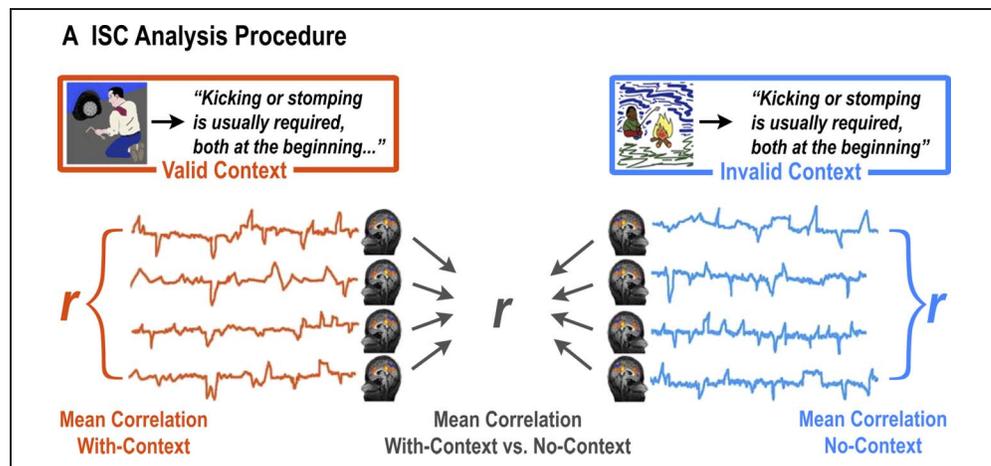


Figure 1. Behavioral procedure and comprehensibility ratings. (A) Participants listened to a total of 12 vignettes. Half of these vignettes were preceded by a picture that provided a valid contextual framework for the vignette (valid context condition); the remaining half were preceded by a nonsequitur picture (invalid context condition). Vignette order was held constant, and experimental condition was counterbalanced across participants. (B) Context improves comprehension. All vignettes were better understood when preceded by a pictorial cue that signaled a valid contextual framework (orange bars), rather than a nonsequitur (blue bars), all p s < .001.

Figure 2. Schematic representation of ISC procedure. BOLD response time courses were extracted from each voxel of a given participant's brain. Each voxel was then correlated with the average of the corresponding voxel in all other participants' brains (within the same condition). This approach indexes the extent to which the same neural response time course is observed reliably across individuals. These response reliabilities were examined separately for the valid context condition (orange) and the invalid context condition (blue). For comparison, we also correlated each time course with the average time courses in the opposite condition (gray).



Magdeburg, Germany). These headphones function optimally in the bore of the scanner and reduce acoustic scanner noise.

fMRI Preprocessing

fMRI data were preprocessed and aligned using version 1.8.6 of the BrainVoyager QX software package (Brain Innovation; Maastricht, The Netherlands). Preprocessing of functional scans included 3-D motion correction, linear trend removal, slice scan-time correction, and high-pass filtering (frequencies below three cycles per functional run removed). A Gaussian spatial filter of 6-mm width at half-maximum value was applied to correct for structural heterogeneity between brains, which were aligned to standard Talairach coordinates. Voxels with low mean BOLD values (>4 standard deviations below the gray and white matter mean) were excluded from analysis.

Time courses were trimmed such that only volumes acquired during the auditory vignettes themselves were included in analysis (i.e., BOLD responses to the pictures were excluded). Therefore, all stimuli served as their own controls, and the only difference across the two conditions was the knowledge-based context that participants brought to bear on the vignettes. To remove transients at the time of stimulus onset (which would artificially inflate ISCs), we discarded the first eight functional volumes (12 sec) from each vignette. Visual inspection of the data confirmed that this procedure removed all apparent transients.

ISC Analysis

ISC analysis was used to index the extent to which a given neural region responded similarly across individuals.

BOLD response time courses for each voxel in a given participant's brain were correlated with the corresponding voxel in all other brains over the duration of each vignette (see Figure 2). In comparison to standard general linear model (GLM) analyses, which typically assume a canonical hemodynamic response function, the ISC method makes no a priori assumptions concerning the specific timing of BOLD responses to each vignette. Rather, the ISC approach asks to what extent the same response time course is reliably observed across all participants. We compared these reliabilities in the valid context condition and in the invalid context condition. In regions where all participants' responses are similar (correlated) regardless of the presence or absence of a valid contextual cue, one may surmise that the responses are locked to the processing of the stimulus, but in a context-independent way. However, in regions where all participants' responses are similar (correlated) only during the valid context condition, one may conclude that the reliability of processing in these areas is context dependent. This approach accounts for the fact that BOLD response time courses can be specific to a given brain region and also to a particular vignette.

