NEURAL CORRELATES OF RELATIONAL REASONING AND THE SYMBOLIC DISTANCE EFFECT: INVOLVEMENT OF PARIETAL CORTEX

E. C. HINTON,[a]* S. DYMOND,[b] U. VON HECKER[c] AND C. J. EVANS[d]

[a]*Wales Institute of Cognitive Neuroscience, School of Psychology, Cardiff University, Park Place, Cardiff, CF10 3AT Wales, UK*

[b]*Department of Psychology, Swansea University, Singleton Park, Swansea, SA2 8PP Wales, UK*

[c]*School of Psychology, Cardiff University, Park Place, Cardiff, CF10 3AT Wales, UK*

[d]*CUBRIC, School of Psychology, Cardiff University, Park Place, Cardiff, CF10 3AT Wales, UK*

Abstract—A novel, five-term relational reasoning paradigm was employed during functional magnetic resonance imaging to investigate neural correlates of the symbolic distance effect (SDE). Prior to scanning, participants learned a series of more-than (E>D>C>B>A) or less-than (A<B<C<D<E) ordered premise pairs. During scanning, inferential tests presented the premise pairs, adjacent, mutually entailed tasks (e.g., D<E and B>A) and nonadjacent one-step (A<C, B<D, C<E, C>A, D>B and E>C) and two-step (A<D, B<E, D>A and E>B) combinatorial entailed tasks. In terms of brain activation, the SDE was identified in the inferior frontal cortex, dorsolateral prefrontal cortex, and bilateral parietal cortex with a graded activation pattern from adjacent to one-step and two-step relations. We suggest that this captures the behavioural SDE of increased accuracy and decreased reaction time from adjacent to two-step relations. One-step relations involving endpoints A or E resulted in greater parietal activation compared to one-step relations without endpoints. Novel contrasts found enhanced activation in right parietal and prefrontal cortices during mutually entailed tasks only for participants who had learned all less-than relations. Increased parietal activation was found for one-step tasks that were inconsistent with prior training. Overall, the findings demonstrate a crucial role for parietal cortex during relational reasoning with a spatially ordered array. © 2010 IBRO. Published by Elsevier Ltd. All rights reserved.

Key words: fMRI, transitive inference, relational reasoning, symbolic distance, more-than, less-than.

A fundamental ability shared by humans and animals is the capacity to make flexible inferences on the basis of newly learned information (Vasconcelos, 2008). This ability is often studied using transitive inference (TI) paradigms, a type of reasoning whereby, for example, if A>B and B>C, then A>C. There is a long history of employing TI based paradigms to study reasoning (Riley, 1976). One such paradigm is the serial order or *n*-term series TI task, in which participants receive training to order stimuli into hierarchical order on the basis of adjacent stimulus pairs alone (e.g., A>B, B>C, C>D, D>E: A-B-C-D-E). At test, the correct answer to novel non-adjacent relations, such as AC and BD, can then be inferred.

Inferential test performance, it is argued, relies on the use of an integrated, or holistic representation or array, containing all information concerning the ordered relations between the pairs of stimuli presented in training (Leth-Steensen and Marley, 2000). Evidence of the formation of such an integrated ordered array is shown through the symbolic distance effect (SDE; Acuna et al., 2002; Colombo and Frost, 2001; Frank et al., 2005; Zalesak and Heckers, 2009) whereby accuracy increases and response latencies decrease the further apart stimulus pairs are along the array. For example, at test participants are more accurate and faster to respond to AD (a symbolic distance of two-steps) than AC (symbolic distance of one-step). The main objective of the present research is to demonstrate the neural correlates of the SDE, in order to further understand how manipulating integrated representations during reasoning is processed by the brain. The paradigm used here facilitates the construction of such a representation by using a novel TI procedure from the behavioral literature that provides a dimensional reference along which stimuli may be ordered. To our knowledge, the only study to have demonstrated the neural correlates of a behavioural SDE (Zalesak and Heckers, 2009) used a paradigm that relied less on dimensional instructions and stimuli, and concentrated its analysis on the role of the hippocampus.

Our aim is to study a network of brain regions previously shown to support spatial processing of items in an integrated array during TI (Davis, 1992; Van Opstal et al., 2008). Studies to date have focused on imaging different aspects of TI, including both training and test phases (Nagode and Pardo, 2002) with a focus on hippocampal activation during specific (i.e., BD) trials (Greene et al., 2006), emergence of inferences during test (Van Opstal et al., 2008), inference vs. support processes at test (Acuna et al., 2002), differential activation associated with pairs separated by symbolic distance of one vs. two steps (Heckers et al., 2004; Zalesak and Heckers, 2009), and the effects of

*Corresponding author. Tel: +44-0-29-208-70717; fax: +44-0-29-208-74858.

E-mail address: hintone@cardiff.ac.uk (E. C. Hinton).

extensive training on left angular gyrus activation (Van Opstal et al., 2009). Much of the TI research has demonstrated the important role of the hippocampus in such processing (Greene et al., 2006; Heckers et al., 2004; Nagode and Pardo, 2002; Preston et al., 2004; Schlund et al., 2008; Zalesak and Heckers, 2009); however, the wider reasoning literature, using TI and other paradigms, has also implicated the prefrontal cortex (relational integration and complexity and working memory) and parietal cortex (spatial and attentional processing, and working memory) as key regions involved in making inferences based on integrated representations of ordered stimuli (Acuna et al., 2002; Christoff et al., 2001; Goel and Dolan, 2001; Heckers et al., 2004; Van Opstal et al., 2008, 2009; Wendelken et al., 2008).

In this study, we employed a novel 5-term series paradigm during functional magnetic resonance imaging (fMRI). As explained in more detail below, this paradigm has two main advantages over TI paradigms that have been used so far to investigate the SDE (Reilly et al., 2005; Whelan et al., 2006). First, the present paradigm facilitates the use of an analogue quantitative dimension (more–less) in order to construct an integrated, ordered representation of the stimuli. This is advantageous in order to investigate activation patterns in brain areas, such as the parietal cortex, known to be associated with spatial ordering functions. Second, by using a continuous quantitative dimension, the present paradigm facilitates the use of novel mutually entailed (e.g., if B>A, then A<B; see Table 1) and combinatorial relations (e.g., if B>A and C>B, then A<C; see Table 1) not usually studied in the TI literature. Testing for such reciprocal relations is necessary to establish the generality of a constructed order representation.

According to relational frame theory (Hayes et al., 2001), a contemporary behavioral theory of language and cognition, successful TI performance involves deriving/inferring comparative relations. Derived comparative relations are involved whenever one event is responded to in terms of a quantitative or qualitative relation along a specified dimension with another event, and have been most widely studied with more-than/less-than relations (e.g., Berens and Hayes, 2007; Vitale et al., 2008; Whelan et al., 2006). First learned with nonarbitrary stimuli that differ along a specified physical dimension, such as size, comparative relations may be applied to any arbitrary stimuli, given appropriate contextual cues. For instance, consider a young child who learns that "X is more than Y." Subsequently, he or she may when asked, "which is less than?" respond "Y," without any further training. This arbitrary relational response is controlled solely by the contextual cues "more than" and "less than," not by any physical relations, and can be applied to any stimuli regardless of their physical properties. In this way, derived comparative relations may be involved in TI-like reasoning processes. For example, if we learn that A is more than B (A>B) and B is more than C (B>C), then we may derive that B is less than A (B<A), C is less than B (C<B), A is more than C (A>C) and C is less than A (C<A). Participants come to apply their relational learning histories of more than and less than, newly established with the contextual cues so as to not be influenced by previous learning, to a series of arbitrary stimuli (e.g., A>B, etc.) in a way that resembles premise pair learning from TI research.

