Adv in Health Sci Educ (2016) 21:935–952
DOI 10.1007/s10459-015-9649-2

Working memory, reasoning, and expertise in medicine—insights into their relationship using functional neuroimaging

Pam Hruska[1] · Olav Krigolson[2] · Sylvain Coderre[3] · Kevin McLaughlin[3] · Filomeno Cortese[4] · Christopher Doig[5] · Tanya Beran[1] · Bruce Wright[6] · Kent G. Hecker[1,7]

Received: 22 October 2015 / Accepted: 25 October 2015 / Published online: 4 November 2015

Abstract Clinical reasoning is dependent upon working memory (WM). More precisely, during the clinical reasoning process stored information within long-term memory is brought into WM to facilitate the internal deliberation that affords a clinician the ability to reason through a case. In the present study, we examined the relationship between clinical reasoning and WM while participants read clinical cases with functional magnetic resonance imaging (fMRI). More specifically, we examined the impact of clinical case difficulty (easy, hard) and clinician level of expertise (2nd year medical students, senior gastroenterologists) on neural activity within regions of cortex associated with WM (i.e., the prefrontal cortex) during the reasoning process. fMRI was used to scan ten second-year medical students and ten practicing gastroenterologists while they reasoned through sixteen clinical cases [eight straight forward (easy) and eight complex (hard)] during a single 1-h scanning session. Within-group analyses contrasted the easy and hard cases which were then subsequently utilized for a between-group analysis to examine effects of expertise (novice > expert, expert > novice). Reading clinical cases evoked multiple neural

✉ Kent G. Hecker
kghecker@ucalgary.ca

[1] Department of Community Health Sciences, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

[2] Neuroscience Program, Centre for Biomedical Research, and School of Exercise Science, Physical, and Health Education, University of Victoria, Victoria, BC, Canada

[3] Undergraduate Medical Education, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

[4] Seaman Family MR Research Centre, Hotchkiss Brain Institute, University of Calgary, Calgary, AB, Canada

[5] Department of Critical Care Medicine, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

[6] Division of Medical Sciences, University of Victoria, Victoria, BC, Canada

[7] Department of Veterinary Clinical and Diagnostic Sciences, Faculty of Veterinary Medicine, University of Calgary, Calgary, AB, Canada

activations in occipital, prefrontal, parietal, and temporal cortical regions in both groups. Importantly, increased activation in the prefrontal cortex in novices for both easy and hard clinical cases suggests novices utilize WM more so than experts during clinical reasoning. We found that clinician level of expertise elicited differential activation of regions of the human prefrontal cortex associated with WM during clinical reasoning. This suggests there is an important relationship between clinical reasoning and human WM. As such, we suggest future models of clinical reasoning take into account that the use of WM is not consistent throughout all clinical reasoning tasks, and that memory structure may be utilized differently based on level of expertise.

Keywords Clinical reasoning · Working memory · Functional magnetic resonance imaging · Novice expert studies

Introduction

Effective and safe patient care depends on sound clinical reasoning and diagnosis (Croskerry 2009; Durning et al. 2010), therefore proficiency in both reasoning and decision making are important abilities for a physician (Elstein et al. 1978). Reasoning and decision-making stages are distinct from one another in that *clinical reasoning* is the activity prior to or during attempts to solve a medical problem, whereby a clinician weighs and sorts through assessment details obtained from medical history, physical assessment and test results. Subsequent to this, *clinical decision-making* is the stage during which the clinician chooses between competing options or courses of action to assign a final diagnosis and determine the plan of care (Simmons 2010). The processes of reasoning and decision-making are complimentary to one another and require integration of information from basic science, medical knowledge and clinical experience (Norman 2005). In this study, we focus on clinical reasoning and how this cognitive process is supported by long term and working memory.