Time courses were cropped as described above to remove picture cues and fixations. They were then normalized (standardized to zero mean and unit variance) to control for mean signal differences across participants, and reconcatenated such that each time course consisted of all 12 cropped vignettes, in a constant order, either in the valid context condition or the invalid context condition. Correlation maps were then calculated separately for the valid context and invalid context conditions. With data mapped into a common Talairach space, the time course of each voxel in each brain was correlated with the average time course of all other same-condition brains at that voxel. These correlations, one per participant, were

then averaged into a mean r value. The resulting maps (Figure 3) index the functional similarity of participants' brains with and without a valid context. For comparison's sake, we also calculated the correlation between each voxel in a given brain and the average of spatially corresponding voxels in brains from the opposite condition (Figure 2). For brain regions that are insensitive to knowledge-based context (e.g., because they process low-level properties of the stimuli), identical reliabilities would be expected in all three conditions (valid context, invalid context, and cross-condition comparison).

Controlling for False Positives

To control for false positives, we employed a family-wise error rate correction procedure that was calibrated to reject chance correlations with 99.99% confidence. The empirically observed correlations (valid context and invalid context ISCs) were compared against a null distribution of chance correlations. These chance correlations were calculated from phase-randomized bootstrapped data. For each condition, a phase randomization of each voxel's time course was carried out via inverted Fourier transformation (which scrambles the signal's phase, while leaving its power spectrum intact). For each voxel, a distribution of 10,000 ISCs was generated using precisely the same procedures employed in the analyses described above—that is, by calculating a Pearson correlation coefficient between each voxel's BOLD time course in one participant's brain and the mean of that voxel's BOLD time course in the remaining participants' brains. We selected the largest "chance" ISC value from the null distribution of all voxels in each iteration, repeating this bootstrapping procedure 10,000 times to generate a null distribution of maximum noise correlation values. This distribution of chance correlations was approximately Gaussian. We therefore compared the empirically observed correlations against this null distribution to compute p values. The family-wise error rate (FWER)

was defined as the top 0.01% of the null distribution of the maximum correlation values exceeding a given threshold ($r = .08$), which was used to threshold the veridical map shown in Figure 3 (Nichols and Holmes, 2002).

Between-condition Differences

We employed a nonparametric bootstrapping procedure to directly compare the reliability of brain responses in the valid versus invalid context conditions. In the case of ISC, bootstrapping procedures provide a more principled test of between-condition differences than does a t test-based contrast (because comparing each participant's response time course to the average responses of all other participants may preserve some temporal dependencies). To take into account temporal dependencies in the data (Bullmore et al., 2001), we employed the circular block bootstrap procedure (Politis & Romano, 1992). Specifically, blocks of a fixed length were resampled with replacement from each BOLD time course to create a bootstrapped sample containing the same number of time points as the original data. Because ISC considers every time point to be a multivariate observation consisting of each participant's BOLD response at that time point, the same blocks of time were resampled for each participant. For each bootstrapped sample, the statistic of interest—the difference between mean within-group ISC for the two conditions—was calculated in the exact same manner as described above for the within-group ISC calculations. This bootstrap procedure was repeated 10,000 times to obtain an approximate distribution for the mean within-group ISC difference. Finally, this distribution was tested against the null hypothesis that the difference is not positive by calculating p values as the proportion of bootstrapped differences that were equal to or below zero (Figure 4).

Block lengths were determined via an algorithm developed by Politis and White (2004; see correction in Patton, Politis, & White, 2009), which takes into account the

Figure 3. Reliability of brain responses within each condition. The BOLD time course of each voxel was correlated across listeners to produce maps of ISC. Separate surface maps show the areas exhibiting statistically reliable responses for each condition. A = anterior; P = posterior; A1+ = primary auditory cortex; L = left; SMG = supramarginal gyrus; AG = angular gyrus. FWER-corrected, $p < .0001$.

