SHORT COMMUNICATION

WILEY | EJN European Journal of Neuroscience FENS

Neural basis of goal-driven changes in knowledge activation

Wei-Chun Wang[1] | Nadia M. Brashier[2] | Erik A. Wing[3] | Elizabeth J. Marsh[3] |
Roberto Cabeza[3]

[1]University of California, Berkeley, Berkeley, California

[2]Harvard University, Cambridge, Massachusetts

[3]Duke University, Durham, North Carolina

Correspondence
Wei-Chun Wang, Helen Wills Neuroscience Institute, University of California, Berkeley, Berkeley, CA.
Email: wcwang@berkeley.edu

Abstract

Depending on a person's goals, different aspects of stored knowledge are accessed. Decades of behavioral work document the flexible use of knowledge, but little neuroimaging work speaks to these questions. We used representational similarity analysis to investigate whether the relationship between brain activity and semantic structure of statements varied in two tasks hypothesized to differ in the degree to which knowledge is accessed: judging truth (*semantic task*) and judging oldness (*episodic task*). During truth judgments, but not old/new recognition judgments, a left-lateralized network previously associated with semantic memory exhibited correlations with semantic structure. At a neural level, people activate knowledge representations in different ways when focused on different goals. The present results demonstrate the potential of multivariate approaches in characterizing knowledge storage and retrieval, as well as the ways that it shapes our understanding and long-term memory.

KEYWORDS

goals, knowledge, representational similarity analysis, semantic memory

1 | INTRODUCTION

Humans know much more than is available at any one moment in time – as examples, you might forget a technical term only for it to pop to mind later, recall only the first letter of an acquaintance's name, or see a robin without immediately thinking of its weight (though you could estimate it if needed). Oftentimes knowledge comes to mind automatically, even when instructed to ignore meaning (e.g., Stroop task; MacLeod, 1991) and other times retrieval fails (e.g., tip-of-the-tongue experiences; Brown, 1991). Depending on the context, different features of a given object or concept take on salience; for example, you can focus on how a word sounds or

on its definition (i.e., *levels of processing*; Craik & Lockhart, 1972). We are adept at resolving lexical ambiguities, quickly interpreting the word "bank" differently depending on whether we are thinking of rivers or money (e.g., Swinney & Hakes, 1976). While decades of behavioral evidence document the flexible use of knowledge, previous neuroimaging studies target semantic processing, not differential activation of underlying representations (Binder, Desai, Graves, & Conant, 2009). We used representational similarity analysis (RSA) to examine how knowledge is represented in the brain, as well as whether different goals influence the activation of these representations.

Representational similarity analysis allows us to take a multivariate approach to quantifying the semantic distance between stimuli (Kriegeskorte, Mur, & Bandettini, 2008). Within a given brain region or searchlight sphere, this method contrasts the voxel-level activity associated with pairs of stimuli, such as pictures of objects (Connolly et al., 2012) or words (Carota, Kriegeskorte, Nili, & Pulvermüller, 2017). This results in a

---

brain representational dissimilarity matrix (RDM), which we can compare to the RDM obtained from a given model (e.g., category membership of objects, semantic similarity among words). Correlating brain and model RDMs tells us whether a voxel "codes for" or disregards distinctions between items along the specified dimension. This technique allows comparisons of semantic structures across modalities and experimental conditions (Devereux, Clarke, Marouchos, & Tyler, 2013). RSA has even demonstrated the emergence of new representations following learning of novel objects (Clarke, Pell, Ranganath, & Tyler, 2016), but little work has explored representational differences as a function of task goals.

The current study tested the extent to which two tasks activated representations reflecting semantic structure, as measured by word co-occurrence statistics. After studying known (e.g., *The capital of France is Paris*) and unknown (e.g., *The inhabitable part of the world is the ecumene*) facts outside the scanner, participants alternated between a semantic (*is this statement true or false*) and an episodic (i.e., *is this statement old or new*) task. The recognition memory test likely activates knowledge (e.g., Jacoby, Shimizu, Daniels, & Rhodes, 2005; Morris, Bransford, & Franks, 1977) without requiring participants to evaluate it; judging truth explicitly encourages both steps. Given that we carefully matched the conditions (i.e., same ratio of old/new and known/unknown facts, 6-point response scale), any differences in activated representations reflect the goals at hand. We predicted that activity patterns would correlate more with semantic structure in the semantic than the episodic task, in regions previously associated with semantic memory: ventromedial prefrontal cortex, posterior cingulate, and left-lateralized inferior frontal gyrus, ventral parietal cortex, lateral and anterior temporal cortex, parahippocampal cortex, and fusiform gyrus (Binder et al., 2009; Visser, Jefferies, & Ralph, 2010).