Memory and reasoning

Though memory and reasoning have been considered separate topics in literature and contemplated from insulated experimental paradigms and theoretical models, recent work exploring many different types of tasks demonstrate a relationship between these cognitive activities (Heit et al. 2012; Süß et al. 2002). Relationships between memory and reasoning are noted in dual-task demands, tasks requiring manipulation of memory content, and tasks where memory content must be coordinated for integration into a new domain (Rottschy et al. 2012; Süß et al. 2002). Parallels to each of these mentioned tasks are found in clinical reasoning and as such, it can be implied memory processes and clinical reasoning are interdependent (Ericsson and Kintsch 1995; Heit et al. 2012; Süß et al. 2002). Indeed, a physician's ability to access stored information from long-term memory (LTM) is of critical importance so that appropriate knowledge schemas (McLaughlin et al. 2006) and/or illness scripts (Charlin et al. 2000; Coderre et al. 2009) can be mobilized for evaluation of the presenting medical problem (Pelaccia et al. 2011). The need to access LTM during clinical reasoning implicates working memory (WM) as being crucial to the reasoning process. Specifically, in order for a physician to use LTM during reasoning the relevant

LTMs need to be brought forward into WM (Baddeley 1992, 1996; Baddeley et al. 2014; Rottschy et al. 2012) where they can be accessed and manipulated in order to clinically reason and diagnose clinical cases. In other words, the clinical reasoning process is dependent upon and sub-served by WM.

Studies using functional magnetic resonance imaging (fMRI) have demonstrated that the neural locus of WM is prefrontal regions of the human cortex (Ranganath et al. 2003; Ruff et al. 2003; Simons and Spiers 2003). The prefrontal cortex (PFC) is thought to be where the central executive resides, which is the high level system that controls WM and, thus, the retrieval and access of LTM (Carpenter et al. 2000; Collette and Van der Linden 2002). More specific localizations that may also be relevant include the dorsolateral prefrontal cortex (DLPFC) for goal maintenance, executive control (Sanfey and Chang 2008) and selective attention (Rosler et al. 2009), as well as the ventrolateral prefrontal cortex (VLPFC) for attention control and simple memory recall—purportedly in the format of the phonological loop and the visuospatial sketchpad (Baddeley 2000; Rosler et al. 2009).

While previous neuroimaging research serves as a basis for understanding neural areas implicated in WM in reasoning tasks from other realms, there is need for further information about the neural underpinnings in clinical reasoning and the role of WM in this process.

Functional magnetic resonance imaging (fMRI) and medical education

fMRI studies conducted within medical education have mostly investigated the decision-making phase (Downar et al. 2011; Durning et al. 2014; Melo et al. 2011), or are visuo-spatial in nature (Bahrami et al. 2014; Haller and Radue 2005). There are, however, two relatable pieces of work to date discussing clinical reasoning that is non-visual in nature. The first is an fMRI study that attempted to identify functional differences in analytic versus non-analytic reasoning. fMRI images obtained in this work were done during three contrasting phases; reading, which was treated as baseline neural activation, answering, and reflecting. Immediately after scanning, participants engaged in a think aloud protocol where they were asked to describe the process they used to answer each question. Results were interpreted to suggest that greater activation in the PFC is associated with analytical reasoning, based on the assumption that analytical reasoning was represented by incorrect answers, guessing, and deep thought (Durning et al. 2012). In a follow up study, the diagnostic thinking inventory (DTI) tool was used in conjunction with an fMRI study to determine if it could serve as a proxy measure for neural areas associated with analytic and non-analytic thinking. Findings were interpreted to suggest DTI memory structure scores were related to proposed neural structures for non-analytical reasoning (left inferior parietal lobule, left ventromedial prefrontal cortex, and left DLPFC), whereas scores related to flexibility in thinking were related to areas associated with analytical reasoning (bilateral ventromedial prefrontal cortex and the right parahippocampal gyrus), (Durning et al. 2015b).

The second fMRI study explored non-analytic (non-declarative) reasoning by ten internal medicine interns (novices) and seventeen board-certified staff internists (experts) (Durning et al. 2015a). In phase one of the experimental task, the multiple-choice question (MCQ) was presented during a reading phase. In phase two, answer options were presented for participants to select from, followed by phase three in which participants were instructed to silently reflect on how they arrived at their diagnosis. Results from this work suggested a common neural network between novices and experts during non-analytical

reasoning, and that experts had increased efficiency (decreased activation) in the PFC. This prefrontal efficiency was interpreted as increased use of non-analytical reasoning processes in experts (Evans and Stanovich 2013).