More-than relations may be easier to learn than less-than relations because they usually appear before less-than relations in young children's repertoires and therefore may remain stronger into adulthood (Hayes et al., 2001). Moreover, Reilly et al. (2005) demonstrated that response latencies to test-trials were significantly faster following training with more-than relations than following training with less-than relations. We therefore included training history (more-than vs. less-than) as a control factor in the present study. Moreover, the inclusion of training history provided a means to investigate any differences in performance and activation arising from deriving comparative relations that contain cues which are the same as or different from that training history. In the present study, the stimulus relations were established intra-experimentally in order to overcome participants' pre-existing learning histories with more than and less than relations, and to ensure the inferences originated from this learning history rather than instructions or observational learning.

The specific goals of this study were three-fold. First, we investigated the neural correlates of the SDE in more detail than has previously been addressed. At test, the longest response latencies and lowest accuracy were pre-

Table 1. Test trials employed during arbitrary relational testing (phase 2) in the all-more and all-less groups

Group	Relation type	Test trial type					
		Specific relations in each group					
All-more	Trained	B>A	C>B	D>C	E>D		
	ME	A<B	B<C	C<D	D<E		
All-less	Trained	A<B	B<C	C<D	D<E		
	ME	B>A	C>B	D>C	E>D		
		Relations common to both groups					
	CE1	A<CE	B<D	C<E	C>A	D>B	E>C
	CE2	A<DE	B<E	D>A	E>B		

"Trained" refers to test trials involving directly trained relations, and the acronym ME, CE1 and CE2 refer to test trials for mutually entailed, and one- and two-node combinatorially entailed relations respectively. The inequality symbols, < (less-than) and > (more-than), denote the contextual cue that was presented. This indicates which comparison should be "selected over" the other, with the reinforced comparison to the left, and the non-reinforced comparison to the right of the inequality symbol.

dicted for trained relations (e.g., A<B), followed by one-step (e.g., A<C) tasks with the fastest responses and greatest accuracy to two-step tasks (e.g., A<D). It was predicted that this distance effect would be reflected in activation in regions involved in TI, namely prefrontal and parietal cortices. Parietal cortex activation was predicted due to its widely accepted involvement in spatial processing of items in an array (e.g., Cohen Kadosh et al., 2008; Dehaene et al., 2003; Pinel et al., 2001; Van Opstal et al., 2008, 2009).

A second aim was the investigation of brain activation during relational tasks with (e.g., A<C, C<E) and without (e.g., B<D) endpoints. There is evidence to suggest that relations containing an endpoint will be processed differently from those that do not contain an endpoint (Leth-Steensen and Marley, 2000). Particularly, endpoints are unique with regards to their reinforcement history because they are always either reinforced or non-reinforced, and may therefore be more salient than stimuli within the linear order that are not endpoints (Heckers et al., 2004; Vasconcelos, 2008; Zalesak and Heckers, 2009).

A third aim was to test whether relational training histories using more-than or less-than relations would result in differential behavioural performance and to explore whether differential neural activity would be associated with using and manipulating the resulting arrays ordered by more than relations versus less than relations when deriving novel comparative relations. If found, differences were hypothesized to occur in a distributed cortical network, shown in previous studies to include the parietal and prefrontal cortex (Heckers et al., 2004; Van Opstal et al., 2008).

EXPERIMENTAL PROCEDURES

Participants

Thirty-four participants were recruited and randomly assigned into two groups for task training purposes; All-More and All-Less. Three participants in the All-More group and seven in the All-Less group did not meet the task mastery criterion and therefore did not continue to the imaging session. Twenty-four participants remained in the study for the fMRI session, 12 in each group; All-More group (m/f=1/11; mean age=24.1 years) and the All-Less group (m/f=3/9; mean age=22.9 years). There was no significant difference in age between the two groups. All participants were right-handed, none had any history of psychiatric or neurological disorders, and none were currently taking psychotropic medications. All participants gave informed consent. The Cardiff University School of Psychology Ethics Committee approved this study.

Stimuli

Two visual stimuli were employed as contextual cues during nonarbitrary relational training and testing to establish the contextual functions of more-than and less-than (see Fig. 1). During nonarbitrary training and testing, the stimulus sets employed consisted of quantities of objects termed *Few* for the smallest amount, *Intermediate* amount (note: not necessarily the midpoint of the smallest and greatest amounts), and *Many* for the greatest amount. For example, one stimulus set was composed of images of one, two, and three tractors. Four stimulus sets were used during nonarbitrary relational training, and were as follows (the
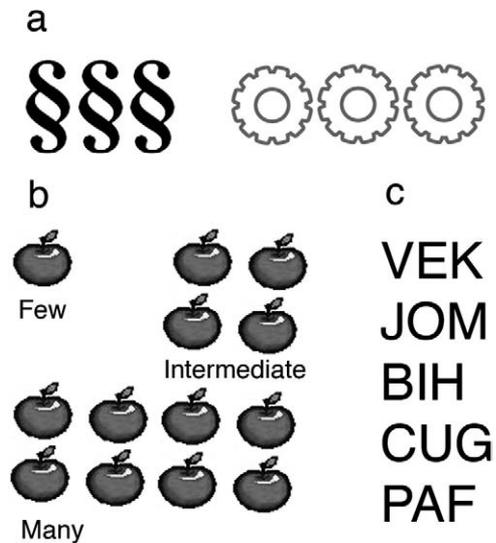

Fig. 1. Stimuli presented during the course of the task: (a) contextual cues for less-than (left-hand image) and more-than (right-hand image); (b) example stimulus set for the nonarbitrary phase (few, intermediate and many); (c) nonsense syllables used in the arbitrary phase [VEK (A), JOM (B), BIH (C), CUG (D), PAF (E)].

quantities of the particular object that composed each image are in parentheses): basketballs (one, two, eight), beakers (one, three, six), tractors (one, two, three), and ladybirds (two, four, eight). Another four stimulus sets were used during nonarbitrary relational testing: turtles (two, three, four), arks (one, two, three), apples (one, four, eight), and traffic lights (one, three, four). Each training and testing set was composed of three stimulus images and two contextual cues, which generated the following six discriminations: LESS THAN [Few/Intermediate], LESS THAN [Few/Many], LESS THAN [Intermediate/Many], MORE THAN [Intermediate/Few], MORE THAN [Many/Intermediate], and MORE THAN [Many/Few]. A total of 24 trial types were generated from these stimulus sets.