2 | MATERIALS AND METHODS

2.1 | Participants

The Duke University Institutional Review Board approved all procedures. Thirty-one native English speakers from Duke University and the surrounding communities participated for monetary compensation. Seven participants were excluded (three due to technical malfunctions with the scanner or testing computer, two fell asleep, one was at chance, and another failed to use the full scale). The final sample included 24 participants (age $M = 23.17$, $SEM = 0.68$; education $M = 15.33$, $SEM = 0.41$; 10 female).

2.2 | Materials

Materials consisted of 360 trivia statements collected from the Internet that referred to known or unknown facts. One third of the sentences were known facts (e.g., *The capital of*

Spain is called Madrid), and the remaining two-thirds comprised of unknown facts (e.g., *The inhabitable part of the world is the ecumene*) and unknown facts with a matching false framing (e.g., *The inhabitable part of the world is the toponym*) that referred to a plausible, but incorrect, alternative (counterbalanced across participants). The false statements address another research question (Wang, Brashier, Wing, Marsh, & Cabeza, 2016, 2018) and were excluded from the RSA analysis. Given that true unknown items still invoked familiar concepts (e.g., that a part of the world is inhabitable), we included true unknown items.

Pilot participants ($N = 47$) rated these items from one (*definitely false*) to six (*definitely true*); they reliably and confidently rated the known facts as true (i.e., $5 = $ *probably true* or $6 = $ *definitely true* responses from >90% of participants) and the unknown facts as unknown (i.e., $3 = $ *guess false* or $4 = $ *guess true* responses from >75% of participants).

2.3 | Procedure

Following informed consent, participants incidentally encoded (outside the scanner) 180 statements twice each by rating their interest on a 6-point scale from *very uninteresting* to *very interesting*. Each statement appeared for 4 s, followed by a fixed 1 s interstimulus interval fixation. Participants then entered the scanner where they alternated between a semantic task (i.e., *is this statement true or false*) and an episodic task (i.e., *is this statement old or new*) in four separate counterbalanced ABBA event-related runs. Participants responded on a 6-point scale (order counterbalanced across participants) in both the semantic (*definitely false*, *probably false*, *guess false*, *guess true*, *probably true*, *definitely true*) and episodic (*definitely new*, *probably new*, *guess new*, *guess old*, *probably old*, *definitely old*) tasks. In each run, participants rated 30 old unknown, 30 new unknown, 15 old known, and 15 new known statements. Each statement appeared for 5 s with a jittered (mean = 3 s; range: 1–8 s) interstimulus interval fixation. The jitter, as well as the randomization of trials, was calculated using the FMRI Toolbox (https://sourceforge.net/projects/fmri-toolbox/files/optimize_design/1.1) based on the Genetic Algorithm (Wager & Nichols, 2003).

2.4 | RDM construction

We constructed a model RDM of semantic dissimilarity for the facts using the website cortical.io. Cortical.io utilizes an unsupervised machine learning algorithm to encapsulate the entirety of semantic space (trained on Wikipedia) into a vector of 16,384 co-occurring concepts, or 'semantic contexts' (Webber, 2015). To calculate semantic similarity, the software extracts the key words of each fact (e.g., nouns) and identifies the semantic contexts (in the vector of 16,384) associated with those key words (i.e., a binary 'semantic
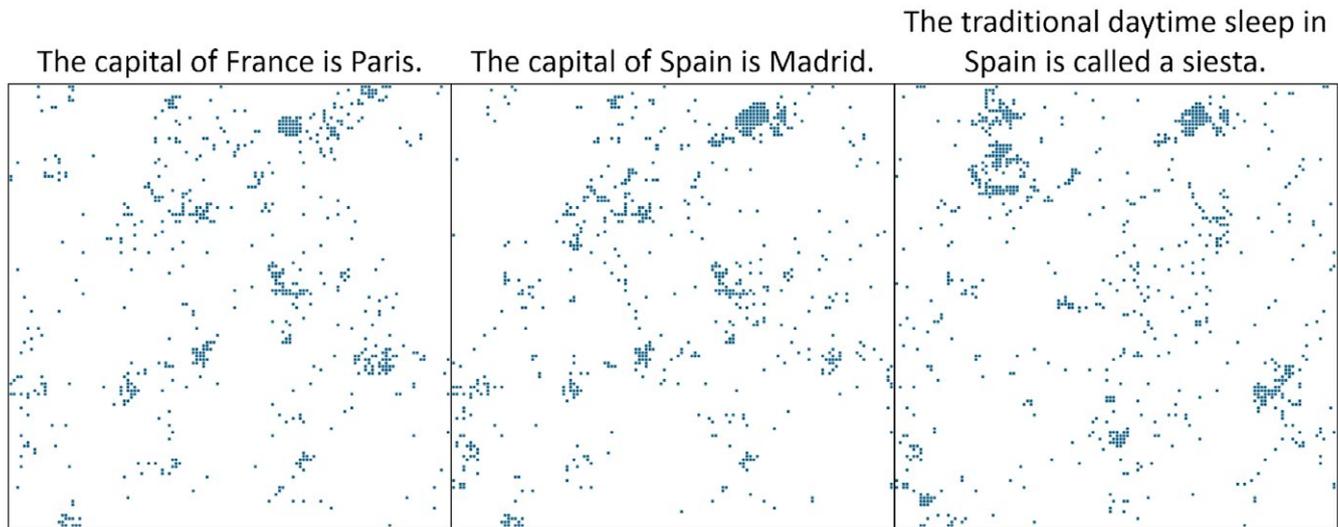
The capital of France is Paris.   The capital of Spain is Madrid.   The traditional daytime sleep in Spain is called a siesta.