Purpose

Here we present the first fMRI study focusing on differences between novice and experts during clinical reasoning to highlight the role of WM in this cognitive process. The specific aim of the present study was to explore neural areas of activation in novice (2nd year medial students) and expert (senior gastroenterologists) clinicians during clinical reasoning tasks and to see whether areas of activation differed when cases were straightforward (easy) or more complex (hard). Our hypotheses were twofold: (1) Common neural areas associated with WM would be activated in novices and experts, with both easy and difficult clinical scenarios because of a general network demonstrating interdependence of WM and reasoning and (2) there would be greater activation of the PFC in novice participants while reading harder cases because of increased demands on WM.

Methods

Participants

Twenty healthy volunteers with normal or corrected-to-normal vision completed the present experiment in full. Ten second-year medical students [eight male, mean (range) age 26.5 (22–38) years, SD = 5.3] were the novice participant group and ten currently practicing gastroenterologists [five male, mean (range) age 39.5 (32–50) years, SD = 4.5] were the expert clinicians. Novice participants were all from the Cumming School of Medicine at the University of Calgary and had completed the gastrointestinal course 1 year prior to this research. Expert participants were all currently practicing gastroenterologists with formal academic teaching responsibilities at the Cumming School of Medicine at the University of Calgary. We restricted participation to right-handed individuals due to predominant language processing being lateralized to the left cerebral cortex (Oldfield 1971; Savoy 2006). Exclusion criteria included inability to complete an fMRI due to scanning safety risks (metal, anxiety, claustrophobia, pregnancy), medical history of traumatic brain injury, history of non-medically induced loss of consciousness for >10 min, psychiatric illness requiring medical treatment, past or present use of psychotropic drugs, history of seizures, any other diagnosis of neurovascular or neurophysiologic abnormality, or use of calcium channel blockers. Participants were free to withdraw from the study at any time. This study was conducted in accordance with the ethical standards outlined by the Calgary Health Ethics Research Board (CHREB), and the Seaman Family MR Research Center at Foothills Medical Centre.

Stimuli and procedures

In the block design of this research, each clinical case presented was carried out as a 'run'. A run is defined as an uninterrupted presentation of an experimental task used in fMRI (Huettel et al. 2009). The experiment was performed using Presentation® (Version 16, www.neurobs.com). Each case presentation began with a centrally presented fixation cross

for 10 s after which a randomly selected case was presented for participants to read for 80 s. There were other aspects to this experiment, such as a clinical decision making component, and these components are to be presented in subsequent papers. In this study we focus solely on the clinical reasoning process, however, it is important to point out that participants were made aware that in phases following reading, they would be required to select the most likely diagnosis as presented in a multiple-choice question (MCQ; clinical diagnosis). Just prior to scanning, participants were made aware of the scanning sequence they would go through in general; details about the type of content being presented on were not exposed to the participants until they were in the scanner. A mock practice case was used to explain the sequence of events in the scanner and to practice use of hand held button pads for answer responses prior to the actual scanning.

While in the scanner, participants read sixteen gastroenterology clinical cases during a single 1-h fMRI scanning session. Each case was approximately 215–219 words of written text in length, and was shown via a mirror on a rear-projected screen situated above and behind the participant's head via the fMRI scanner projection system (Avotec Inc, Florida, USA).

Clinical cases

Clinical cases were optimized for display from previously published written cases for this research (Coderre et al. 2010). Eight of the clinical cases were deliberately made to be "easy" and eight were made to be "hard". For easy questions, the patient's initial contextual (written) data was concordant with the analytical data (lab values) subsequently presented. For hard questions, the patient's contextual data was discordant with the analytical data subsequently presented; see "Appendix 1 and 2" for example clinical cases. Questions were randomized during scans. There were four questions related to each of the following clinical presentations; elevated liver enzymes, diarrhea, dysphagia, and anemia.

