European Journal of Neuroscience, Vol. 22, pp. 1233–1238, 2005

An fMRI study of causal judgments

Ajay B. Satpute,[1] Daniela B. Fenker,[2,3] Michael R. Waldmann,[2] Golnaz Tabibnia,[1] Keith J. Holyoak[1] and Matthew D. Lieberman[1]
[1]Department of Psychology, 1285 Franz Hall, University of California, Los Angeles, CA 90095-1563, USA
[2]Department of Neurology II, Otto-v.-Guericke University, Magdeburg, Germany
[3]University of Göttingen, Germany

Keywords: causality, dorsolateral prefrontal cortex, humans, relations, semantic memory

Abstract

The capacity to evaluate causal relations is fundamental to human cognition, and yet little is known of its neurocognitive underpinnings. A functional magnetic resonance imaging study was performed to investigate an hypothesized dissociation between the use of semantic knowledge to evaluate specifically causal relations in contrast to general associative relations. Identical pairs of words were judged for causal or associative relations in different blocks of trials. Causal judgments, beyond associative judgments, generated distinct activation in left dorsolateral prefrontal cortex and right precuneus. These findings indicate that the evaluation of causal relations in semantic memory involves additional neural mechanisms relative to those required to evaluate associative relations.

Introduction

Knowledge of causal relations is critical for planning, acting and reasoning (Spellman, 1997; Lieberman *et al.*, 2002; for a review see Buehner & Cheng, 2005). Neural studies have begun to investigate the perception of causality (Fugelsang *et al.*, in press), the acquisition of new causal relations (Turner *et al*., 2004; Fugelsang & Dunbar, 2005) and the role of causal understanding in text processing (Mason & Just, 2004). However, the neurocognitive mechanisms underlying evaluation of causal relations already stored in semantic memory have not been directly investigated (unlike many other aspects of semantic memory; e.g. Gabrieli *et al.*, 1998; Wagner *et al.*, 2000). In the present study, we use functional magnetic resonance imaging (fMRI) to examine the brain regions involved in processing stored causal relations.

Behavioral studies of semantic memory have exploited priming paradigms to investigate the organization of semantic memory (e.g. 'hospital' will prime a lexical decision about 'doctor'; see Medin & Rips, 2005). The most general account of priming postulates spreading activation based on associative links between concepts, a process that does not depend on the specific semantic content of the linking relation. Neuroimaging studies have implicated left inferior prefrontal cortex as a locus specific to semantic priming (Gabrieli *et al*., 1998; Wagner *et al*., 2000).

Priming can also be triggered by specific shared relations, even when associative priming is controlled (Spellman *et al*., 2001). A critical computational requirement for representing relations is dynamic role binding (Hummel, 1999; Hummel & Holyoak, 2003) – using working memory to represent and manipulate elements bound to particular roles in structured relations. We hypothesize that evaluating causal relations stored in semantic memory will also depend on the neural processes that underlie role binding. Evaluating a causal relation may require forming a working-memory representation in which specific events are bound to the roles of 'cause' and 'effect'. These relations are asymmetric (Waldmann, 1996). For example, smoking cigarettes causes lung cancer, whereas lung cancer does not cause smoking. Accordingly, assessing the veracity of a causal relation will depend on more than simple associative priming between two concepts in semantic memory; it is also necessary to bind the relevant events into the 'cause' and 'effect' roles.

Research has demonstrated that the dorsolateral prefrontal cortex (DLPFC) is involved in cognitive processes that require role binding, including deductive reasoning (Goel *et al*., 1997; Waltz *et al*., 1999), inductive reasoning (Christoff *et al*., 2001; Kroger *et al*., 2002), analogical reasoning (Morrison *et al*., 2004; Bunge *et al*., 2005) and related working-memory tasks (Cohen *et al*., 1994; Smith & Jonides, 1999). We hypothesized that making causal judgments would disproportionately activate DLPFC. Associative judgments, by contrast, can be made without reference to any specific role bindings. For example, 'smoking' may prime 'cancer', and vice versa, without necessarily representing which is the cause and which the effect. It follows that causal judgments, compared with associative judgments, should distinctly activate DLPFC even when the referents being evaluated are held constant.