FIGURE 1 A visual representation of the semantic fingerprints for three facts. The overlap between each fact reflects common semantic contexts

fingerprint' that is visualized with a 128^2 matrix). Then, the dissimilarity (1-cosine similarity) between the semantic fingerprint for each pair of facts is calculated (i.e., where 0 = identical sentences). See Figure 1 for three representative semantic fingerprints that illustrate shared and unique semantic contexts between facts, and Table 1 for a representative subset of pairwise dissimilarity values. This 'semantic fingerprint' method is relatively new, but has successfully grouped similar firms based on their business descriptions (Ibriyamova, Kogan, Salganik-Shoshan, & Stolin, 2017) and academic authors based on the content of their publications (Han et al., 2017). Thus, it provides a useful tool for extracting semantic similarity beyond just a single word. An in-depth examination of the advantages of this method is beyond the scope of the current study, but for a discussion, see Webber (2016).

We conducted an additional control analysis to ensure that low-level visual similarity could not explain any observed effects. A separate model RDM was created by taking the absolute value of the difference in character count between each pair of facts (i.e., a measure of low-level visual similarity), and this RDM was then correlated with the brain RDM separately in each task.

2.5 | Image acquisition and analysis

Images were collected on a 3T General Electric scanner with an 8-channel head coil at the Duke University Brain Imaging and Analysis Center. Functional MRI (FMRI) images were acquired using a SENSE spiral sequence (64×64 matrix, repetition time = 2000 ms, echo time = 27 ms, field of view = 24 cm, flip angle = 60°) and consisted of 34 axial slices acquired in an interleaved fashion. Slice thickness was 3.8 mm, resulting in $3.75 \times 3.75 \times 3.8$ mm voxels.

Additionally, high-resolution structural images were collected using a 3D, T1-weighted FSPGR sequence (256×256 matrix, 166 slices, 1 mm isotropic voxels).

Data were preprocessed with SPM12 (Wellcome Trust Centre for Neuroimaging). After discarding the first three volumes of each run, the functional data for each participant were slice-time corrected, realigned, and coregistered to their respective anatomical images. The anatomical images were then segmented into separate gray and white matter images that were used to normalize the functional and anatomical images into MNI space. Lastly, the normalized functional data were denoized using the DRIFTER toolbox (Särkkä et al., 2012).

Statistical analyses were performed in SPM12 using the general linear model. A high-pass filter of 128 sec and grand mean scaling were applied to the data, and serial correlations in the time series were accounted for using the autoregressive model (AR[1]). For the univariate analysis, eight separate conditions of interest – old and new known and unknown statements seen during the episodic and the semantic tasks – were modeled on the smoothed data (8-mm isotropic FWHM Gaussian filter). For our RSA analysis, each fact was modeled in a separate general linear model using the Least Squares - Separate approach (one regressor for the trial of interest, and one nuisance regressor for all other trials) on the unsmoothed data, yielding first-level single-trial beta images for each trial in each participant (Mumford, Turner, Ashby, & Poldrack, 2012). Other covariates of no interest included the six motion parameters estimated during realignment, baseline and session effects, global mean, and motion outliers obtained from the Artifact Detection Toolbox (http://www.nitrc.org/projects/artifact_detect), and white matter and CSF signal time courses.