In addition, five consonant-vowel-consonant strings (VEK, JOM, BIH, CUG, PAF) were used during arbitrary relational training and testing phases (Fig. 1c). The strings were labelled as follows: A (VEK), B (JOM), C (BIH), D (CUG), and E (PAF). A 5-series linear relational network was subsequently established by presenting the training pairs AB, BC, CD and DE.

Relational reasoning task: training prior to scanning

Participants were trained on the relational reasoning task prior to the functional imaging. The procedure consisted of two phases, each with interlinked training and testing elements. The task was programmed in Presentation (Neurobehavioral systems, http://www.neurobs.com), which controlled all stimulus presentations and recorded all responses.

Phase 1: nonarbitrary relational training and testing. The purpose of this phase was to establish contextual control over responding by two cues (i.e. more than and less than; Fig. 1a) using stimulus sets of differing physical quantities (example in Fig. 1b). At the start of each trial, one of the contextual cues appeared in the centre, top-third of the screen. After a 1.5 s delay, the two comparison stimuli appeared simultaneously in the lower third of the left- and right-hand side of the screen (positions were counterbalanced across trials). Participants made selections by pressing either the "Z" or "/" keys. When the contextual cue for MORE THAN was presented, choosing the comparison stimulus with the greater quantity produced the feedback "Correct!," while choosing the comparison stimulus with the lesser quantity produced the

feedback "Wrong". When the contextual cue for LESS THAN was presented, choosing the comparison stimulus with the lesser quantity produced the feedback "Correct!," while choosing the comparison stimulus with the greater quantity produced the feedback "Wrong". Feedback was displayed for 1.5 s, and an intertrial interval (ITI) of 1.5 s followed. The contextual cue and both comparisons remained on the computer screen until a response was made. Participants were exposed to this nonarbitrary relational training phase until they made 10 consecutive correct responses.

On reaching this criterion, participants were immediately exposed to the nonarbitrary relational test with the four novel stimulus sets presented in the absence of feedback. Participants were exposed to the nonarbitrary relational test until they made 10 consecutive correct responses. If this criterion was not met within 24 trials, participants were re-exposed to nonarbitrary relational training, followed again by nonarbitrary relational testing.

Phase 2: arbitrary relational training and testing. Phase 2 began immediately following phase 1. The contextual cues first appeared in the centre, top-third of the screen followed by the two comparison stimuli, which consisted of nonsense syllables (Fig. 1c, labelled in the interests of clarity as A (VEK), B (JOM), C (BIH), D (CUG), and E (PAF)). A total of four training tasks were presented to each group. The All-More group were presented with training tasks B>A, C>B, D>C, and E>D (where ">" describes the contextual cue for the reinforced relation "more than") designed to establish the predicted relational network E>D>C>B>A. The All-Less group were presented with training tasks: A<B, B<C, C<D, and D<E (where "<" describes the contextual cue for the reinforced relation "less than") designed to establish the predicted relational network A<B<C<D<E. Participants in both groups were presented with the four training tasks plus a null event (fixation cross) in a quasi-random order, three times each in a block of 15 trials. Participants were exposed to the arbitrary relational training tasks until they made 12 consecutive correct responses.

On reaching this criterion, participants were exposed to the arbitrary relational test during which the four previously trained tasks and 14 novel test tasks were presented once each, in the absence of feedback. Table 1 lists the trial-types presented during this phase. Both groups were tested for mutual (reversed) stimulus relations based on their unique training history, and all received the same one-step and two-step combinatorial entailment tasks. The end-point pairs A<E or E>A were not presented since selections of either A or E were always reinforced in the respective training groups (Vasconcelos, 2008).

The mastery criterion for the arbitrary relational test was a minimum of 15 correct responses out of a block of 18 trials. If this criterion was not met, participants were re-exposed to phase 1 and phase 2 a maximum of three further times. If participants met the mastery criterion for the arbitrary relational test within the above training, they proceeded to the brain scan.

Relational reasoning task: test during fMRI scan

Since the imaging acquisition followed this brief exposure to the inferential task, the activation patterns produced were associated with maintenance of derived relations rather than their initial acquisition (see Schlund et al., 2008) because participants had to pass the behavioural test phase in order to proceed to the scanning session. During the functional imaging, participants were exposed to the arbitrary relational test phase as presented out of the scanner (see Table 1). A trial started with 1.5 s for the contextual cue, followed by the cue and the comparison stimuli for 2 s, with an ITI of 1.5 s. Participants made selections by pressing one of two buttons on an MR compatible button box. A block was made up of 24 test trials: all the possible adjacent trials (trained and mutually entailed) were presented once (eight trials), as were the AC and CE one-step trials (four trials). BD (four trials) and the two-step trials were presented twice (eight trials), in order to equalize the number of trials for our key comparisons (e.g., AC/CE compared to BD). Trials were presented in a random order in a block, along with eight null trials (a fixation "X" on a black background) that formed the baseline. Six blocks of the test trials were presented during the scan, with a jittered fixation cross of 0.5–2 s between blocks. The functional run lasted approximately 15 min, followed by an anatomical scan for a further 7:20 minutes.

Following the imaging session, participants were given a post-imaging questionnaire, designed to ascertain how participants reported solving the task and their awareness of the hierarchy within the task.

Image acquisition

Anatomical and functional images were acquired at the Cardiff University Brain and Research Imaging Centre (CUBRIC), using a General Electric 3 Tesla HDx MRI scanner. Functional images were collected using a gradient-echo echo-planar pulse sequence (TE=35 ms; TR=2500 ms; flip angle=90°; acquisition matrix=64×64; field of view (FOV) 240×240 mm^2; in plane resolution 3.4 mm). The volumes covered the whole brain in 45 slices (thickness 2.8 mm with a 1.0 mm slice gap) and were acquired in line with the anterior commissure/posterior commissure line. A total of 360 volumes were acquired for each participant. Finally, a high-resolution T1-weighted FSPGR anatomical image was acquired (TR=7.9 s; TE=3 ms; inversion time=450 ms; flip angle=20°; acquisition matrix 256×192×176; FOV 256×192×176 mm^3, resulting in 1 mm isotropic voxels).

Data analysis

Data were analysed using the FSL package from FMRIB, University of Oxford (www.fmrib.ox.ac.uk/fsl/). Data were acquired in one run for each participant. Each run was pre-processed and analysed separately, using the following stages: motion correction using MCFLIRT (Jenkinson et al., 2002), non-brain removal using BET (Smith, 2002), spatial smoothing using a Gaussian kernel of FWHM 5 mm, mean-based intensity normalization of all volumes, and high-pass temporal filtering. Time-series statistical analysis was carried out using FILM with local autocorrelation correction (Woolrich et al., 2001).

An event related analysis was performed on the data. Each event began at the onset of the comparison stimuli for that test trial. Only trials correctly answered within 2.3 s from stimulus onset were included in the analysis of the imaging data, resulting in 11% of the total number of trials being excluded. At the first level, each relation type was compared to baseline (fixation) (i.e., trained—fixation, mutually entailed—fixation, one-step-fixation, and two-step-fixation). A specific contrast was performed to investigate the SDE in the fMRI data: as greater activation was expected for trained relations, with linearly decreasing activation for one-step, then two-step relations, the contrast looked for areas of activation showing a linear trend in that direction (contrast values: trained 1, one-step 0, two-step −1). For the comparison between one-step trials with and without end-points, a contrast between AC/CE—BD trials was performed. To investigate differences between trial types from the relational contexts of more than and less than, additional contrasts were performed: (i) trained-mutually entailed; (ii) one-step less than—one-step more than; (iii) two-step less than—two-step more than.