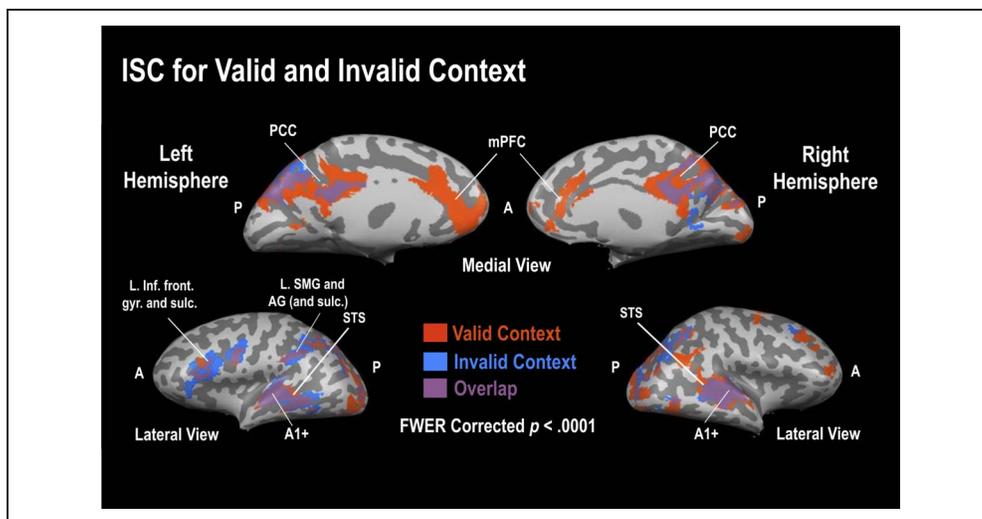
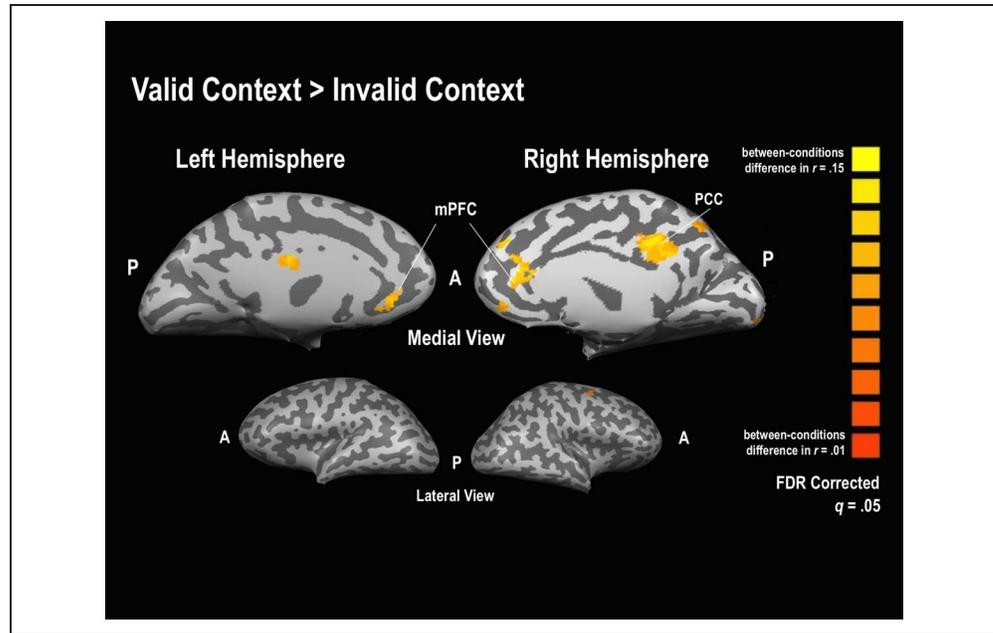


Figure 4. Regions differentially sensitive to context. Surface map shows areas in which reliability is significantly greater (FDR-corrected, $q = .05$) when participants share a valid contextual framework compared with when participants lack such a framework. A = anterior; P = posterior.



autocorrelation structure of a given time course. For each participant, the median of each block length suggested by the algorithm was taken across all voxels. Next, the mean of the previous result was taken across participants (rounded to the nearest integer), leading to a final block length set at seven repetition times. Correction for multiple comparisons was done using the false discovery rate (FDR) procedure given by Benjamini and Hochberg (1995), with a q value of .05.

ROI Analyses

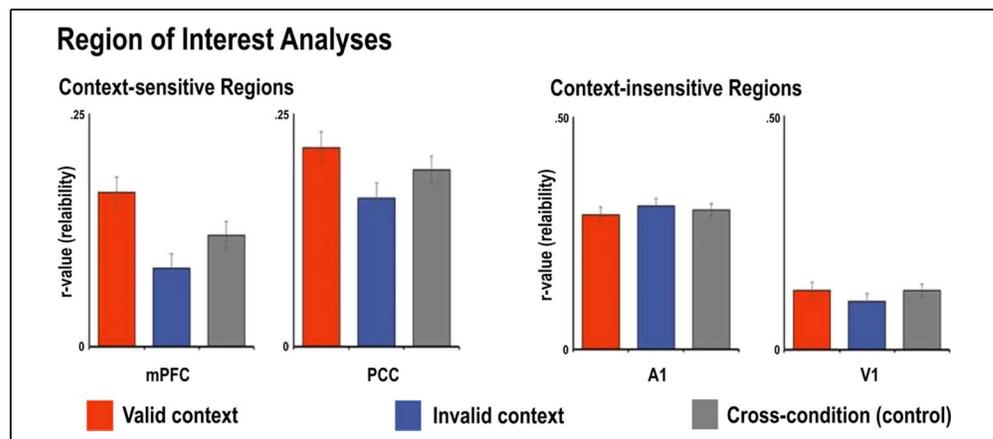
ROIs were defined cytoarchitecturally using the Talairach brain atlas accompanying the BrainVoyager QX software package (BVQX 2.4; Brain Innovation). Mean ISCs were calculated for the valid context and invalid context condi-

tions within each ROI. Cross-condition reliabilities were also calculated for comparison (Figure 5).