A paired sample *t* test examined the univariate difference between episodic and semantic trials. For the RSA

TABLE 1 Representative subset of pairwise dissimilarity (1-cosine similarity) values derived from cortical.io. Green, low dissimilarity; Yellow, high dissimilarity

	The capital city of France is called Paris	The capital city of Spain is called Madrid	A unit for measuring pain intensity is the dol	The rate of change in velocity is acceleration	The inability to sleep is called insomnia	The traditional daytime sleep in Spain is called a siesta
The capital city of France is called Paris	0	0.52	0.94	0.98	0.91	0.85
The capital city of Spain is called Madrid	0.52	0	0.94	0.98	0.93	0.57
A unit for measuring pain intensity is the dol	0.94	0.94	0	0.79	0.74	0.84
The rate of change in velocity is acceleration	0.98	0.98	0.79	0	0.9	0.94
The inability to sleep is called insomnia	0.91	0.93	0.74	0.9	0	0.61
The traditional daytime sleep in Spain is called a siesta	0.85	0.57	0.84	0.94	0.61	0

analysis, Spearman's correlations were computed between each model RDM (e.g., the 1-cosine similarity between facts) and the brain RDM (i.e., the pairwise correlations of the single trial betas) for each participant using an in-house searchlight script (https://github.com/brg015/mfMRI_v2/) with a 3-voxel searchlight sphere, separately for the semantic and episodic tasks. That is, within each searchlight sphere, the brain RDM was correlated with the cortical. io RDM separately for trials in each task (Kriegeskorte et al., 2008). Similarly, the brain RDM was correlated with the character count RDM separately for each task. For group analyses, these correlation maps were then spatially smoothed (8-mm isotropic FWHM Gaussian filter) and analyzed with one-sample and paired sample t tests. All analyses were corrected for multiple comparisons with 3dClustSim (version 18.0.11) using an uncorrected threshold of $p < 0.001$ and a cluster extent of 72 voxels (for a discussion of cluster-level corrections, see Slotnick, 2017).

To compare our findings to existing analyses of semantic memory, we then uploaded the group correlation maps from the semantic RSA analysis into Neurovault, a public repository of neuroimaging data (Gorgolewski et al., 2015). This allowed us to quantitatively compare our pattern of results to meta-analysis maps from Neurosynth, a database of over 10,000 FMRI studies that allows for large-scale analyses of neuroimaging data (Yarkoni, Poldrack, Nichols, Van Essen, & Wager, 2011).

3 | RESULTS

3.1 | RSA results

As shown in Figure 2a, a network of regions including medial prefrontal cortex, posterior cingulate, and left-lateralized inferior frontal gyrus, superior frontal gyrus, ventral parietal cortex, lateral and anterior temporal cortex, posterior hippocampus and parahippocampal cortex, and fusiform gyrus exhibited significant correlations with our semantic RDM during the semantic task. These regions are consistent with meta-analyses of regions activated during semantic memory tasks (Binder et al., 2009; Visser et al., 2010). In contrast, no suprathreshold clusters exhibited significant correlations during the episodic task. A paired sample t test indicated that correlations were stronger for the semantic than the episodic task in posterior cingulate, medial prefrontal cortex, left anterior temporal cortex, and right dorsolateral prefrontal cortex (Figure 2b).