Statistical inference

Subject data obtained in this fMRI research are representative of each participant's brain. Data collected are called voxels, analogous to 3D pixels, and are volumetric in nature (Ashby 2011). Each voxel is represented by 3D coordinates (x, y, z), which are used to identify associated structural areas using brain atlas tools; see this tutorial for a further description of neuroanatomical coordinate systems (Tadel 2015: http://neuroimage.usc.edu/brainstorm/CoordinateSystems). fMRI images are made possible by tracking hemodynamic response to neural activity over time (Huettel et al. 2009; Logothetis 2003). When neurons become active in response to a task or demand, hemodynamic changes of increased blood volume, increased blood flow and alterations in oxygenation occur (Attwell and Iadecola 2002; Heeger and Ress 2002). These changes produce the blood oxygen level dependent (BOLD) signal, which can be simplistically described as a ratio of oxygenated to deoxygenated hemoglobin (Ashby 2011).

Deoxyhemoglobin, being paramagnetic, is more attracted to the magnetic field and interrupts magnetic resonance signals more so than oxyhemoglobin (Heeger and Ress 2002). These differences in magnetic properties are what provide a natural contrast for fMRI data analysis. The underlying assumption in fMRI is that increased oxyhemoglobin concentration indicates nearby neural activity (Savoy 2001). The BOLD signal therefore enables researchers to make inferences about how imposed cognitive tasks impact neural activity (Ashby 2011).

During data analysis for this research, a subtractive approach was used to determine if the BOLD signal was more prominent in task conditions as compared to baseline conditions (Amaro and Barker 2006). With this method, images obtained during tasks of interest (such as in reading/reasoning or decision-making/MCQ phases) were compared against images obtained during baseline tasks (staring at the fixation cross). If BOLD responses during task of interest images meet or exceed the statistical threshold set over the BOLD responses found in baseline tasks, neural activations are attributed to the performance of the task of interest (Amaro and Barker 2006). A cluster threshold of $p < 0.05$ and $z > 2.3$ were selected within the FSL software package for statistical threshold. This means only groups of 50 or more contiguous voxels (clusters) meeting the set statistical threshold would be considered areas of activation (Poldrack et al. 2011).

Functional and structural magnetic resonance imaging (fMRI) data acquisition

Data were acquired on a 3-Tesla GE Discovery MR750 diagnostic magnetic resonance whole body scanner (General Electric Healthcare, Waukesha, WI, USA) at the University of Calgary, Seaman Family MR Research Centre at Foothills Medical Centre. Preliminary imaging consisted of acquiring a T1-weighted 2D spin-echo sequence with the same geometric orientation and voxel size as the subsequent functional images. The functional imaging sequences were gradient-recalled echo, echo-planar imaging (GRE–EPI) sequences in the oblique/axial plane and were acquired in an interleaved, bottom-up slice acquisition [repetition time (TR) = 2000 ms, echo time (TE) = 20 ms, flip angle (FA) = 70°, 37 slices at 3 mm thickness, in-plane resolution of 64 × 64 pixels reconstructed in a field of view (FOV) of 24 cm] using a 12-channel receive-only phased array head coil. Each functional run began with 6 TRs during which no data were acquired to allow for steady-state tissue magnetization. A total of 90 echo-planar imaging volumes were collected in each functional run, and a total of 16 functional runs were collected for each participant. A 3D high-resolution (1 × 1 × 1 mm), T1-weighted axial images were also taken of each participant (FOV = 25.6 cm) for registration of the functional data.

fMRI data processing and analysis

Data were preprocessed using FEAT (Version 6.00), which is part of FSL (FMRIB's Software Library, Version 6.0; FMRIB Analysis Group, Oxford University, UK). Data were motion corrected (Jenkinson 2003), registered by FLIRT (Jenkinson and Smith 2001), and spatially smoothed with a Gaussian kernel of 5.0 mm full width at half maximum. The resulting time series was then convolved using a gamma function. For an overview of fMRI analyses, please refer to Smith (Smith 2004) or this online tutorial (see https://sites.google.com/site/mritutorial/functional-mri-tutorials/tutorial-i-overview-of-fmri-analysis).