Method

To investigate the above hypothesis, we used a relation-judgment paradigm (Fenker *et al*., in press) in which participants determined whether pairs of words were causally related (causal condition) or associatively related (associative condition; see Fig. 1). Importantly, the critical target word pairs were selected to be both causally and associatively related (e.g. 'spark' and 'fire' are both causally related and associatively related). Participants were therefore required to make the same behavioral response to the same word pairs regardless of condition; however, the basis for this response was expected to differ depending on condition. In the causal condition, participants must access the referents in semantic memory, and then use working memory to bind the referents to cause and effect roles and verify their

Correspondence: A. B. Satpute, as above.
E-mail: satpute@psych.ucla.edu

Received 23 February 2005, revised 8 May 2005, accepted 21 June 2005

doi:10.1111/j.1460-9568.2005.04292.x

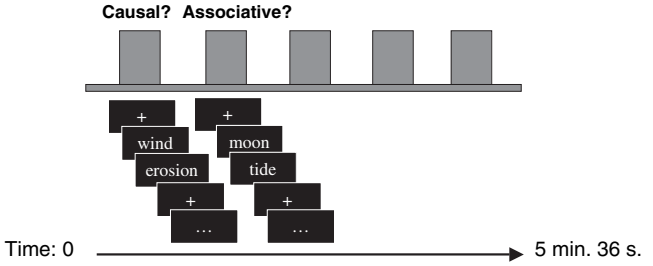

FIG. 1. Schematic representation of the experimental design. The top portion shows the five blocks (two Causal, two Associative, one filler) within each functional run. Each block was preceded by the task cue 'Causal?' or 'Associative?'. The bottom portion shows a sample of trials within a block in which word pairs are causally and associatively related (moon–tide, wind–erosion) or unrelated (e.g. eggs–liar; not shown in figure).

relationship. In the associative condition, participants must still access referents of the words in semantic memory; however, a judgment can then be made without binding referents to cause and effect roles. To avoid carry-over effects that would be expected if participants were asked to switch judgment tasks from trial to trial and to match Fenker *et al.*'s (in press) paradigm as closely as possible, a blocked design was used in which participants judged word pairs for causal relations in some blocks and for associativity in other blocks. If causal processing requires additional computations, such as role binding, that are not required for associative judgments, then the causal condition should distinctly activate DLPFC.

Participants

Twelve right-handed adults (six females; mean age 27.3 years, range 18–48 years; 11 Caucasian, one Asian) with normal or corrected to normal vision participated in the study. Upon arrival at the UCLA Brain Mapping Center, participants signed informed-consent forms in compliance with the UCLA Institutional Review Board's policy. All participants consented to the study and received $30 for participation.

Behavioral tasks

Behavioral tasks were programmed in Superlab® (Cedrus Corporation, San Pedro, CA, USA) and implemented on a Macintosh G3 desktop computer. The stimulus material consisted of 64 weakly associated but causally related word pairs (e.g. 'moon–tide'), 16 weakly associated but non-causal word pairs (e.g. 'ring–emerald') and 30 unrelated filler word pairs (e.g. 'eggs–liar'). Word pairs were selected from the University of South Florida (USF) Word Association Norm list (Nelson *et al.*, 1998) coupled with an additional norming study to select causally related word pairs that were equated in both forward and backward directions (i.e. 'cause' word appearing before or after the 'effect' word in each pair) in terms of the strength of associative relations (forward and backward strength < 0.01 based on the USF norms). The test pairs were presented in Arial Black font, size 24 point, on a white background. The words were created as pict files in Canvas 6.0 graphics software (ACD Systems, Miami, FL, USA).

Word pairs were organized into ten blocks each consisting of 11 trials. Eight blocks each contained eight causally related word pairs. Because word pairs were presented one word at a time, four of these blocks presented the cause word first and four presented the effect word first. Every block also included three unrelated filler word pairs to ensure that participants attended to the task. Prior to each block a prompt of either the word 'CAUSAL?' or the word 'ASSOCIATIVE?' appeared for 4000 ms. If the prompt 'CAUSAL?' appeared, subjects had to decide whether or not there was a causal relation between the two words; if the prompt was 'ASSOCIATIVE?' they had to decide whether or not these words were associated. Causally related word pairs were thus presented in both the associative and the causal conditions. Because causally related word pairs were also selected to be weakly related in the USF association norms, participants should make a 'yes' response to the word pairs in both the causal and the associative condition.