In comparison, the character count RDM yielded significant correlations in bilateral visual cortex in both the semantic (Figure 3a) and the episodic (Figure 3b) tasks. A paired sample t test revealed no significant differences between the two tasks, suggesting that they did not differ in the extent to which they evoked low-level visual processing.
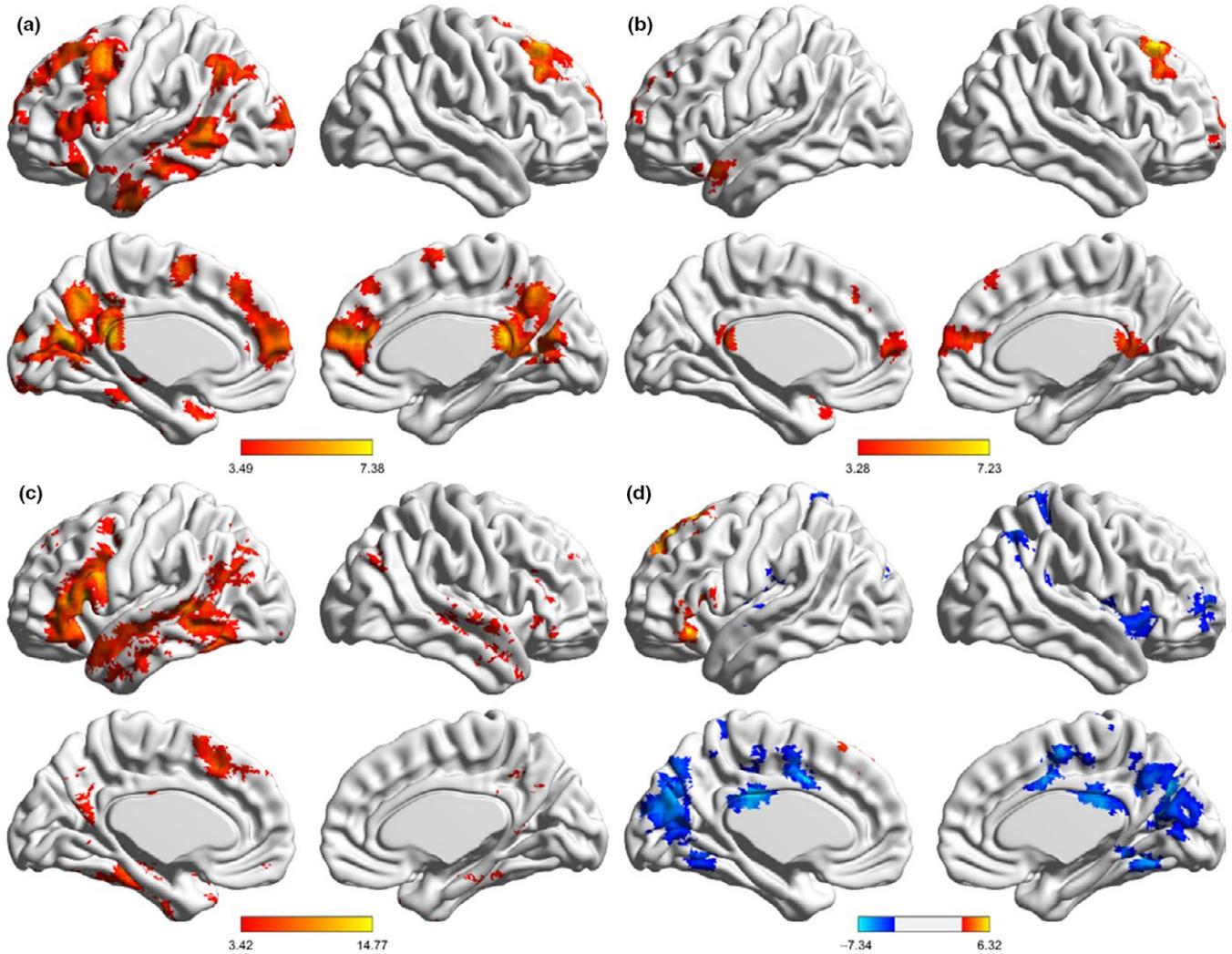
FIGURE 2 (a) Regions exhibiting significant correlations with the semantic representational dissimilarity matrix (RDM) during the semantic task. No suprathreshold clusters were observed with this RDM during the episodic task. (b) Paired sample *t* test showing stronger correlations for the semantic than the episodic task. (c) Neurosynth meta-analysis map for the term *semantic*. (d) Paired sample *t* test showing greater univariate activity for semantic than episodic trials (warm colors) and for episodic than semantic trials (cool colors)

3.2 | Comparison to existing studies

To better understand the pattern of observed correlations, we compared our group correlation maps from the semantic RSA analysis to Neurosynth meta-analysis maps (Yarkoni et al., 2011). Figure 2c shows the Neurosynth meta-analysis map for the term *semantic*, which was the term most strongly associated with our RSA results during the semantic task. Table 2 shows the 10 strongest correlations between Neurosynth meta-analysis maps (excluding anatomical regions or networks, e.g., *posterior cingulate*) and our group map for both the semantic task (http://neurosynth.org/decode/?neurovault=52941), and the episodic task (http://neurosynth.org/decode/?neurovault=52945).

Notably, the terms most strongly associated with our group map for the episodic task also related to semantic memory and/or language, but at a lower magnitude (e.g., 0.16 vs. 0.40 for the term *semantic*). Thus, our pattern of RSA results not only complement previous meta-analyses of semantic memory (Binder et al., 2009; Visser et al., 2010), but also are similar to data-driven meta-analyses derived from Neurosynth.