In line with standard MRI analysis procedures (Huettel et al. 2009), we next employed contrast analyses to allow meaningful interpretation of our experimental results. In essence, contrast analyses isolate differences between functional images (e.g., a paired samples t test). The reasoning for this approach is simple—examination of a functional image associated with the viewing of an experimental stimulus would reveal activations across the entire human brain, and as such, a contrast or difference approach is used to isolate differences between two functional images while similar regions of activation are effectively cancelled out.

With that in mind, for our first-level analysis each of the 16 functional runs (i.e., the reading of a single case) for each participant were analyzed separately. In these analyses the functional images for case reading (the first 20 s of the 80 s reading phase) were contrasted with the functional images for the first fixation phase (10 s; read > fixation) to highlight neural regions where activations were greater for reading than for viewing of the fixation image. The resulting "contrast" statistical maps (i.e., parameter estimate maps and variance maps) were subsequently forwarded to a second-level fixed-effects analysis. In the second-level analysis average contrast images were constructed for the reading of easy (8) and hard (8) clinical cases for each participant. After averaging, additional levels of contrasts were conducted for each participant to assess differences between the reading of easy and hard clinical cases. Specifically, contrasts were conducted to examine activations that were greater for the reading of hard cases relative to the reading of easy cases (hard > easy) and the reverse, contrasts to examine activations that were greater for the reading of easy cases relative to the reading of hard cases (easy > hard). Within the FSL MRI analysis package one has to implement both sets of contrasts—unlike a *t* test that can give both positive and negative *t* statistic's we required that FSL only generate positive activation values. As such, one must run contrasts in both directions to assess all statistical effects. We also note here that all functional runs were given equal weighting in the model.

A final group level mixed-effects analysis using FLAME 1 + 2 [a method used due to the smaller sample size in the present experiment (McCarthy 2015: http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/GLM)] was conducted wherein images from the second-level contrasts for each of the two groups of participants were combined separately to model group level differences related to expertise (e.g., novice versus expert). Thus, at this final stage of analysis we contrasted group level effects—a contrast to assess activations that were greater for novices than for experts (novices > experts) and a contrast to assess activations that were greater for experts than for novices (experts > novices). Again, for users unfamiliar with fMRI analysis FSL carries over lower level contrasts. As such, we were able to examine interactions between expertise (novice, expert) within case difficulty (easy, hard) and examine interactions between expertise (novice, expert) and case difficulty (easy, hard) after the group level analysis.

Statistically significant clusters of activation were initially identified on the entire group statistical map by using a voxel-wise threshold to $z > 2.3$ ($p < 0.05$) and the FSL cluster analysis procedure. However, given our outlined a priori hypotheses, we also conducted region of interest (ROI) analyses (Poldrack 2007) to focus on specific neural regions. In these analyses, ROIs were defined within the prefrontal cortex (DLPFC, VLPFC) and similar statistical criteria were used to evaluate activation: a voxel-wise threshold to $z > 2.3$ ($p < 0.05$) and a criteria of at least 30 contiguous voxels (Worsley et al. 1992).

To summarize the analyses, within the novice and expert groups we first wanted to identify neural areas of activation associated with clinical reasoning. We then were interested in whether there were differences in neural activation between novices and experts on easy questions in isolation and hard questions in isolation. Finally, we used an ROI analysis to identify if there were any interactions between case difficulty and level of expertise.

We used the Montreal Neurological Institute (MNI) coordinates of the voxel within the cluster with the maximum *z* statistic to determine the most probable anatomical label for the cluster from the Harvard–Oxford Cortical Structural Atlas packaged in FSL. We have also included the reading versus fixation contrast for both easy and hard clinical cases for both groups to show whole brain activation.