Exploratory GLM Analyses

An exploratory whole-brain analysis was performed using a random effects general linear model. Preprocessing procedures were identical to those employed in the ISC analysis. For each participant, events corresponding to the valid context and invalid context conditions were defined according to the onset and offset of each auditory stimulus (vignette) in each condition. These events were then convolved with a canonical hemodynamic response function to create regressors for each condition. Percent signal change transformations were applied to each time course. A liberal FDR correction ($q < .10$)

Figure 5. ROI analyses. ROIs were defined anatomically using a Talairach atlas. Cortical midline structures showed greater ISC when participants shared a valid interpretational context (orange) than when they lacked such a context (blue). Cross-condition reliabilities are shown in gray. Primary sensory regions were insensitive to context, showing statistically identical reliabilities across conditions. mPFC = medial pFC (BA 10); PCC = posterior cingulate cortex (BA 31); A1 = primary auditory cortex (BA 41); V1 = primary visual cortex (BA 17).



was applied to control for false positives. In addition, each of the ROIs defined above was interrogated using a FWER-corrected t test contrast of responses in the valid and invalid context conditions.

RESULTS

To investigate the neural processes underlying the application of knowledge-based context, we presented participants with a series of vignettes, with either a valid or invalid contextual cue (Figure 1A), and examined differences in comprehension as a function of context (Figure 1B). ISC maps were calculated (Figure 2) to identify brain regions that responded reliably across participants, both with and without context (Figure 3). These maps were then directly contrasted using a nonparametric bootstrapping procedure to identify regions whose reliability differed as a function of context (Figure 4). Finally, we visualized the differences in the correlation coefficients in anatomically defined ROIs (Figure 5).

Behavioral Results

Participants' comprehension ratings confirmed that vignettes were more comprehensible when presented with valid contextual cues than without (Figure 1B; see Methods), $M_{(\text{valid context})} = 9.3$ (1.2); $M_{(\text{invalid context})} = 3.2$ (1.3), $t(11) = 20.90$, $p < .001$, $d = 4.96$, confirming that the valid contextual cues facilitated the understanding of the text. The improvement was robust for all 12 vignettes (all $ps < .001$). Treating participants, rather than vignettes, as the unit of analysis yielded comparable results, $M_{(\text{valid context})} = 9.3$ (1.0); $M_{(\text{invalid context})} = 3.2$ (1.8), $t(23) = 16.80$, $p < .001$, $d = 4.25$. We next measured the reliability of the response time courses of neural activation for each vignette as a function of the validity of the contextual cue.

Reliability of Responses for the Valid and Invalid Contextual Cues

Our first objective with respect to the neural data was to map the ISC for both the valid context condition and the invalid context condition (Figure 3). Consistent with previous research (e.g., Honey, Thompson, Lerner, & Hasson, 2012; Hasson, Yang, Vallines, Heeger, & Rubin, 2008; Hasson et al., 2004), maximally reliable correlations were observed in primary auditory cortex and adjacent tonotopic regions (Romanski & Averbek, 2009). Multiple regions known to contribute to language processing also responded reliably in both conditions, including bilateral STS, and the left inferior frontal, supramarginal, and angular gyri, all of which have been linked to linguistic processing to varying extents (Vigneau et al., 2006; Turkeltaub, Eden, Jones, & Zeffiro, 2002; Ferstl & von Cramon, 2001; Robertson et al., 2000; Fiez & Petersen, 1998; Binder et al., 1997; Huettner, Rosenthal, & Hynd, 1989). Several "extralinguistic" regions,

including the dorsal and ventral mPFC, PCC, TPJ, left inferior frontal sulcus, and inferior and superior occipital gyri, were also reliably responsive in both conditions (Honey et al., 2012; Wilson, Molnar-Szakacs, & Jacoboni, 2008; Xu, Kemeny, Park, Frattali, & Braun, 2005).

Visual inspection of Figure 3 suggests a broader extent of reliable responses in the valid context condition as compared with the invalid context condition—particularly along the cortical midline. Consistent with this observation, a calculation of overall volume of significant correlation (after threshold correction) revealed that brains sharing a valid context had approximately one-third (32%) more correlated voxels than brains lacking a valid context. To test whether this overall volumetric difference reflected statistically significant differences between conditions, we conducted bootstrap-based analyses comparing the valid context and invalid context conditions.