3.3 | Univariate results

A paired sample *t* test between episodic and semantic trials revealed greater activity for semantic than episodic trials in left inferior and superior frontal gyrus, and greater activity for episodic than semantic trials in right prefrontal cortex and
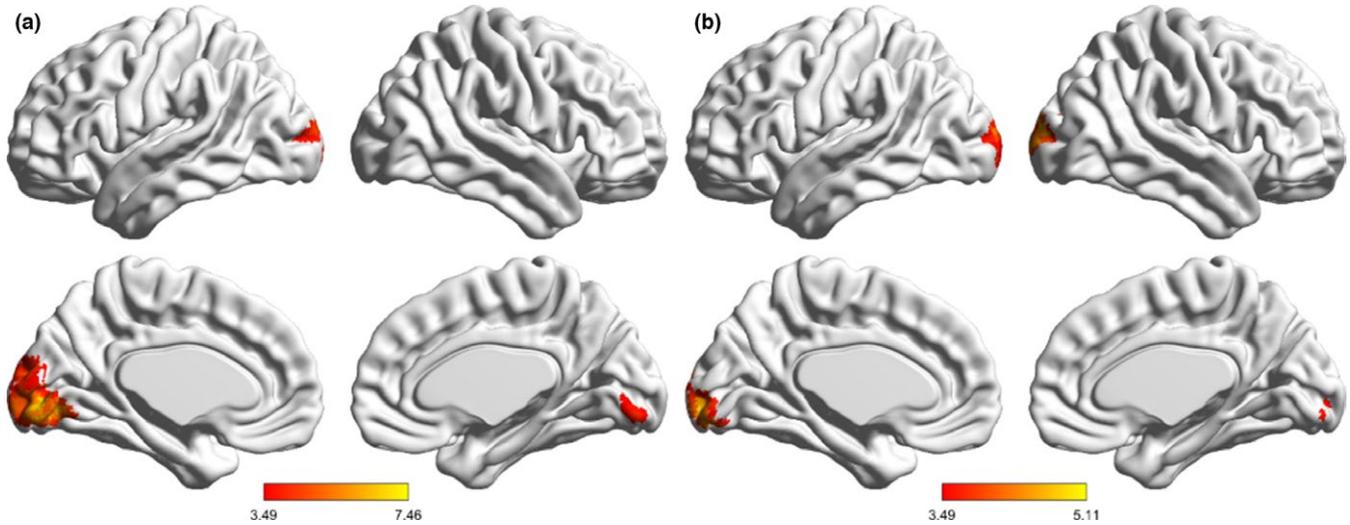
FIGURE 3 Regions exhibiting significant correlations with the character count representational dissimilarity matrix (RDM) (a) during the semantic task and (b) during the episodic task

TABLE 2 Top 10 strongest correlations between semantic RSA maps for both tasks and meta-analysis maps from Neurosynth

Term	Semantic	Episodic	No of studies
Semantic[a]	0.400	0.164	844
Language	0.378	0.184	855
Word	0.361	0.192	782
Words	0.361	0.182	778
Sentence	0.324	0.156	266
Lexical	0.316	0.158	272
Reading[b]	0.308	0.194	427
Sentences	0.306	0.121	307
Phonological	0.304	0.152	310
Syntactic	0.289	0.149	145

[a]Strongest correlation with the semantic task. [b]Strongest correlation with the episodic task.

bilateral parieto-occipital cortex (Figure 2d), suggesting that differences between tasks in the semantic RSA results were not due to univariate activity differences.

4 | DISCUSSION

The present research demonstrates that changing one's goals influences the extent to which neural activation corresponds to semantic structure. When explicitly encouraged to retrieve and evaluate relevant facts (*semantic task*), participants exhibited activity in distributed regions implicated in semantic memory that correlated with "semantic space." This mapping was absent when participants made old/new recognition judgments (*episodic task*). Notably, the correlations during the episodic task still resemble "semantic" meta-analysis maps

from Neurosynth, but to a lesser degree than in the semantic task. A direct correlation of the two maps (https://www.neurovault.org/images/compare/52941/52945) reveals highly overlapping semantic structures, regardless of whether people attended to truth or oldness (occipital cortex: $r = 0.65$, parietal cortex: $r = 0.58$, temporal cortex: $r = 0.48$, frontal cortex: $r = 0.45$). In other words, people retrieved relevant knowledge even when making recognition judgments, just to a lesser extent than during truth judgments.

Crucially, our results are consistent with evidence that multivariate data are not merely a byproduct of univariate activity differences (Jimura & Poldrack, 2012). Within regions where activation patterns correlated with semantic structure during truth judgments, a range of univariate patterns emerged. Posterior regions exhibited greater univariate activity for the episodic than the semantic task, and left frontal areas showed the opposite pattern; no differences emerged in temporal and parietal regions, even at a relatively liberal threshold ($p < 0.05$, uncorrected). Additionally, a control RSA of character count revealed significant correlations in visual cortex in both tasks, suggesting that there are no discernible confounds between them.