Results

Our initial analyses were focused on separate examinations of the group activation maps for the novice and expert clinicians while they read the easy and hard clinical cases. By doing this, we hoped to identify regions of interest (i.e., differential activation between easy and hard cases and/or novices and experts) for subsequent analyses. As expected, reading clinical cases (both easy and hard) evoked significant changes in hemodynamic activity in multiple brain regions for both novice and expert clinicians (see Fig. 1; Tables 1, 2 for neural areas significantly activated). Specifically, the top row in Fig. 1 shows the areas of activation in the right and left hemisphere of the novice brain while reading easy and hard clinical cases and these areas of activation correspond to the data provided in Table 1. In Fig. 1 areas of activations associated with reading easy cases are shown in blue, areas of activations associated with reading hard cases are shown in red and the overlap is shown in purple. So, the largest cluster size in Table 1 for reading easy cases is in the Occipital Fusiform Gyrus of the right hemisphere and the largest cluster size for reading hard cases is in the Superior Division of the Lateral Occipital Cortex of the right

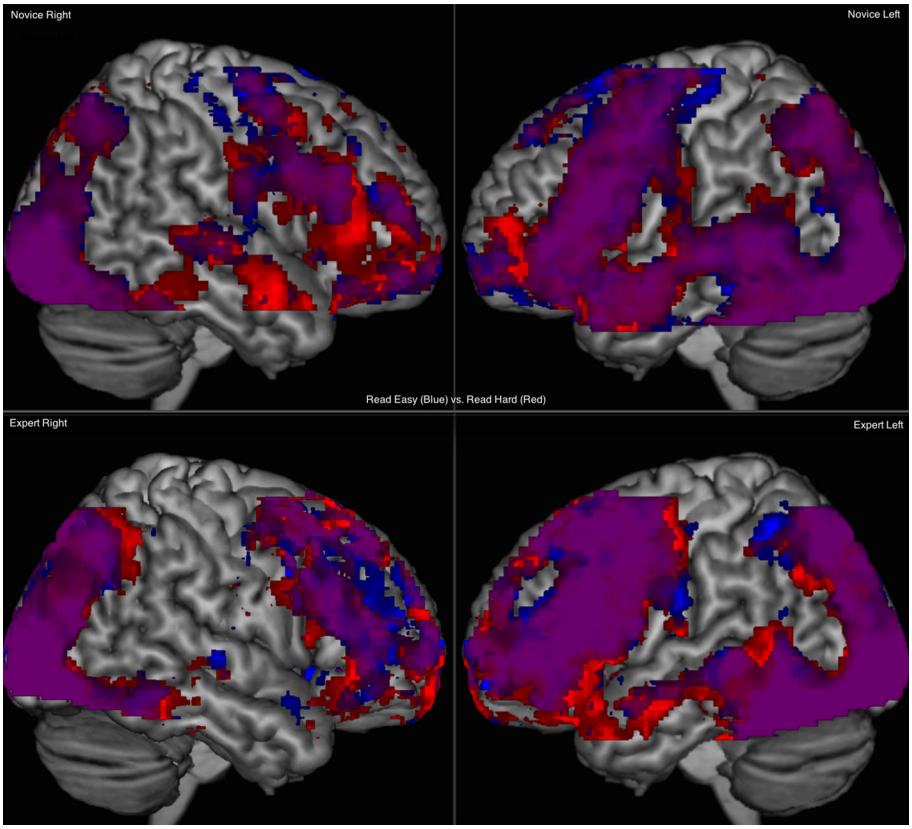

Fig. 1 Combined neural areas of activation in clinical reasoning for novices (*top row*) and experts (*bottom row*). *Top row* (*Right hemisphere* & *Left hemisphere*): novice brain. *Bottom row* (*Right hemisphere* & *Left hemisphere*): expert brain. Reading easy cases (*blue*) versus reading hard cases (*red*). Common areas of activation (*purple*). (Color figure online)

Table 1 Common areas of neural activation (clusters) for novices reading easy and hard clinical cases