Regions Sensitive to Context

A valid context > invalid context comparison based on circular block bootstrapping identified regions of mPFC (BA 10) and PCC (BA 31) as being significantly more reliable when participants shared a valid contextual framework within which to interpret the vignettes. These regions correspond to the "midline core" of the default network, as identified by Andrews-Hanna, Reidler, Huang, and Buckner (2010). At more relaxed thresholds, smaller regions were also observed in bilateral TPJ, which is also a commonly identified component of the DMN. No regions survived the reverse contrast (invalid context > valid context; FDR-corrected; Figure 4). As with the ISC maps then, the midline core of the DMN—that is, the mPFC and PCC—demonstrated the most robust sensitivity to context. To verify this observation, we turned to ROI analyses using anatomically defined regions.

ROI Analysis

ROIs in mPFC (BA 10) and PCC (BA 31) were defined cytoarchitecturally using the Talairach atlas accompanying the BrainVoyager QX software package (BVQX 2.4; Brain Innovation). For comparison, we also interrogated set of ROIs corresponding to low-level sensory regions, which were not expected to demonstrate contextual sensitivity: bilateral primary auditory cortex (A1+; BA 41) and bilateral primary visual cortex (V1; BA 17). For each ROI, mean ISCs were calculated for the valid context and invalid context conditions. Cross-condition reliabilities were also calculated for comparison (see Methods; Figure 5). As predicted, A1+ and V1 showed no functional sensitivity to knowledge-based context, with statistically identical reliabilities emerging in these regions regardless of condition ($ps > .20$). In agreement with the voxel-wise statistical map (Figure 4), the midline core of the default system demonstrated significantly greater reliability when participants shared

context, as compared with when they did not (mPFC: $p < .001$; PCC: $p < .05$, one-tailed, uncorrected).

Exploratory GLM Analyses

Each ROI was also interrogated using a standard event-related contrast comparing the valid context and invalid context conditions (see Methods). No differences in response amplitude were observed in any of these regions: mPFC: $t(23) = 0.45$, $p = .656$; PCC: $t(23) = 0.09$, $p = .928$; V1: $t(23) = 1.07$, $p = .296$; A1+: $t(23) = 0.35$, $p = .730$. Likewise, a whole-brain random effects contrast carried out with FDR correction ($q < .10$, independence assumption) revealed no significant differences in response amplitude between the valid context and invalid context conditions in any brain region. At a reduced threshold of $p < .01$ uncorrected, a number of regions emerged, including a circumscribed region of PCC (BA 31) at $[5, -35, 42]$. As noted in the Introduction (and elaborated in the Discussion), a number of previous studies employing similar contrasts have identified regions of PCC—although there is considerable variability across these studies in the reported locations of these activations (Maguire et al., 1999: $[0, -46, 28]$; Martín-Loeches et al., 2008: $[-9, -49, 37]$ and $[-3, -76, 40]$; Ferstl et al., 2008: $[-8, -51, 33]$ and $[-5, -30, 36]$).

DISCUSSION

The ability to integrate incoming information with relevant contextual knowledge is a crucial component of virtually all cognitive functions, including perception, memory, decision-making, and social cognition. It is therefore natural to ask what brain regions support this integration. We approached this question by measuring the intersubject alignment of neural time courses as participants listened to auditory vignettes preceded either by valid or invalid contextual cues. Stimuli were identical across conditions at all time points entered into the neural analysis; only participants' contextual understanding of these stimuli was manipulated. Cognitively, the provision of a valid contextual framework led to significantly improved comprehension of the vignettes. Neurally, brains that shared such a framework demonstrated greater ISC in the midline core of the default network (mPFC and PCC).

Previous investigations provide partial convergence with these findings. Maguire and colleagues (1999) observed greater mean PET response in PCC to texts accompanied by a valid (rather than invalid) contextual cue. However, a contemporaneous fMRI study with a similar design failed to reproduce this effect (St. George et al., 1999). A later fMRI study (Martín-Loeches et al., 2008) observed a diverse set of regions for a similar contrast: although the main findings localized to the angular gyrus, examination of the full set of activations reveals circumscribed clusters within precuneus/PCC. Moreover, an ALE-based meta-analysis by

Ferstl and colleagues (2008) found that greater PCC and mPFC response tended to be elicited by linguistic stimuli that were more, rather than less, coherent (including, for example, pairs of sentences that were thematically related vs. unrelated); however, some studies employing very similar designs that were not included in this meta-analysis are not fully consistent with this suggestion (Siebörger, Ferstl, & von Cramon, 2007; St. George et al., 1999).

As we have suggested, the mixed nature of this evidence may be, in part, a symptom of applying event-related averaging methods to stimuli for which those methods are not optimally suited (Ben-Yakov et al., 2012). This idea is reflected in the statistical dissociation observed in the present data: When analyzed using event-related averaging methods (i.e., GLM), the present experiment shows no differential sensitivity to context in any region at conventional thresholds and only a small activation in PCC at a far more liberal threshold; however, an ISC approach suggests that PCC and mPFC become preferentially aligned when individuals share a valid contextual framework.