Behavioral work clearly shows that people do not always access their own knowledge – *marginal knowledge* is stored in memory, but cannot be retrieved (e.g., foreign vocabulary words after a class ends; Berger, Hall, & Bahrick, 1999). Even to the extent that facts *are* accessible, they are not always successfully applied, such as when people fail to notice any problem with the question, *How many animals of each kind did Moses take on the ark?* (Erickson & Mattson, 1981). Understanding the neural basis of these phenomena is crucial, since there are important downstream consequences. Failing to notice a contradiction with stored knowledge (e.g., that Noah, not Moses, took animals on the ark) makes it more likely to be repeated

later (Bottoms, Eslick, & Marsh, 2010). On the flip side, successfully bringing related knowledge to bear allows us to organize new information (e.g., Bransford & Johnson, 1972) and better remember it later (i.e., *schematic support*; Castel & Craik, 2003).

While neuroimaging studies identify brain regions that support knowledge retrieval, the focus is often on *process* rather than representation. These studies reveal a common network supporting the retrieval of facts, recent instances, and personal memories (Burianova, McIntosh, & Grady, 2010), as well as point to the crucial role the ventromedial prefrontal cortex plays in schema-based learning (Gilboa & Marlatte, 2017). More recently, however, RSA revealed the "semantic atlas" used to extract meaning from speech (Huth, de Heer, Griffiths, Theunissen, & Gallant, 2016). Similar to these characterizations of the representations underlying language comprehension (e.g., Carota et al., 2017), our work suggests that current goals differentially activate the *representations* of facts and confirms the importance of multivariate approaches when considering knowledge retrieval.

ACKNOWLEDGEMENTS

CONFLICTS OF INTEREST

The authors have no conflicts of interest to report.

AUTHOR CONTRIBUTIONS

WCW collected and analyzed the data and wrote the paper. NMB designed the study, collected the data, and revised the manuscript. EAW provided input on data analysis and revised the manuscript. EJM designed the study and revised the manuscript. CB designed the study and revised the manuscript. All authors approved the final manuscript.

DATA ACCESSIBILITY

Group RSA maps are available online at https://neurovault.org/collections/2865/.

ORCID

Wei-Chun Wang ⓘ http://orcid.org/0000-0003-1788-2727

REFERENCES

Berger, S. A., Hall, L. K., & Bahrick, H. P. (1999). Stabilizing access to marginal and submarginal knowledge. *Journal of Experimental Psychology Applied*, *5*, 438–447. https://doi.org/10.1037/1076-898X.5.4.438

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, *19*, 2767–2796. https://doi.org/10.1093/cercor/bhp055

Bottoms, H. C., Eslick, A. N., & Marsh, E. J. (2010). Memory and the Moses illusion: Failures to detect contradictions with stored knowledge yield negative memorial consequences. *Memory*, *18*, 670–678. https://doi.org/10.1080/09658211.2010.501558

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. https://doi.org/10.1016/S0022-5371(72)80006-9

Brown, A. S. (1991). A review of the tip-of-the-tongue experience. *Psychological Bulletin*, *109*, 204–223. https://doi.org/10.1037/0033-2909.109.2.204

Burianova, H., McIntosh, A. R., & Grady, C. L. (2010). A common functional brain network for autobiographical, episodic, and semantic memory retrieval. *NeuroImage*, *49*, 865–874. https://doi.org/10.1016/j.neuroimage.2009.08.066

Carota, F., Kriegeskorte, N., Nili, H., & Pulvermüller, F. (2017). Representational similarity mapping of distributional semantics in left inferior frontal, middle temporal, and motor cortex. *Cerebral Cortex*, *27*, 294–309.

Castel, A. D., & Craik, F. I. M. (2003). The effects of aging and divided attention on memory for item and associative information. *Psychology and Aging*, *18*, 873–885. https://doi.org/10.1037/0882-7974.18.4.873

Clarke, A., Pell, P. J., Ranganath, C., & Tyler, L. K. (2016). Learning warps object representations in the ventral temporal cortex. *Journal of Cognitive Neuroscience*, *28*, 1010–1023. https://doi.org/10.1162/jocn_a_00951

Connolly, A. C., Guntupalli, J. S., Gors, J., Hanke, M., Halchenko, Y. O., Wu, Y.-C., … Haxby, J. V. (2012). The representation of biological classes in the human brain. *Journal of Neuroscience*, *32*, 2608–2618. https://doi.org/10.1523/JNEUROSCI.5547-11.2012

Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*, 671–684. https://doi.org/10.1016/S0022-5371(72)80001-X

Devereux, B. J., Clarke, A., Marouchos, A., & Tyler, L. K. (2013). Representational similarity analysis reveals commonalities and differences in the semantic processing of words and objects. *Journal of Neuroscience*, *33*, 18906–18916. https://doi.org/10.1523/JNEUROSCI.3809-13.2013

Erickson, T. D., & Mattson, M. E. (1981). From words to meaning: A semantic illusion. *Journal of Verbal Learning and Verbal Behavior*, *20*, 540–551. https://doi.org/10.1016/S0022-5371(81)90165-1