Cerebral hemisphere	Neural area of activation: anatomical labels	Cluster size (number of activated voxels)	Max Z (maximum statistic value)	MNI coordinates (locations of activated voxels using 3D coordinates)		
				X	Y	Z
Reading easy cases (corresponds to blue colour in the top row of Fig. 1)						
Right[a]	Occipital fusiform gyrus	22,695	6.04	36	23	25
Right[a]	Middle frontal gyrus	533	3.97	20	80	51
Right	Precentral gyrus	256	4.92	27	61	68
Right	Juxtapositional lobule cortex	233	3.58	42	66	59
Left	Frontal pole	195	3.86	57	94	32
Right	Frontal pole	157	4.2	27	88	46
Right	Heschel's gyrus (includes H1 + H2)	155	3.58	20	55	39
Right	Postcentral gyrus	84	3.49	11	58	52
Left	Frontal orbital cortex	80	4.03	62	79	24
Right	Paracingulate gyrus	71	3.42	39	72	57
Left	Superior frontal gyrus	53	4.16	51	71	70
Right	Frontal pole	53	3.59	28	90	29
Reading hard cases (corresponds to red colour in the top row of Fig. 1)						
Right[a]	Lateral occipital cortex, superior division	25,279	7.04	35	23	61
Right[a]	Frontal pole	1038	4.96	21	83	40
Right	Frontal pole	365	4.28	26	90	31
Left	Paracingulate gyrus	337	3.62	47	67	60
Right	Frontal pole	211	5.29	25	88	43
Right	Middle temporal gyrus, anterior division	182	3.79	15	65	25
Right	Superior temporal gyrus, posterior division	178	3.73	11	55	37
Left	Superior frontal gyrus	68	5.51	48	66	72
Right	Postcentral gyrus	66	3.33	12	58	49
Right	Temporal pole	54	3.28	22	67	26

The same anatomical label appears more than once in some instances. The reason for this is simple—the anatomical labels refer to fairly large areas of the brain and our results demonstrate small clusters of significant activations within these larger anatomical areas. For instance, there were 2 main clusters of activations found in the right frontal pole when novices read easy questions; each cluster has an identified location within the frontal pole, as located by the MNI coordinates in 3D space, and the size of the clusters activated within the frontal pole are different; one cluster was found to be 157 voxels in size, while the other is only 53 voxels

[a] Cluster list of activations

Table 2 Common of neural activation (clusters) for experts reading easy and hard clinical cases

Cerebral hemisphere	Neural area of activation: anatomical labels	Cluster size (number of activated voxels)	Max Z (maximum statistic value)	MNI coordinates (locations of activated voxels using 3D coordinates)		
				X	Y	Z
Reading easy cases (corresponds to blue colour in the bottom row of Fig. 1)						
Left[a]	Inferior temporal gyrus, temporooccipital part	12,538	5.14	69	33	28
Left[a]	Middle frontal gyrus	4736	4.95	70	79	48
Right[a]	Middle frontal gyrus	1079	4.49	18	75	53
Left	Superior temporal gyrus, posterior division	311	3.74	70	49	34
Left	Paracingulate gyrus	272	3.51	48	81	54
Left	Frontal pole	184	3.67	58	96	37
Right	Frontal pole	169	4.65	35	97	40
Right	Precentral gyrus	148	3.8	24	65	47
Right	Insular cortex	77	3.26	30	74	34
Right	Frontal pole	60	3.37	42	93	50
Reading hard cases (corresponds to red colour in the bottom row of Fig. 1)						
Left[a]	Inferior temporal gyrus, temporooccipital part	18,308	5.23	69	39	23
Left[a]	Superior frontal gyrus	8800	5.19	59	77	64
Left	Frontal pole	132	3.92	59	92	46
Right	Precentral gyrus	123	3.54	29	62	55
Left	Middle temporal gyrus, anterior division	122	3.62	74	61	27
Right	Frontal pole	99	4.05	39	98	31
Right	Frontal pole	69	3.49	36	96	43
Right	Frontal pole	69	3.34	19	83	35
Right	Frontal pole	62	3.74	31	96	39
Left	Frontal orbital cortex	58	3.6	58	80	24

[a] Cluster list of activations

hemisphere. In the quadrant labeled "Novice Right" of Fig. 1 the Occipital lobe is primarily purple, indicating (unsurprisingly) that this region is significantly engaged while reading both hard and easy clinical cases. Similarly, the bottom row of Fig. 1 shows the areas of activation in the right and left hemisphere of the expert brain and these correspond to Table 2. Combining these data, there were common neural areas of activity in both novice and experts while reasoning through clinical cases, including the right middle frontal gyrus, right precentral gyrus, left frontal pole, right frontal pole, left frontal orbital cortex, and left superior frontal gyrus.