As previously noted, these regions accumulate information over long timescales (Honey et al., 2012; Lerner et al., 2011; Hasson et al., 2008), a property that is essential for the integration of present information with previously acquired contextual knowledge. These same regions appear across numerous research areas as core elements of what is commonly called the default-mode network (Raichle et al., 2001; Shulman et al., 1997) or the social cognition network (Mitchell, 2008; Amodio & Frith, 2006). This network appears to play a critical role in several important cognitive functions that are broadly related to sense making (i.e., integrating new information with a previously established situational context), including comprehension of extended texts (e.g., Mar, 2011; Yarkoni, Speer, & Zacks, 2008), social cognition (e.g., Mitchell, 2009; Lieberman, 2007), memory (e.g., Ranganath & Ritchey, 2012), and self-projection (Buckner & Carroll, 2007).

Each of these functions may contribute to contextual processing. In the present studies, most of the vignettes involved thinking about people (social cognition), and all required recalling the previous context (memory) and linking it to present stimuli (likely drawing upon personal experiences related to those stimuli, e.g., changing a tire, picking flowers, carving a pumpkin, wrapping a present). Imagining oneself in various past or counterfactual situations (self-projection) may be a fundamental component of such mnemonic processes, and these particular vignettes were especially likely to promote such projection, as they were spoken in the second person ("you"). It is difficult, however, to project oneself into a set of disjointed statements that lack a coherent situational framework. Valid contextual cues provide a coherent situational framework into which one may project oneself—which may help to explain why participants are able to better understand and remember hypothetical situations that include context (Bransford & Johnson, 1972).

Conclusion

The integration of present information into existing knowledge structures is fundamental to the way in which we experience the world. Understanding how this integration takes place has long been a goal of researchers in psychology (Johnson-Laird, 1983; Van Dijk & Kintsch, 1983) and linguistics (Goodwin & Duranti, 1992). More recently, neuroscience has begun to reveal how the brain uses contextual information to facilitate certain forms of cognition, such as the rapid processing of objects (Bar, 2004) and words (Rissman et al., 2003; Mummery et al., 1999). However, the question of how the brain accomplishes such integration at higher levels of processing—as in the application of world knowledge in natural language perception—remains largely unexplored. The present experiment begins to address this gap, suggesting that this process is partly subserved by core components of the DMN—regions implicated in various processes related to sense making (Mar, 2011; Mitchell, 2009; Ferstl et al., 2008; Yarkoni et al., 2008; Buckner & Carroll, 2007). In contrast, early visual and auditory cortices, which process relatively low-level stimulus features, as well as many language-related areas showed no sensitivity to high-level contextual cues.

One as-yet-unanswered question is how various memory structures—particularly those situated in the medial-temporal lobe—might contribute to high-level context representation.² Although the application of context-based knowledge clearly involves the selective retrieval of information from memory (Ericsson & Kintsch, 1995), the specific neural processes linking memory and context remain largely obscure. However, given the strong anatomical and functional connections that exist between the medial-temporal lobe and DMN and their joint-importance to situating the self (Ranganath & Ritchey, 2012; Buckner et al., 2008), future investigations may reveal that reciprocal processing across these two networks plays a critical role in the representation and application of high-level situational context.

Acknowledgments

This work was supported by Defense Advanced Research Projects Agency-Broad Agency Announcement 12-03-SBIR Phase II “Narrative Networks” and by the National Science Foundation Graduate Research Fellowship Program. We thank Dmitry Smirnov, Iiro Jääskeläinen, and Mikko Sams for helpful comments on the manuscript.

Reprint requests should be sent to Uri Hasson, 243 Princeton Neuroscience Institute, Washington Road, Princeton University, Princeton, NJ 08540, or via e-mail: hasson@princeton.edu.

Notes

1. During pretesting, a separate sample of participants was given materials similar to those used in the actual experiment. Participants read (rather than listened to) each vignette, either with or without the valid contextual cue (counterbalanced across participants). They then rated their understanding of the vignette

on a 1–7 scale (1 = “I was totally confused”; 7 = “It was all totally clear”) and tried to guess what the vignette was about.

2. In this study, we report context-based effects in cortical areas with long processing timescales. Perhaps surprisingly, we did not observe such context-based effects in either the medial-temporal lobe or the hippocampal formation, regions known to be involved in the encoding and retrieval of information from long-term memory (Lepage, Habib, & Tulving, 1998). On possibility is that the specific task we employed did not require the retrieval of information from long-term memory, given that the cues were presented a few seconds before the beginning of each vignette. However, as always, one needs to be careful with the interpretation of negative results, which may arise because of other factors such as lack of power, inconsistent field of view, or reduced SNR within certain brain regions.