At the second level, data from all participants was combined into a higher-level group analysis for each contrast, using a mixed effects group analysis—FLAME (stage 1 only) (Beckmann et al., 2003; Woolrich et al., 2004). Z statistic images were thresholded using Gaussian Random Field (GRF)-theory based maximum cluster thresholding with a corrected significance threshold of P=0.05 and cluster-corrected at z>2.3 (Worsley et al., 1992). Registration to high resolution and standard images was carried out using FLIRT (Jenkinson and Smith, 2001; Jenkinson et al., 2002). Data are presented in MNI space.

Due to the previous focus on hippocampal activation during TI tasks, a region of interest analysis was conducted using the left and right hippocampal masks from the Harvard–Oxford Subcortical Structural Atlas to investigate the role of the hippocampus during the present relational reasoning paradigm. Percentage signal change was extracted from this region during the different relations for each person using Featquery.

RESULTS

The results were analysed according to three main aims: to examine the neural basis of the SDE, to compare trials with and without endpoints, and to investigate the effect of the different training histories of more than and less than on behavioural performance and associated brain activation.

Behavioural data

Training.   Only data from those who took part in the fMRI scan are included in the following analyses ($n=24$). The mean (and SD) number of training trials required for each group to pass each phase were as follows: Nonarbitrary training All-more=161.4 (287.4), All-less=68.2 (45.5); Nonarbitrary test All-more=34.7 (12.4), All-less=32.8 (16.7); Arbitrary training All-more=208.0 (95.1), All-less=172.0 (78.7); Arbitrary test All-more=49.5 (17.4), All-less=48.0 (17.7). There was no significant difference between the All-more and All-less groups on any of the training measures.

Accuracy.   Fig. 2a shows the mean percentage correct scores for each relation type between training groups during the inferential test phase. A $2\times3\times2$ ANOVA (training history: All-more or All-less; inferential step: none (e.g., A<B), one (e.g., A<C), two (e.g., A<D); relation: same as

or different to training) was conducted using the percentage correct scores on the task. A significant main effect of inferential step was found ($F(2,44)=22.69$, $P<.001$), whereby accuracy increased the further apart the elements were along the array. Pairwise comparisons showed there was a significant difference between no-steps (i.e., adjacent) and one-step ($P<.001$), and no-steps and two-steps ($P<.001$), but no significant difference between one- and two-steps ($P=.504$). There was no significant difference in accuracy between the two training groups, or between relations same as or different to training, and no significant interactions.

Accuracy on mutually entailed trials was highly correlated with performance on one-step trials for the All-Less group only ($r^2=.629$; $r=.793$, $P=0.001$). Further analyses for this group revealed that performance accuracy on C>B and D>C mutually entailed trials correlated with D>B one-step combinatorial entailed trials ($r=.819$, $P=.001$; $r=.709$, $P=.001$, respectively). This finding that performance on two of the relevant mutually entailed trials was highly correlated with performance on the related one-step task mirrors a similar finding by Greene et al. (2006, p. 1162) who found that BD performance at test was correlated with BC and CD performance at training.

Reaction time.   Fig. 2b shows the mean reaction time to answer each relation type between training groups. Only correct responses were included in this analysis. A $2\times3\times2$ ANOVA (training history: All-More or All-Less; inferential step: none (e.g., A<B), one (e.g., A<C), two (e.g., A<D); relation: same as or different to training) was conducted using the RTs. A significant main effect of inferential step
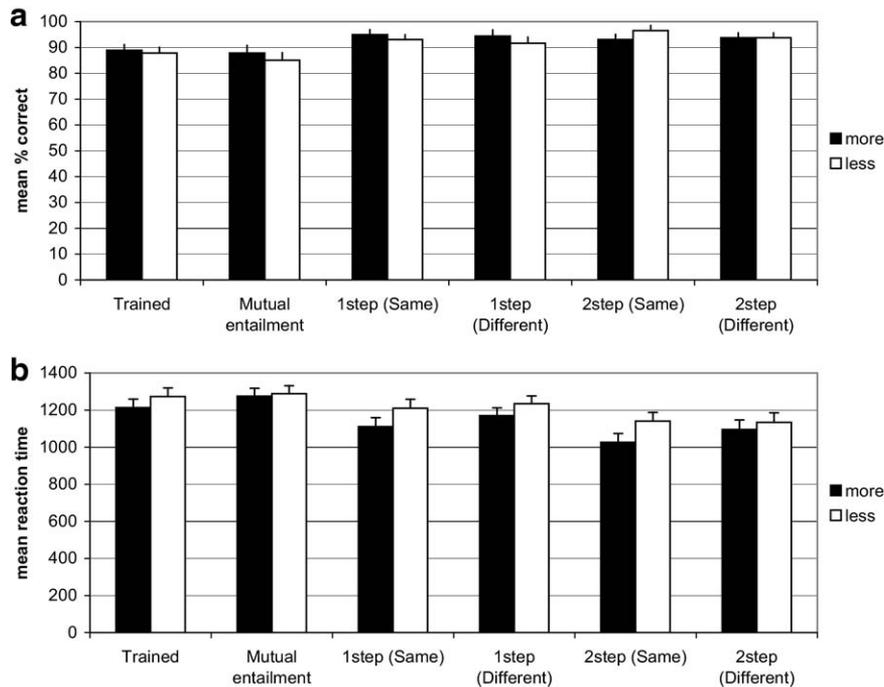