The remaining two blocks consisted of eight weakly associated non-causal word pairs and three unrelated filler word pairs, preceded by the 'ASSOCIATIVE?' cue. These blocks were shams that were included to prevent participants from using assessment of causal relations as a strategy to judge word pairs in the associative condition. The order of trials within a block was randomized. The ten blocks were block-randomized into two sets of five blocks each. After each block there was a 16-s break followed by the next prompt.

At the beginning of each trial, a fixation cross appeared in the center of the screen. After 1000 ms, the cross disappeared and a blank screen was presented for 500 ms. The first word of the item pair was then presented for 1000 ms, followed by the second word (which replaced the first word). Thus, the interstimulus interval (ISI) was 0 and the stimulus onset asynchrony (SOA) was 1000 ms. The second word remained on the screen for up to 1500 ms until the participant pressed one of the two response keys. If participants viewed the item as related in the specified way, they were asked to press a response-box button underneath their index finger indicating a 'yes' response; otherwise they were to press a button underneath their middle finger indicating a 'no' response. Each trial lasted 4 s, which matched the TR of the functional imaging runs. Response times and errors were recorded by the Superlab® software. An initial practice session consisting of 26 practice trials (three associative, 12 causal, 11 unrelated) with feedback was run prior to the experimental blocks.

Imaging procedure

Upon obtaining consent, participants were given detailed instructions concerning the behavioral task described above. They were also told that at times a fixation cross would appear on the screen and that they should watch the cross until the next cue and set of word pairs appeared. Participants were instructed to respond as quickly as they could without sacrificing accuracy.

Participants were positioned on the scanner bed in a supine position. They wore earplugs to dampen scanner noise and headphones to enable communication with experimenters. Stimuli were presented through goggles connected to a Macintosh G3 computer. Behavioral responses were recorded using a two-button response box. To prevent head motion, participants' heads were restrained with foam padding and a piece of surgical tape was placed across each participant's forehead. Participants were instructed to lie as still as was comfortably possible, especially during task scans.

Once inside the scanner, approximately 8 min was spent obtaining preliminary scans to set scanner parameters and slice locations. Structural and functional scans were then obtained. Participants completed the behavioral task twice, corresponding to two functional runs, each lasting 5 min and 36 s with a rest interval of approximately 2 min between scans.

Functional imaging equipment and parameters

Imaging data were acquired using a 3-T GE whole-body MRI scanner with an upgrade for echo-planar imaging (EPI; Advanced NMR

Systems, Inc., Wilmington, MA, USA) and equipped with a standard head volume coil. For each participant, a high-resolution structural T2-weighted echo planar imaging volume (spin-echo; TR = 4000 ms; TE = 54 ms; matrix size 128 × 128; 26 axial slices; 3.125-mm in-plane resolution; 4 mm thick, skip 1 mm) was acquired and was co-planar with the functional scans. Functional scans (echo planar T2*-weighted gradient-echo, TR = 4 s, TE = 25 ms, flip angle = 90°, matrix size 64 × 64, 19 axial slices, 3.125-mm in-plane resolution; 4 mm thick, skip 1 mm, ascending slice sequence) spanned nearly the entire brain except some regions of the brain stem. Each scan consisted of 84 acquisitions.

Results

Reaction times and errors

Two 2 × 2 ANOVAs were conducted with condition and counterbalancing as independent factors, and reaction time (RT) for correct trials and accuracy as dependent variables. The analysis of reaction times showed no significant effects across conditions ($P > 0.2$). Accuracy data revealed only that participants made more errors in the causal judgment task than in the associative judgment task ($F_{1,11} = 8.99$, $MSE = 0.0264$, $P < 0.02$), indicating that the causal condition was more difficult than the associative condition. Mean RTs for causal and associative judgment conditions (± SD) were 924 ± 29 and 891 ± 34 ms, respectively. Mean error rates were 15% (0.031) and 8% (0.015), respectively.