REFERENCES

- Amodio, D. M., & Frith, C. D. (2006). Meeting of minds, the medial frontal cortex and social cognition. *Nature Reviews of Neuroscience*, *7*, 268–277.
- Andrews-Hanna, J. R., Reidler, J. S., Huang, C., & Buckner, R. L. (2010). Evidence for the default network’s role in spontaneous cognition. *Journal of Neurophysiology*, *104*, 322–335.
- Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience*, *5*, 617–629.
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate, a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *57*, 289–300.
- Ben-Yakov, A., Honey, C. J., Lerner, Y., & Hasson, U. (2012). Loss of reliable temporal structure in event-related averaging of naturalistic stimuli. *Neuroimage*, *63*, 501–506.
- Binder, J. R., Frost, J. A., Hammeke, T. A., Cox, R. W., Rao, S. M., & Prieto, T. (1997). Human brain language areas identified by functional magnetic resonance imaging. *Journal of Neuroscience*, *17*, 353–362.
- Bransford, J. D., & Johnson, M. K. (1972). Contextual prerequisites for understanding. Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior*, *11*, 717–726.
- Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The brain’s default network. *Annals of the New York Academy of Sciences*, *1124*, 1–38.
- Buckner, R. L., & Carroll, D. C. (2007). Self-projection and the brain. *Trends in Cognitive Science*, *11*, 49–57.
- Bullmore, E., Long, C., Suckling, J., Fadili, J., Calvert, G., Zelaya, F., et al. (2001). Colored noise and computational inference in neurophysiological (fMRI) time series analysis: Resampling methods in time and wavelet domains. *Human Brain Mapping*, *12*, 61–78.
- Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, *102*, 211–245.
- Ferstl, E. C., Neumann, J., Bogler, C., & Von Cramon, D. Y. (2008). The extended language network, a meta-analysis of neuroimaging studies on text comprehension. *Human Brain Mapping*, *29*, 581–593.
- Ferstl, E. C., & von Cramon, D. Y. (2001). The role of coherence and cohesion in text comprehension: An event-related fMRI study. *Brain Research. Cognitive Brain Research*, *11*, 325–340.
- Fiez, J. A., & Petersen, S. E. (1998). Neuroimaging studies of word reading. *Proceedings of the National Academy of Sciences, U.S.A.*, *95*, 914–921.
- Goodwin, C., & Duranti, A. (1992). Rethinking context: An introduction. In A. Duranti & C. Goodwin (Eds.), *Rethinking*