Fig. 2. Mean accuracy and latency for each relational type. The mean percentage correct and reaction time to each relational task for each training group. "Same" refers to relational tasks with contextual cues same as training and "Different" refers to relational tasks different to training. Error bars represent standard errors.
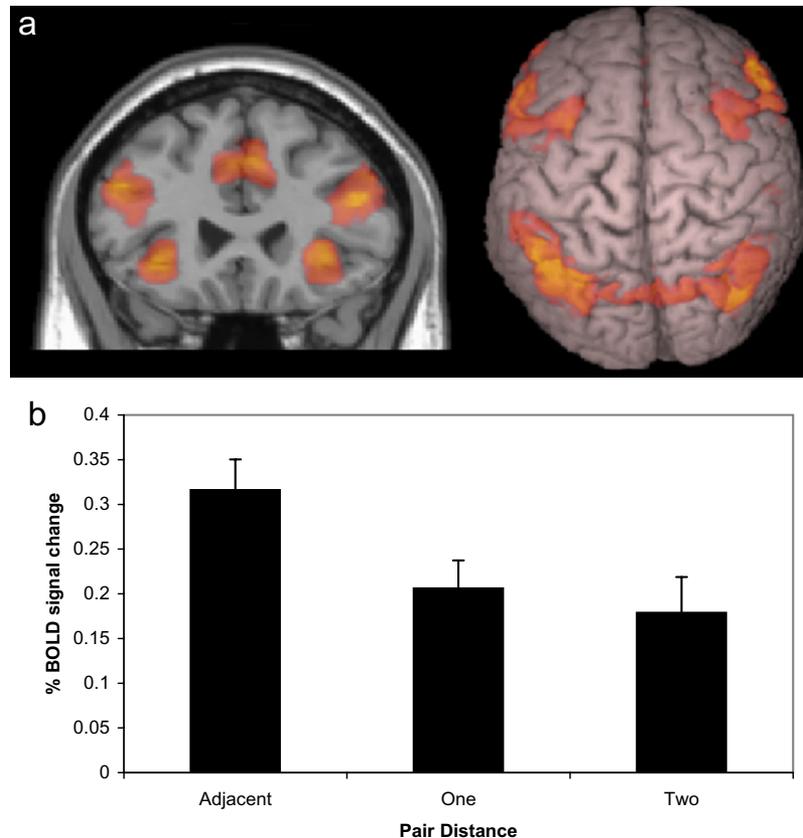
Fig. 3. (a) Coronal section and a 3D rendered image of activity on the surface of the brain showing the regions with greatest activation associated with adjacent relations and the least activation with two-step relations. Bilateral parietal cortex (left peak $z=6.54$, -38, -56, 42; right peak $z=5.17$, $x=35$, $y=-55$, $z=41$), bilateral middle frontal gyrus (DLPFC; left peak $z=5.01$, -46, 22, 28; right peak $z=5.25$, 46, 26, 24), paracingulate gyrus (BA6; peak $z=5.31$, -6, 14, 44), and bilateral inferior frontal cortex (left peak $z=4.99$, $x=-33$, $y=25$, $z=-3$; right peak $z=5.04$, 38, 18, -8). (b) Percentage BOLD signal change extracted from these regions for adjacent, one- and two-step tasks. Error bars represent standard errors. For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.

was found ($F(2,44)=4.17$, $P<.001$), whereby RT decreased linearly as elements were further apart on the array. Pairwise comparisons confirmed a significant difference ($P<.001$) between each level. A significant main effect of relation type was found ($F(1,22)=4.42$, $P=.047$), whereby RTs were significantly slower for relations that were different to the training received: All-More group; more than relations (same as trained)$=1117.26$ (141.51) ms, less than relations (different to trained)$=1180.58$ (140.46) ms; All-Less group; more than relations (different to trained)$=1219.38$ (170.58) ms, less than relations (same as trained)$=1208.26$ (181.64) ms.

Post-imaging questionnaire. All participants correctly reported the contextual cues as meaning "more than" and "less than," respectively. All but two participants (one from each group) correctly learned the hierarchy/array such that they had organized the arbitrary symbols into the correct order. The two participants who had not learned the correct array incorrectly ordered two items.

fMRI data

Neural correlates of the symbolic distance effect. As a significant SDE was observed in the behavioural data,

with decreasing RTs and increasing accuracy from adjacent relations (e.g., A<B, D>C) to two inferential steps (e.g., A<D, B>E), a similar analysis was performed on the associated imaging data. A whole brain corrected analysis with a cluster-based threshold of $z>2.3$, $P<.05$ showed that such a linear trend was found in a number of regions (Fig. 3a), which showed greatest activation for the adjacent relations, followed by one-step relations and with the least activation for the two-step relations. Fig. 3b shows the percentage BOLD signal change extracted from these regions. Fig. 3a shows the brain regions with this pattern of activation. The peaks of the clusters were as follows: bilateral parietal cortex (left peak $z=6.54$, $x=-38$, $y=-56$, $z=42$; right peak $z=5.17$, $x=35$, $y=-55$, $z=41$), bilateral inferior frontal cortex (left peak $z=4.99$, $x=-33$, $y=25$, $z=-3$; right peak $z=5.04$, $x=38$, $y=18$, $z=-8$), dorsolateral prefrontal cortex (left peak $z=5.01$, $x=-46$, $y=22$, $z=28$; right peak $z=5.25$, $x=46$, $y=26$, $z=24$), and paracingulate gyrus (BA6; peak $z=5.31$, $x=-6$, $y=14$, $z=44$).

Activation differences between relational reasoning with and without endpoints. Previous work has suggested that there may be a difference in inferential test performance based on whether an end point from the array

(A or E in the case of a 5-term task) is present. Combinatorial one-step queries provide a case where this can be investigated as A<C, C>A, C<E and E>C include end points whereas B<D and D>B relations do not (Table 1). In line with our hypothesis, significantly greater activation was found in association with the AC/CE trials compared to BD trials in a region encompassing lateral occipital cortex, supramarginal/angular gyrus ($z=3.72$, $x=34$, $y=-74$, $z=32$), as well as another cluster in the middle frontal gyrus (BA6; $z=3.61$, $x=36$, $y=10$, $z=42$).

In an attempt to further understand these activation differences, paired t-tests were conducted on the percentage correct and RT data between performance on AC/CE and BD trials. While performance accuracy did not differ, the difference in reaction time between AC/CE and BD trials marginally missed significance ($P=0.08$). The mean reaction time to AC/CE trials was marginally faster at 1161.26 (145.65) ms compared to 1223.01 (214.35) ms for BD trials.

Comparisons based on training history. Basic contrasts between trained, mutually entailed, one-step and two-step relations compared to baseline (fixation) were conducted to ascertain if there were differences in activation on the basis of type of training using a whole-brain corrected cluster-based threshold ($z>2.3$, $P<.05$); however, no significant differences were found between the two training groups on any of the relation types. The overall pattern of activation for each of the different relations compared to baseline was similar, with a distributed network of regions including prefrontal and parietal cortex, as well as occipital cortex and cerebellum (details of this analysis are not reported due to the large extent of activation). Notably, bilateral hippocampal activation was seen when comparing each relation type to baseline (see below for further ROI analysis on this activation).

In a novel contrast, the activation associated with trained and mutually entailed relations was compared within each group. No significant differences were seen in the All-More group. In the All-Less group, however, greater activation was found in the right parietal (angular gyrus/ supramarginal gyrus, peak $z=3.27$, $x=46$, $y=-46$, $z=42$) and prefrontal cortices (middle frontal gyrus, BA6, peak $z=3.64$, $x=38$, $y=8$, $z=50$) during the mutually entailed relations compared to the trained relations. These regions overlap with those found in the endpoint analysis above. However, when the two groups were directly compared with respect to mutually entailed vs. trained relations, there were no significant differences in activation.