Neuroimaging analyses

Neuroimaging data were processed using statistical parametric mapping (SPM'99; Wellcome Department of Imaging Neuroscience). Structural images were normalized to a standard T1 template image in MNI space provided by the SPM'99 software package using a 12-parameter affine transformation. Functional images were corrected for head motion using a six-parameter affine 'rigid-body' transformation, normalized (12-parameter affine transformation) to an EPI template image in MNI space provided by the SPM'99 software package, and smoothed using an 8-mm FWHM Gaussian kernel. The first three acquisitions of each scan were excluded from further analyses. A blocked design convolved with a hemodynamic response function was used to model the imaging data and contrast images were generated by running *t*-tests for each subject. Single subject contrast files were then gathered and subjected to *t*-tests for group comparisons.

Analyses were conducted at a threshold of $P < 0.001$ uncorrected, *k*-extent = 30. This combination of intensity and extent thresholds produces a per voxel false positive probability of < 0.000001 (Forman *et al*., 1995). MNI coordinates were converted to Talairach coordinates (Talairach & Tournoux, 1988; Brett, 1999). Because of technical difficulties, data from one participant were not included in the analyses.

Our goal was to decompose neural regions involved in causal processing into those that are common with and distinct from associative processing. It was hypothesized that causal processing would overlap with a broad network of structures involved in associative processing, but that that there would also be activations in DLPFC unique to causal processing.

In order to assess the brain regions that were active during both causal and associative processing, we carried out an inclusion analysis. An inclusion mask was created by finding those clusters of voxels that were more active for associative processing than fixation and then examining only these clusters in a second analysis of causal processing compared with fixation. In essence, a cluster of voxels must be significantly active in both associative and causal processing, relative to fixation, to be significant in the inclusion analysis. A number of regions that might be expected during a linguistic semantic task (Cabeza & Nyberg, 2000) were activated, including fusiform gyrus, anterior cingulate cortex, superior parietal lobule, and inferior and middle prefrontal gyri including some DLPFC activity, primarily in the left hemisphere (see Table 1 and Fig. 2).

To examine the regions unique to causal and associative processing, we directly compared the neural activity produced in these two conditions. The causal condition produced greater activation than the associative condition in left DLPFC in a region more anterior than those found in the inclusion analysis, and additionally produced activation in right precuneus (see Table 2 and Fig. 2). The associative condition produced greater activation than the causal condition in right superior temporal gyrus (STG; see Fig. 2). In comparison with the inclusion analysis, all of the activations in the causal vs. associative comparison were located outside of those areas that were common to both causal and associative processing, suggesting that these regions were distinct from the network of shared areas. In other words, it seems that the causal condition selectively produced more activity in this region (see Fig. 2).

It is impossible to determine on the basis of a main effect alone whether the activations found in the causal vs. associative task conditions were due to an increase in the causal condition, a decrease in the associative condition or some combination of both. Accordingly, we examined whether each condition alone vs. fixation produced activations or deactivations in the aforementioned areas. Although significant activations at the a-priori threshold were not found, a more lenient threshold ($P < 0.005$) revealed activation in DLPFC in the causal vs. fixation comparison [(Talaraich coordinates: −38,44,16), $k = 9$, $Z = 2.89$] but not in the associative minus fixation condition. Moreover, for this area of DLPFC the associative vs. fixation comparison was not significant at any threshold, and the area was less active during the associative condition than during fixation. Precuneus was more active in the causal condition than during fixation [(24,−61,33), $k = 33$, $Z = 3.71$] and more active during fixation than in the associative condition [(20,−63,31), $k = 10$, $Z = 3.00$]. Finally, STG activity in the associative minus causal condition comparison was due to a deactivation in the causal condition relative to fixation [(57,−55,19), $k = 13$, $Z = 2.83$] rather than to an increase during the associative trial blocks. Altogether, these analyses tentatively suggest that (1) activity in an anterior region of DLPFC was primarily due to increased activation during performance of the causal task; (2) STG activity was primarily due to decreased activation during performance of the causal task; and (3) precuneus activity was due to both increased

TABLE 1. Areas common to causal vs. fixation and associative vs. fixation comparisons

Location (BA)	*x*	*y*	*z*	*k*	*Z*
Left DLPFC (45/46)	−51	32	15	140	4.4
Left middle frontal gyrus (6)	−34	−5	54	151	4.6
Left inferior frontal gyrus (44)	−55	8	5	140	5.11
Left inferior frontal gyrus (44)	−48	9	29	290	4.60
Left inferior frontal gyrus (47)	−51	19	−8	173	3.54
Left superior parietal lobule (7)	−32	−53	58	52	4.56
Left anterior cingulate (24)	−4	2	46	153	4.09
Left fusiform (37)	−42	−66	−5	164	4.07
Left cerebellum	−44	−53	−19	253	4.93
Right cerebellum	28	−58	−29	307	5.18