Gilboa, A., & Marlatte, H. (2017). Neurobiology of schemas and schema-mediated memory. *Trends in Cognitive Sciences*, *21*, 618–631. https://doi.org/10.1016/j.tics.2017.04.013

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

Han, H., Yao, C., Fu, Y., Yu, Y., Zhang, Y., & Xu, S. (2017). Semantic fingerprints-based author name disambiguation in

Chinese documents. *Scientometrics*, *111*, 1879–1896. https://doi.org/10.1007/s11192-017-2338-6

Huth, A. G., de Heer, W. A., Griffiths, T. L., Theunissen, F. E., & Gallant, J. L. (2016). Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature*, *532*, 453–458. https://doi.org/10.1038/nature17637

Ibriyamova, F., Kogan, S., Salganik-Shoshan, G., & Stolin, D. (2017). Using semantic fingerprinting in finance. *Applied Economics*, *49*, 2719–2735. https://doi.org/10.1080/00036846.2016.1245844

Jacoby, L. L., Shimizu, Y., Daniels, K. A., & Rhodes, M. G. (2005). Modes of cognitive control in recognition and source memory: depth of retrieval. *Psychonomic Bulletin and Review*, *12*, 852–857. https://doi.org/10.3758/BF03196776

Jimura, K., & Poldrack, R. A. (2012). Analyses of regional-average activation and multivoxel pattern information tell complementary stories. *Neuropsychologia*, *50*, 544–552. https://doi.org/10.1016/j.neuropsychologia.2011.11.007

Kriegeskorte, N., Mur, M., & Bandettini, P. (2008). Representational similarity analysis–connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, *2*, 4. https://doi.org/10.3389/neuro.06.004.2008

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, *109*, 163–203. https://doi.org/10.1037/0033-2909.109.2.163

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior*, *16*, 519–533. https://doi.org/10.1016/S0022-5371(77)80016-9

Mumford, J. A., Turner, B. O., Ashby, F. G., & Poldrack, R. A. (2012). Deconvolving BOLD activation in event-related designs for multi-voxel pattern classification analyses. *NeuroImage*, *59*, 2636–2643. https://doi.org/10.1016/j.neuroimage.2011.08.076

Särkkä, S., Solin, A., Nummenmaa, A., Vehtari, A., Auranen, T., Vanni, S., & Lin, F.-H. (2012). Dynamic retrospective filtering of physiological noise in BOLD fMRI: dRIFTER. *NeuroImage*, *60*, 1517–1527. https://doi.org/10.1016/j.neuroimage.2012.01.067

Slotnick, S. D. (2017). Cluster success: fMRI inferences for spatial extent have acceptable false-positive rates. *Cognitive Neuroscience*, *8*, 150–155. https://doi.org/10.1080/17588928.2017.1319350

Swinney, D. A., & Hakes, D. T. (1976). Effects of prior context upon lexical access during sentence comprehension. *Journal of Verbal Learning and Verbal Behavior*, *15*, 681–689. https://doi.org/10.1016/0022-5371(76)90060-8

Visser, M., Jefferies, E., & Ralph, M. L. (2010). Semantic processing in the anterior temporal lobes: A meta-analysis of the functional neuroimaging literature. *Journal of Cognitive Neuroscience*, *22*, 1083–1094. https://doi.org/10.1162/jocn.2009.21309

Wager, T. D., & Nichols, T. E. (2003). Optimization of experimental design in fMRI: A general framework using a genetic algorithm. *NeuroImage*, *18*, 293–309. https://doi.org/10.1016/S1053-8119(02)00046-0

Wang, W. C., Brashier, N. M., Wing, E. A., Marsh, E. J., & Cabeza, R. (2016). On known unknowns: fluency and the neural mechanisms of illusory truth. *Journal of Cognitive Neuroscience*, *28*, 739–756. https://doi.org/10.1162/jocn_a_00923

Wang, W. C., Brashier, N. M., Wing, E. A., Marsh, E. J., & Cabeza, R. (2018). Knowledge supports memory retrieval through familiarity, not recollection. *Neuropsychologia*, *113*, 14–21. https://doi.org/10.1016/j.neuropsychologia.2018.01.019

Webber, F. D. S. (2015). *Semantic Folding Theory and its Application in Semantic Fingerprinting*. arXiv preprint, arXiv:1511.08855.

Webber, F. D. S. (2016). *Arguments for Semantic Folding and Hierarchical Temporal Memory Theory*. Retrieved from http://www.cortical.io/static/downloads/Semantic-Folding-Arguments_V01.pdf.

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*, 665–670. https://doi.org/10.1038/nmeth.1635