Comparison of same/different relational tasks. During combinatorial one-step relations (A<C, B<D, C<E, C>A, D>B, E>C), an interesting pattern was observed in the comparison between relations that were same or different to training. The All-More group showed greater activation in a region of the parietal cortex (Fig. 4a) to one-step less
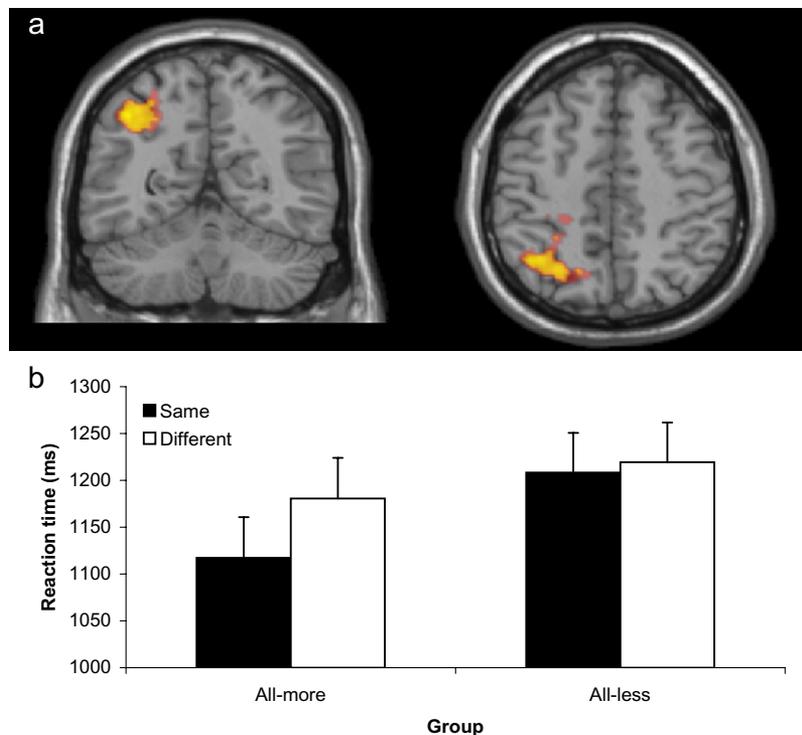

Fig. 4. (a) Coronal and sagittal sections showing the region in the left hemisphere with greater activation associated with relations different to training in both groups. This area of activation encompasses the superior parietal lobule, supramarginal and angular gyrus and lateral occipital cortex (Harvard–Oxford Cortical Structural Atlas). The peak of activation was found at -40, -58, 44, $z=3.31$. (b) Percentage BOLD signal change extracted from this ROI, showing greater activation in this region for relations inconsistent with training in both groups. Error bars represent standard errors. For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.

than relations (i.e., those different to training: A<C, B<D, C<E), whereas the All-Less group showed greater activation in the same region for one-step more than relations, again those different to their training (i.e., C>A, D>B, E>C). Fig. 4b shows the percentage BOLD signal change in this parietal region for both groups.

Region of interest analysis on hippocampus. No significant difference in left or right hippocampal activation was seen between All-More and All-Less training groups. In addition, no significant differences in left or right hippocampal activation was found between relational types (trained, mutually entailed, combinatorial one-step or two-step), or when split into less than or more than relations. Correlations were conducted between RT and accuracy measures for each trial type and associated left and right hippocampal activation (total correlational tests conducted=16). A significant correlation was found between activation in the left hippocampus during trained trials and RTs on those trials ($r^2=0.208$; $r=0.456$, $P=0.025$); however, no further relationships were found between performance measures and left or right hippocampal activation during the other trial types (ME, one- and two-step). As a caveat, this result should be considered preliminary given the number of correlations conducted.

DISCUSSION

This novel investigation into the neural correlates of relational reasoning has yielded some noteworthy results. First, a clear correspondence between relational distance and activation was found, with the neural correlates of the SDE shown as *graded* activation across relational distance in regions such as the parietal and prefrontal cortex; indeed, an analogue representation of relational distance. The SDE was demonstrated for a wider range of distances than before (adjacent vs. one-step vs. two-step in the present study, as compared to one-step vs. two-step; see Zalesak and Heckers, 2009). These regions are likely to reflect working memory contributions to inference making, as well as spatial functions associated with manipulating the linear order representation (Cohen Kadosh et al., 2008; Libben and Titone, 2008; Wendelken et al., 2008). Second, greater activation in parietal and prefrontal cortices was also shown during test trials with endpoints of the array compared to those without. These results highlight the importance of these cortical regions in relational reasoning and the extent to which they are recruited by the current task. Third, relational test trials with cues different to training required greater parietal activity in conjunction with longer RTs, possibly due to the spatial operation required to reverse the correct answer from training and/or difficulty with deriving the correct response in the different context. Finally, training history had almost no effect on brain activation patterns. The one exception, that is, greater parietal and prefrontal activation in the less-than group for mutually entailed relations, must await further research to be clarified.

Neural correlates of the distance effect

As discussed previously, there is considerable behavioural evidence for a distance effect in serial order reasoning (e.g., Acuna et al., 2002; Colombo and Frost, 2001; Frank et al., 2005; Leth-Steensen and Marley, 2000; Riley, 1976; Zalesak and Heckers, 2009). We also found evidence for such an effect as predicted, with the longest RTs and poorest accuracy for adjacent relations, followed by one-step tasks and the fastest responses and greatest accuracy for two-step tasks. We predicted that this distance effect would be reflected in the brain activation pattern in regions involved in relational reasoning and in spatial functions. Accordingly, a similar linear trend was found in inferior frontal cortex, dorsolateral prefrontal cortex, and bilateral parietal cortex activity, with the greatest activation in these regions for adjacent relations and the least activation for two-step relations. Ventrolateral PFC has been shown to be involved in the specification and maintenance of retrieved information, whereas the dorsolateral PFC has been associated with verification and evaluation of such information (Simons and Spiers, 2003). The above results are in line with these functions, given greater maintenance of the adjacent relations would be expected from the training to the test phase, and more thorough verification of the relational pairs may be required when the elements are adjacent in the array.

Indeed, this SDE in the parietal cortex could be due to greater difficulty to spatially distinguish elements closer together in the array once integrated (Zalesak and Heckers, 2009). Zalesak and Heckers (2009) also observed activation in frontal, temporal, bilateral parietal and right angular gyrus to be associated with the SDE, regions similar to those found in the present study. Importantly, the 6-term array used by Zalesak and Heckers (2009) permitted a comparison of novel inference trials without endpoints directly as a function of step distance, whereas the present 5-term task necessitated a comparison of step size on inferences made to trials with and without endpoints (see also, Heckers et al., 2004; Moses et al., 2006). Future neuroimaging research into transitive inference with the present paradigm should consider adopting a 6- or 7-term array.

The pattern of activation in frontal and parietal cortices showing this distance effect is in fact similar to that found previously in association with many different cognitive demands (Duncan and Owen, 2000). Performance of a rapid visual information-processing task, for example, evoked activation in a similar network of frontal and parietal regions (Lawrence et al., 2003). The authors suggested this activation may be due to the sustained attention and working memory demands of such a task involving "maintenance, updating and verification of serial stimuli" (p. 1034). We duly note that this argument could also be applied to the current dataset; working memory has been shown to be involved in relational reasoning (Libben and Titone, 2008). The graded activation found in association with the SDE in the present study, however, is unlikely to be explained exclusively in terms of reduced working memory

demand; the composition of the relations (two items and a cue), and hence the memory load, is the same for all relations. Moreover, these results are unlikely to solely be a function of general task difficulty but related to specific, increased discriminability required when elements are closer in the array. As several authors have argued, larger step-size (e.g., B–D) may be more discriminable than smaller ones (e.g., C–D) on a dimensional representation, such as an array (Holyoak and Patterson, 1981; Huttenlocher, 1968; Leth-Steensen and Marley, 2000).