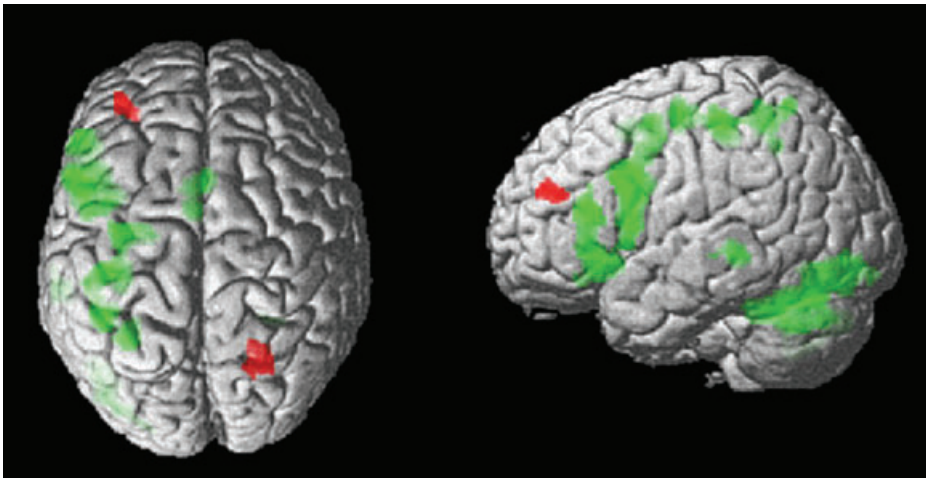

FIG. 2. Rendered image of regions common to causal and associative processing (green areas), left DLPFC and right precuneus showing more activity in the causal condition relative to the associative condition (red areas), and right superior temporal gyrus showing more activity in the associative condition than the causal condition (blue area).

TABLE 2. Locations of activations across causal vs. associative conditions

Comparison and location (BA)	*x*	*y*	*z*	*k*	*Z*
Causal–associative					
Left DLPFC (46)	−30	40	18	55	3.95
Left DLPFC (46)	−36	48	20	55	3.82
Right precuneus (7/19)	24	−61	31	96	4.03
Midbrain	−2	−22	−16	30	3.23
Causal–associative with mean RT difference regressed out					
Left DLPFC (46)	−30	40	18	42	3.87
Left DLPFC (46)	−36	48	20	42	3.62
Right precuneus (7/19)	24	−63	31	101	4.05
Right mid frontal gyrus (8)	51	12	38	38	4.17
Causal–associative with error rates regressed out					
Left DLPFC (46)	−36	48	22	80	4.08
Left DLPFC (46)	−30	48	20	80	3.91
Right precuneus (7/19)	24	−61	31	79	4.06
Associative–causal					
Right superior temporal gyrus (22)	61	−53	21	33	6.16
Associative–causal with mean RT difference regressed out					
Right superior temporal gyrus (22)	61	−53	21	35	6.07
Associative–causal with error rates regressed out					
Right superior temporal gyrus (22)	61	−53	21	32	5.84

activation during performance of the causal task and decreased activation during performance of the associative task.

Because task difficulty could potentially provide an alternative explanation for differences between the causal condition and the associative condition, we performed an analysis of covariance (ANCOVA) to remove statistically the effect of each subject's mean RT difference between the causal and associative task conditions. As before, the causal condition produced greater activation in the same areas of left DLPFC and right precuneus than did the associative condition (see Table 2). The associative condition continued to produce greater activity in right STG. A parallel ANCOVA with differences in error rates instead of reaction times as a covariate revealed a similar pattern of activations in the left DLPFC and right precuneus for the causal condition, and right STG for the associative condition. These results suggest that task difficulty cannot account for the results obtained.