- context, language as an interactive phenomenon (pp. 1–42). New York: Cambridge University Press.
- Hagoort, P., & van Berkum, J. (2007). Beyond the sentence given. *Philosophical Transactions of the Royal Society, Series B*, *362*, 801–811.
- Hasson, U., Avidan, G., Gelbard, H., Vallines, I., Harel, M., Minshew, N., et al. (2009). Shared and idiosyncratic cortical activation patterns in autism revealed under continuous real-life viewing conditions. *Autism Research*, *2*, 220–231.
- Hasson, U., Malach, R., & Heeger, D. J. (2010). Reliability of cortical activity during natural stimulation. *Trends in Cognitive Science*, *14*, 40–48.
- Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., & Malach, R. (2004). Intersubject synchronization of cortical activity during natural vision. *Science*, *303*, 1634–1640.
- Hasson, U., Yang, E., Vallines, I., Heeger, D. J., & Rubin, N. (2008). A hierarchy of temporal receptive windows in human cortex. *Journal of Neuroscience*, *28*, 2539–2550.
- Honey, C. J., Thompson, C. R., Lerner, Y., & Hasson, U. (2012). Not lost in translation, neural responses shared across languages. *Journal of Neuroscience*, *32*, 15277–15283.
- Huettner, M. I., Rosenthal, B. L., & Hynd, G. W. (1989). Regional cerebral blood flow (rCBF) in normal readers, bilateral activation with narrative text. *Archives of Clinical Neuropsychology*, *4*, 71–78.
- Johnson, M. K., Doll, T. J., Bransford, J. D., & Lapinski, R. H. (1974). Context effects in sentence memory. *Journal of Experimental Psychology*, *103*, 358.
- Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference and consciousness*. Cambridge, UK: Cambridge University Press.
- Lepage, M., Habib, R., & Tulving, E. (1998). Hippocampal PET activations of memory encoding and retrieval: The HIPER model. *Hippocampus*, *8*, 313–322.
- Lerner, Y., Honey, C. J., Silbert, L. J., & Hasson, U. (2011). Topographic mapping of a hierarchy of temporal receptive windows using a narrated story. *Journal of Neuroscience*, *31*, 2906–2915.
- Lieberman, M. D. (2007). Social cognitive neuroscience: A review of core processes. *Annual Review of Psychology*, *58*, 259–289.
- Maguire, E. A., Frith, C. D., & Morris, R. G. M. (1999). The functional neuroanatomy of comprehension and memory: The importance of prior knowledge. *Brain*, *122*, 1839–1850.
- Mar, R. (2011). The neural bases of social cognition and story comprehension. *Annual Review of Psychology*, *63*, 103–134.
- Martín-Loeches, M., Casado, P., Hernández-Tamames, J. A., & Álvarez-Linera, J. (2008). Brain activation in discourse comprehension, a 3t fMRI study. *Neuroimage*, *41*, 614–622.
- Mitchell, J. P. (2008). Contributions of functional neuroimaging to the study of social cognition. *Current Directions in Psychological Science*, *17*, 142–146.
- Mitchell, J. P. (2009). Social psychology as a natural kind. *Trends in Cognitive Sciences*, *13*, 246–251.
- Mummary, C. J., Shallice, T., & Price, C. J. (1999). Dual-process model in semantic priming, a functional imaging perspective. *Neuroimage*, *9*, 516–525.
- Nichols, T. E., & Holmes, A. P. (2002). Nonparametric permutation tests for functional neuroimaging: A primer with examples. *Human Brain Mapping*, *15*, 1–25.
- Patton, A., Politis, D. N., & White, H. (2009). Correction to “automatic block-length selection for the dependent bootstrap” by D. Politis and H. White. *Econometric Reviews*, *28*, 372–375.
- Politis, D. N., & Romano, J. P. (1992). A circular block-resampling procedure for stationary data. *Exploring the limits of bootstrap* (pp. 263–270). New York: Wiley & Sons.
- Politis, D. N., & White, H. (2004). Automatic block-length selection for the dependent bootstrap. *Econometric Reviews*, *23*, 53–70.
- Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences, U.S.A.*, *98*, 676–682.
- Ranganath, C., & Ritchey, M. (2012). Two cortical systems for memory-guided behavior. *Nature Reviews Neuroscience*, *13*, 713–726.
- Rissman, J., Eliassen, J. C., & Blumstein, S. E. (2003). An event-related fMRI investigation of implicit semantic priming. *Journal of Cognitive Neuroscience*, *15*, 1160–1175.
- Robertson, D. A., Gernsbacher, M. A., Guidotti, S. J., Robertson, R. R., Irwin, W., Mock, B. J., et al. (2000). Functional neuroanatomy of the cognitive process of mapping during discourse comprehension. *Psychological Science*, *11*, 255–260.
- Romanski, L. M., & Averbeck, B. B. (2009). The primate cortical auditory system and neural representation of conspecific vocalizations. *Annual Reviews Neuroscience*, *32*, 315–346.
- Shulman, G. L., Fiez, J. A., Corbetta, M., Buckner, R. L., Miezin, F. M., Raichle, M. E., et al. (1997). Common blood flow changes across visual tasks, II. Decreases in cerebral cortex. *Journal of Cognitive Neuroscience*, *9*, 648–663.
- Sieborger, F. T., Ferstl, E. C., & von Cramon, D. Y. (2007). Making sense of nonsense: An fMRI study of task induced inference processes during discourse comprehension. *Brain Research*, *1166*, 77–91.
- Smirnov, D., Glerean, E., Lahnakoski, J. M., Salmi, J., Jääskeläinen, I. P., Sams, M., et al. (2012, June). *Inferior frontal gyrus subserves context-dependent comprehension of spoken narrative*. Poster presented at the 18th Annual Meeting of the Organization for Human Brain Mapping, Beijing, China.
- St. George, M., Kutas, M., Martinez, A., & Sereno, M. I. (1999). Semantic integration in reading, engagement of the right hemisphere during discourse processing. *Brain*, *122*, 1317–1325.
- 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*, 765–780.
- Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academy Press.
- Vigneau, M., Beaucousin, V., Hervé, P.-Y., Duffau, H., Crivello, F., Houde, O., et al. (2006). Meta-analyzing left hemisphere language areas, phonology, semantics, and sentence processing. *Neuroimage*, *30*, 1414–1432.
- Wilson, S. M., Molnar-Szakacs, I., & Iacoboni, M. (2008). Beyond superior temporal cortex, intersubject correlations in narrative speech comprehension. *Cerebral Cortex*, *18*, 230–242.
- Xu, J., Kemeny, S., Park, G., Frattali, C., & Braun, A. (2005). Language in context, emergent features of word, sentence, and narrative comprehension. *Neuroimage*, *25*, 1002–1015.
- Yarkoni, T., Speer, N. K., & Zacks, J. M. (2008). Neural substrates of narrative comprehension and memory. *Neuroimage*, *41*, 1408–1425.