Relational reasoning with and without endpoints

Both relational frame and associative learning theory predict that RTs will be faster for trials with endpoints. While one-step trials with endpoints were faster than trials without endpoints, the data marginally missed significance to fully support this prediction. The main finding here, however, is that greater brain activity was found in frontal, parietal and occipital regions associated with the one-step combinatorial relations with endpoints (A<C, C>A, C>E, E<C) compared to trials without endpoints (B<D and D>B). This is of particular note, given that the focus of many relational reasoning studies (e.g., Greene et al., 2006) is rather on the BD trials themselves, as they provide a clear instance of inference. This activation, therefore, may represent processing associated with utilizing a serially ordered representation, rather than inference per se. The endpoints of an array may provide the anchors on which an array is formed and subsequent judgments are made (Trabasso and Riley, 1975; Leth-Steensen and Marley, 2000); therefore, it is plausible that processing relations with endpoints, perhaps through their greater salience (see Leth-Steensen and Marley, 2000), is associated with greater activation of relevant areas. Indeed, the middle frontal gyrus including BA6 has previously been linked to processing of sequences (Hanakawa et al., 2002; Schubotz et al., 2004), and the posterior parietal cortex has been long since associated with numerical operations, using visuospatial processing of elements (de Hevia et al., 2008).

This argument is also consistent with the activation differences in right frontal and parietal cortex from the mutually entailed-trained relations contrast, which partially overlap with the above endpoint contrast. These differences suggest that compared to trained trials, greater processing of the array sequence is needed to solve mutually entailed trials (Dickins et al., 2001; Schlund et al., 2007, 2008).

The analysis comparing trials with and without endpoints provide an informative example in which greater activation in the parietal cortex is associated with marginally faster reaction times. This is in contrast to the argument that activity in this region is a reflection of response selection processes related to longer reaction times, rather than relating to processing of the ordered representation (Göbel et al., 2004).

Training history

If pre-experimental learning histories with more-than relations continues to exert an influence over adult reasoning

abilities, then RFT predicts improved inferential performance on more-than trials relative to less-than trials. However, an alternative view is taken from associative learning theory, which does not predict any performance differences between the two training schedules (Wynne, 1998). According to this account, the trained associative strengths reinforced to elements in the array are unaffected by training schedule, because during nonarbitrary training more-than and less-than cues are equally associated with reinforcement. As a result, the cues share equal associative strength during the subsequent test phases. Associatively, then, training B>A or A<B should be indistinguishable, leading to comparable performance at test for both groups. Therefore, the broad finding that there was no effect of type of training history on relational reasoning is consistent with associative learning accounts.

Relations same as training or different to training

Greater activation in the left parietal cortex and longer RTs were found when inferring test trials with contextual cues different to training. These effects could be a result of greater derivation difficulty for the trials with cues different to training; where "difficulty" could be an instance of greater processing load when a reversal/reordering of the pre-learned relation between any two elements is required, rather than greater relational difficulty per se, as the number of arguments in the relation stays the same but the contextual cue changes. The finding that an additional spatial operation, in this case of reordering the elements in the array according to contextual cue, results in greater parietal activation supports the finding of Kroger et al. (2002), in which linear increases in bilateral posterior parietal activation was found, notably in the same region as this study, as relational complexity involving greater spatial manipulation demands increased. Moreover, activity in this region has also been associated with the amount of interference experienced with contextually incongruent semantic stimuli (Hoenig and Scheef, 2009), and with tasks involving learned context-dependent rules (Rawley and Constantinidis, 2009). However, in accordance with Christoff et al. (2001), given the longer response latencies associated with test trials with cues different to training, the greater parietal activity found here could be a reflection of longer duration of processing incorporating visual inspection of the relation or attentional engagement (Göbel et al., 2004), as well as the spatial operations required to process the correct response.

Role of the hippocampus

As noted above, much of the TI neuroimaging literature to date has focused on the role of the hippocampus. There is still some debate regarding the exact function of this region in relational reasoning, as for example, Greene et al. (2006) suggest an associative role in the formation of integrated representations through in-context discrimination in study phases and inference at test. Greene et al. (2006) found that the extent of hippocampal activation differentially predicted test performance with greater activation during inference; a finding supported by Zalesak

and Heckers (2009), who also observed greater right hippocampal activation for the one-step pairs B>D and C>E compared to the two-step pair B>E, suggesting a general scaling of the extent to which the hippocampus is recruited during relational reasoning. Van Opstal et al. (2008), on the other hand, found no relationship between activation and performance, leading to the suggestion of a more general role for the hippocampus as a pattern separator.

The data from the present study lend some support to the idea of a more general role for the hippocampus during relational reasoning rather than inference per se, as no differences in activation were found between relational types, and the only relationship with performance was limited to presentation of the previously trained trials during the test phase. These findings are more in accordance with the idea that the main role of the hippocampus during reasoning test phases, such as in that under investigation in this study, is in maintaining the integrated relational structure (Preston et al., 2004; Schlund et al., 2008).

CONCLUSION

In summary, we have demonstrated a crucial role of the parietal cortex, along with prefrontal areas, during different aspects of relational reasoning: in representing the linear changes in performance akin to the SDE, when end-point elements are present in the relational task, and when the relational context is different to previous training. The common element of these aspects of relational reasoning is that accessing and utilizing a spatially ordered representation of the newly learned array is required.

REFERENCES

Acuna BD, Eliassen JC, Donoghue JP, Sanes JN (2002) Frontal and parietal lobe activation during transitive inference in humans. Cereb Cortex 12:1312–1321.

Beckmann C, Jenkinson M, Smith S (2003) General multi-level linear modelling for group analysis in fMRI. Neuroimage 20:1052–1063.

Berens NM, Hayes SC (2007) Arbitrarily applicable comparative relations: experimental evidence for a relational operant. J Appl Behav Anal 40:45–71.

Christoff K, Prabhakaran V, Dorfman J, Zhao Z, Kroger JK, Holyoak KJ, et al. (2001) Rostrolateral prefrontal cortex involvement in relational integration during reasoning. Neuroimage 14:1136–1149.

Cohen Kadosh R, Lammertyn J, Izard V (2008) Are numbers special? An overview of chronometric, neuroimaging, developmental and comparative studies of magnitude representation. Prog Neurobiol 84:132–147.

Colombo M, Frost N (2001) Representation of serial order in humans: a comparison to the findings with monkeys (*Cebus apella*). Psychon Bull Rev 8:262–269.

Davis H (1992) Transitive inference in rats (*Rattus norvegicus*). J Comp Psychol 106:342–349.

Dehaene S, Piazza M, Pinel P, Cohen L (2003) Three parietal circuits for number processing. Cogn Neuropsychol 20:487–506.

de Hevia MD, Vallar G, Girelli L (2008) Visualizing numbers in the mind's eye: the role of visuo-spatial processes in numerical abilities. Neurosci Biobehav Rev 32:1361–1372.