Discussion

The present study is the first to examine the neural regions involved in accessing and evaluating causal relations. For identical word pairs, the results demonstrated a dissociation between causal and associative processing: causal processing led to increased activity in DLPFC and precuneus, and decreased activity in STG. Furthermore, these results withstood multiple corrections for task difficulty. Although causal and associative processing shared several regions of activity in common, including a posterior region of DLPFC, only the causal condition produced activations in an anterior region of DLPFC (see Fig. 2). This finding was still reliable after correcting for task difficulty as indexed by either RTs or error rates. The DLPFC activation, found in the present study to be unique to causal processing, fits well with previous evidence obtained with other high-level cognitive tasks. The DLPFC appears to be engaged by reasoning and working-memory

tasks that require the maintenance and manipulation of relations (Robin & Holyoak, 1995; Waltz *et al*., 1999; Morrison *et al*., 2004). The selective activation of DLPFC observed in the present study when making causal judgments is particularly compelling because it was obtained using identical word pairs that were judged under both causal and associative instructions. Our findings argue against associationist accounts of human causal knowledge (e.g. Corlett *et al*., 2004), which would predict that the neural substrate of causal judgments should be a subset of the regions involved in associative judgments.

The difference in causal and associative processing also led to activation in a lateral portion of the right precuneus. Precuneus has been hypothesized to be involved in multimodal episodic retrieval (Buckner *et al*., 1996; Kraus *et al*., 1999). In the present investigation, activity in the right precuneus was located further lateral to the midline than in most earlier studies. A similar area has been found to be associated with reasoning complexity, independent of task difficulty (Kroger *et al*., 2002).

In the present study, DLPFC and precuneus may be involved in the application of abstract causal relational structures. Even though both causal and associative judgments necessarily depend on semantic knowledge, causal judgments require additional neurocognitive resources in order to represent and manipulate the specific roles of a causal relation (i.e. cause and effect). By contrast, simple associative judgments do not require the processing of specific relational roles. Associative responses can be made on the basis of a spreading activation process that does not require explicit representation of roles and relations. In support of this hypothesis, the activations found in the anterior portion of DLPFC and precuneus were distinctly involved in the causal task, and were not active when performing the associative task, relative to a fixation baseline.

Because the present study focused only on causal relations, it is difficult to assess whether the activity found was specific to causal relations or more generally to all asymmetric relations (e.g. 'is bigger than', 'is a type of'). Further studies that include causal relations, non-causal asymmetric relations, and symmetric associative relations would be able to tease this possibility apart. However, a chief advantage of the present study is that the same word pairs were used across task conditions; it would be difficult to use a similar design to compare causal relations with other non-causal asymmetric relations because word pairs would have to satisfy each of the relations being tested.

Owing to the preliminary nature of the present study and the dearth of neuroimaging studies of causal reasoning, many alternative accounts of the DLPFC activity are possible. For instance, this region may reflect response inhibition. In particular, during the causal task participants may attempt to inhibit responding on the basis of associativity prior to applying a causal schema (Deacon, 1997). This additional response inhibition may underlie the increased RT in the causal condition. However, even after regressing RT differences out, the neural differences were still robust, suggesting that response inhibition cannot account for the present results. Another possibility is that DLPFC activity is associated with evaluating theory-inconsistent data (Fugelsang & Dunbar, 2005). This explanation does not easily account for the current data because most of the word pairs used (eight pairs out of 11 per block) were theory-consistent within a block. Moreover, regressing out error rates, which may in part reflect trials in which participants believed the stimulus to be theory-inconsistent, did not appreciably alter the activity in DLPFC.

The results of the present study contribute to the recent body of work that explores the neural basis of causal reasoning (Fugelsang & Dunbar, 2005). Given the central importance of causal judgments in prediction of future events and decisions about appropriate interventions, causal relations constitute an especially significant vehicle for ascertaining how neural processes carry out the access and application of relational structures. An understanding of causality, the 'cement of the universe' (Kant, 1781/1965; Mackie, 1974), may provide key insights towards a mechanistic understanding of the neural basis of human reasoning.

Acknowledgements

We thank the staff at the UCLA Brain Mapping Center for their assistance and Brooke McGowan for comments on an earlier draft. This research was funded by a McDonnell-Pew Cognitive Neuroscience Grant (JSMF 99-25 CN-QUA.05) awarded to Matthew Lieberman and a Deutscher Akademischer Austauschdienst HSP III Grant awarded to Daniela Fenker. We also appreciate the support provided by the Brain Mapping Medical Research Organization, Brain Mapping Support Foundation, Pierson–Lovelace Foundation, the Ahmanson Foundation, Tamkin Foundation, Jennifer Jones-Simon Foundation, Capital Group Companies Charitable Foundation, Robson Family, William M. and Linda R. Dietel Philanthropic Fund at the Northern Piedmont Community Foundation, Northstar Fund, and National Center for Research Resources grants RR12169, RR13642 and RR08655. Preliminary reports of this research were presented at the 2004 annual meeting of the Cognitive Neuroscience Society (San Francisco) and at the 28th International Congress of Psychology (Beijing, 2004).