Dickins DW, Singh KD, Roberts N, Burns P, Downes JJ, Jimmieson P, Bentall RP (2001) An fMRI study of stimulus equivalence. Neuroreport 12:405–411.

Duncan J, Owen AM (2000) Common regions of the human frontal lobe recruited by diverse cognitive demands. Trends Neurosci 23:475–483.

Frank MJ, Rudy JW, Levy WB, O'Reilly RC (2005) When logic fails: implicit transitive inference in humans. Mem Cognit 33:742–750.

Göbel SM, Johansen-Berg H, Behrens T, Rushworth MFS (2004) Response-selection-related parietal activation during number comparison. J Cogn Neurosci 16:1536–1551.

Goel V, Dolan RJ (2001) Functional neuroanatomy of three-term relational reasoning. Neuropsychologia 39:901–909.

Greene AJ, Gross WL, Elsinger CL, Rao SM (2006) An fMRI analysis of the human hippocampus: inference, context, and task awareness. J Cogn Neurosci 18:1156–1173.

Hanakawa T, Honda M, Sawamoto N, Okada T, Yonekura Y, Fukuyama H, et al. (2002) The role of rostral Brodmann area 6 in mental-operation tasks: an integrative neuroimaging approach. Cereb Cortex 12:1157–1170.

Hayes SC, Barnes-Holmes D, Roche B (2001) Relational frame theory: a post-Skinnerian account of human language and cognition. New York, NY: Plenum Press.

Heckers S, Zalesak M, Weiss AP, Ditman T, Titone D (2004) Hippocampal activation during transitive inference in humans. Hippocampus 14:153–162.

Hoenig K, Scheef L (2009) Neural correlates of semantic ambiguity processing during context verification. Neuroimage 45:1009–1019.

Holyoak KJ, Patterson KK (1981) A positional discriminability model of linear order judgments. J Exp Psychol Hum Percept Perform 7:1283–1302.

Huttenlocher J (1968) Constructing spatial images: a strategy in reasoning. Psychol Rev 75:550–560.

Jenkinson M, Smith S (2001) A global optimisation method for robust affine registration of brain images. Med Image Anal 5:143–156.

Jenkinson M, Bannister P, Brady J, Smith S (2002) Improved optimisation for the robust and accurate linear registration and motion correction of brain images. Neuroimage 17:825–841.

Kroger JK, Sabb FW, Fales CL, Bookheimer SY, Cohen MS, Holyoak KJ (2002) Recruitment of anterior dorsolateral prefrontal cortex in human reasoning: a parametric study of relational complexity. Cereb Cortex 12:477–485.

Lawrence NS, Ross TJ, Hoffmann R, Garavan H, Stein EA (2003) Multiple neuronal networks mediate sustained attention. J Cogn Neurosci 15:1028–1038.

Leth-Steensen C, Marley AAJ (2000) A model of response time effects in symbolic comparison. Psychol Rev 107:62–100.

Libben M, Titone D (2008) The role of awareness and working memory in human transitive inference. Behav Processes 77:43–54.

Moses SN, Villate C, Ryan JD (2006) An investigation of learning strategy supporting transitive inference performance in humans compared to other species. Neuropsychologia 44:1370–1387.

Nagode JC, Pardo JV (2002) Human hippocampal activation during transitive inference. Neuroreport 13:939–944.

Pinel P, Dehaene S, Riviere D, leBihan D (2001) Modulation of parietal activation by semantic distance in a number comparison task. Neuroimage 14:1013–1026.

Preston AR, Shrager Y, Dudukovic NM, Gabrieli JDE (2004) Hippocampal contribution to the novel use of relational information in declarative memory. Hippocampus 14:148–152.

Rawley JB, Constantinidis C (2009) Neural correlates of learning and working memory in the primate posterior parietal cortex. Neurobiol Learn Mem 91:129–138.

Reilly T, Whelan R, Barnes-Holmes D (2005) The effect of training structure on the latency of responses to a five-term linear chain. Psychol Rec 55:233–249.

Riley C (1976) The representation of comparative relations and the transitive inference task. J Exp Child Psychol 22:1–22.

Schlund MW, Hoehn-Saric R, Cataldo MF (2007) New knowledge derived from learned knowledge: functional-anatomic correlates of stimulus equivalence. J Exp Anal Behav 87:287–307.

Schlund MW, Cataldo MF, Hoehn-Saric R (2008) Neural correlates of derived relational responding on tests of stimulus equivalence. Behav Brain Funct 4:6.

Schubotz RI, Sakreida K, Tittgemeyer M, von Cramon DY (2004) Motor areas beyond motor performance: deficits in serial prediction following ventrolateral premotor lesions. Neuropsychology 18: 638–645.

Simons JS, Spiers HJ (2003) Prefrontal and medial temporal lobe interactions in long-term memory. Nat Rev Neurosci 4:637–648.

Smith S (2002) Fast robust automated brain extraction. Hum Brain Mapp 17:143–155.

Trabasso T, Riley C (1975) On the construction and use of representations involving linear order. In: Information processing and cognition: The Loyola symposium (Solso R, ed). Hillsdale, NJ: Erlbaum.

Van Opstal F, Verguts T, Orban GA, Fias W (2008) A hippocampal-parietal network for learning an ordered sequence. Neuroimage 40:333–341.

Van Opstal F, Fias W, Peigneux P, Verguts T (2009) The neural representation of extensively trained ordered sequences. Neuro-image 47:367–275.

Vasconcelos M (2008) Transitive inference in non-human animals: an empirical and theoretical analysis. Behav Processes 78:313–334.

Vitale A, Barnes-Holmes Y, Barnes-Holmes D, Campbell C (2008) Facilitating responding in accordance with the relational frame of comparison: systematic empirical analyses. Psychol Rec 58:365–390.

Wendelken C, Nakhabenko D, Donohue SE, Carter CS, Bunge SA (2008) "Brain is to thought as stomach is to??": investigating the role of rostrolateral prefrontal cortex in relational reasoning. J Cogn Neurosci 20:682–693.

Whelan R, Barnes-Holmes D, Dymond S (2006) The transformation of consequential functions in accordance with the relational frames of more-than and less-than. J Exp Anal Behav 86:317–335.

Woolrich M, Ripley B, Brady J, Smith S (2001) Temporal autocorrelation in univariate linear modelling of fMRI data. Neuroimage 14:1370–1386.

Woolrich M, Behrens T, Beckmann C, Jenkinson M, Smith S (2004) Multi-level linear modelling for fMRI group analysis using Bayesian inference. Neuroimage 21:1732–1747.

Worsley K, Evans A, Marrett S, Neelin P (1992) A three-dimensional statistical analysis for CBF activation studies in human brain. J Cereb Blood Flow Metab 12:900–918.

Wynne CDL (1998) A minimal model of transitive inference. In: Models of action (Staddon JER, ed), pp 269–307. Mahwah, NJ: Lawrence Erlbaum Associates.

Zalesak M, Heckers S (2009) The role of the hippocampus in transitive inference. Psychiatry Res 172:24–30.