Abbreviations

DLPFC, dorsolateral prefrontal cortex; EPI, echo-planar imaging; fMRI, functional magnetic resonance imaging; ISI, interstimulus interval; RT, reaction time; SOA, stimulus onset asynchrony; STG, superior temporal gyrus; USF, University of South Florida.

References

Brett, M. (1999) The MNI brain and the Talairach atlas. http://www.mrc-cbu.cam.ac.uk/Imaging/mnispace.html.

Buckner, R.L., Raichle, M.E., Miezin, F.M. & Petersen, S.E. (1996) Functional anatomic studies of memory retrieval for auditory words and visual pictures. *J. Neurosci.*, **10**, 6219–6235.

Buehner, M.J. & Cheng, P.W. (2005) Causal learning. In Holyoak, K.J. & Morrison, R.G. (Eds), *Cambridge Handbook of Thinking and Reasoning*. Cambridge University Press, Cambridge, pp. 143–168.

Bunge, S.A., Wendelken, C., Badre, D. & Wagner, A.D. (2005) Analogical reasoning and prefrontal cortex: evidence for separable retrieval and integration mechanisms. *Cereb. Cortex*, **15**, 239–249.

Cabeza, R. & Nyberg, L. (2000) Imaging cognition II: an empirical review of 275 PET and fMRI studies. *J. Cognitive Neurosci*. **12**, 1–47.

Christoff, K., Prabhakaran, V., Dorfman, J., Zhao, Z., Kroger, J.K., Holyoak, K.J. & Gabrieli, J.D.E. (2001) Rostrolateral prefrontal cortex involvement in relational integration during reasoning. *Neuroimage*, **14**, 1136–1149.

Cohen, J.D., Forman, S.D., Braver, T.S., Casey, B.J., Servan-Schreiber, D. & Noll, D.C. (1994) Activation of the prefrontal cortex in a nonspatial working memory task with functional MRI. *Hum. Brain Mapp.*, **1**, 293–304.

Corlett, P.R., Aitken, M.R.F., Dickinson, A., Shanks, D.R., Honey, G.D., Honey, R.A.E., Robbins, T.W., Bullmore, E.T. & Fletcher, P.C. (2004) Prediction error during retrospective revaluation of causal associations in humans: fMRI evidence in favor of an associative model of learning. *Neuron*, **44**, 877–888.

Deacon, T.W. (1997) *The Symbolic Species: the Co-Evolution of Language and the Brain*. W. W. Norton, New York.

Fenker, D.B., Waldmann, M.R. & Holyoak, K.J. (in press) Accessing causal relations in semantic memory. *Mem. Cognition*, in press.

Forman, S.D., Cohen, J.D., Fitzgerald, M., Eddy, W.F., Mintun, M.A. & Noll, D.C. (1995) Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. *Magnet. Reson. Med.*, **33**, 636–647.

Fugelsang, J.A. & Dunbar, K.E. (2005) Brain-based mechanisms underlying complex causal thinking. *Neuropsychologia*, **43**, 1204–1213.

Fugelsang, J.A., Roser, M.E., Corballis, P.M., Gazzaniga, M.S. & Dunbar, K.E. (2005) Brain mechanisms underlying perceptual causality. *Cognitive. Brain. Res.*, **24**, 41–47.

Gabrieli, J.D., Poldrack, R.A. & Desmond, J.E. (1998) The role of left prefrontal cortex in language and memory. *Proc. Natl. Acad. Sci. USA*, **95**, 87–115.

Goel, V., Gold, B., Kapur, S. & Houle, S. (1997) The seat of reason: a localization study of deductive and inductive reasoning using PET (O15) blood flow technique. *Neuroreport*, **8**, 1305–1310.

Hummel, J.E. (1999) Binding problem. In Wilson, R.A. & Keil, F.C. (Eds), *The MIT Encyclopedia of the Cognitive Sciences*. MIT Press, Cambridge, MA, pp. 85–86.

Hummel, J.E. & Holyoak, K.J. (2003) A symbolic-connectionist theory of relational inference and generalization. *Psych. Rev.*, **110**, 220–264.

Kant, I. (1781/1965) *Critique of Pure Reason*. Macmillan, London.

Kraus, B.J., Schmidt, D., Mottaghy, F.M., Taylor, J., Halsband, U., Herzog, H., Tellmann, L. & Muller-Gartner, H.W. (1999) Episodic retrieval activates the precuneus irrespective of the imagery content of word pair associates. *Brain*, **122**, 255–263.

Kroger, J.K., Saab, F.W., Fales, C.L., Bookheimer, S.Y., Cohen, M.S. & Holyoak, K.J. (2002) Recruitment of anterior dorsolateral prefrontal cortex in human reasoning: a parametric study of relational complexity. *Cereb. Cortex*, **12**, 477–485.

Lieberman, M.D., Gaunt, R., Gilbert, D.T. & Trope, Y. (2002) Reflexion and reflection: a social cognitive neuroscience approach to attributional inference. *Adv. Exp. Soc. Psychol.*, **34**, 199–249.

Mackie, J.L. (1974) *The Cement of the Universe*. Oxford University Press, Oxford.

Mason, R.A. & Just, M.A. (2004) How the brain processes causal inferences in text. *Psychol. Sci.*, **15**, 1–7.

Medin, D.L. & Rips, L.J. (2005) Concepts and categories: memory, meaning, and metaphysics. In Holyoak, K.J. & Morrison, R.G. (Eds) *Cambridge Handbook of Thinking and Reasoning*. Cambridge University Press, Cambridge, pp. 37–72.

Morrison, R.G., Krawczyk, D.C., Holyoak, K.J., Hummel, J.E., Chow, T.W., Miller, B.L. & Knowlton, B.J. (2004) A neurocomputational model of analogical reasoning and its breakdown in Frontotemporal Lobar Degeneration. *J. Cognitive Neurosci.*, **16**, 260–271.

Nelson, D.L., McEvoy, C.L. & Schreiber, T.A. (1998) The University of South Florida word association, rhyme, and word fragment norms. http://www.usf.edu/FreeAssociation/.

Robin, N. & Holyoak, K.J. (1995) Relational complexity and the functions of prefrontal cortex. In Gazzaniga, M.S. (Ed.), *The Cognitive Neurosciences*. MIT Press, Cambridge, MA, pp. 987–997.

Smith, E.E. & Jonides, J. (1999) Storage and executive processes in the frontal lobes. *Science*, **283**, 1657–1661.

Spellman, B.A. (1997) Crediting causality. *J. Exp. Psychol. Gen.*, **126**, 323–348.

Spellman, B.A., Holyoak, K.J. & Morrison, R.G. (2001) Analogical priming via semantic relations. *Mem. Cognition*, **29**, 383–393.

Talairach, J. & Tournoux, P. (1988) *Coplanar Stereotaxic Atlas of the Human Brain.* Thieme Medical, New York.

Turner, D.C., Aitken, M.R.F., Shanks, D.R., Sahakian, B.J., Robbins, T.W., Schwarbauer, C. & Fletcher, P.C. (2004) The role of the lateral frontal cortex in causal associative learning: exploring preventative and super-learning. *Cereb. Cortex*, **14**, 872–880.

Wagner, A.D., Koustaal, W., Maril, A., Schacter, D.L. & Buckner, R.L. (2000) Task-specific repitition priming in left inferior prefrontal cortex. *Cereb. Cortex*, **10**, 1176–1184.

Waldmann, M.R. (1996) Knowledge-based causal induction. In Shanks, D.R., Holyoak, K.J. & Medin, D.L. (Eds), *The Psychology of Learning and Motivation*, **Vol. 34**: Causal Learning. Academic Press, San Diego, CA, pp. 47–88.

Waltz, J.A., Knowlton, B.J., Holyoak, K.J., Boone, K.B., Mishkin, F.S., de Menezes Santos, M., Thomas, C.R. & Miller, B.L. (1999) A system for relational reasoning in human prefrontal cortex. *Psychol. Sci.*, **10**, 119–125.