To assess whether areas of activation differed between novices and experts when cases were straightforward (easy) or more complex (hard) we compared novice-expert differences (novice > expert; expert > novice) for both the easy and hard questions. Analyses revealed that there were significantly greater activations in the left temporal pole and anterior division of the left middle temporal gyrus for novices relative to experts when reading both easy and hard clinical cases (see Fig. 2; Table 3). In other words, activation in the left anterior temporal lobe was greater for novices than for experts (red and blue colour in Fig. 2)—but was the same for both the reading of easy and hard clinical cases. There were no clusters that reached significance when we contrasted experts > novices for both easy and hard questions.

In a final series of analyses, we conducted specific ROI analyses to examine activity in the PFC (specifically DLPFC and VLPFC). Here, we first assessed areas of activation within expertise level (hard > easy; easy > hard), the results of these analyses were then combined to determine if neural areas of activation could be identified based upon expertise (novice > expert; expert > novice). The results of these analyses (see Fig. 2; Table 4) revealed that for novices, reading more difficult clinical cases resulted in greater prefrontal activation than reading easy clinical cases, and further, this activation was greater in novices than in experts.

Discussion

The relationship between clinical reasoning and WM was explored using fMRI. The significant findings from this study were: (1) There were common neural areas of activity in both novice and experts while reasoning through clinical cases, including the right middle frontal gyrus, right precentral gyrus, left frontal pole, right frontal pole, left frontal orbital cortex, and left superior frontal gyrus; (2) There were greater activations in the left anterior temporal lobe for novices relative to experts when reading both easy and hard clinical

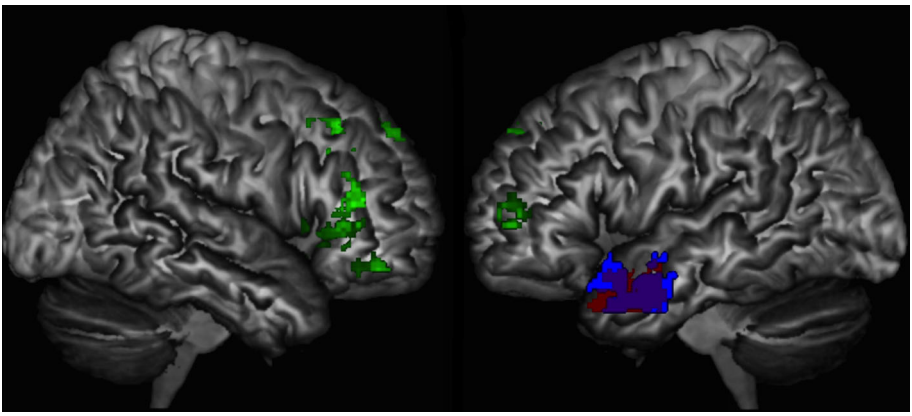

Fig. 2 novice-expert differences in clinical reasoning. There was significantly more activation in the left temporal lobe of the novices for both easy (*blue*) and hard (*red*) clinical cases [novice > expert, easy cases (*blue*); novice > expert, hard cases (*red*)]. Region of interest analysis contrasting expertise and case difficulty indicate significantly more prefrontal cortex activity in novices for hard cases (*green*),[novice > expert, hard > easy cases (*green*)]. For specific anatomical labels refer to Table 4. (Color figure online)

Table 3 Novice > expert activations for reading easy and hard cases

Cerebral hemisphere	Neural area of activation: anatomical labels	Cluster size (number of activated voxels)	Max Z (maximum statistic value)	MNI coordinates (locations of activated voxels using 3D coordinates)		
				X	Y	Z
Easy cases (corresponds to blue colour in Fig. 2)						
Left	Temporal pole	472	4.15	−56	16	−8
Hard cases (corresponds to red colour in Fig. 2)						
Left	Middle temporal gyrus, anterior division	560	5.28	−58	−2	−18

Table 4 Areas of activation in regions of interest: novice > expert; hard > easy

Cerebral hemisphere	Neural area of activation: anatomical labels	Cluster size (number of activated voxels)	Max Z (maximum statistic value)	MNI coordinates (locations of activated voxels using 3D coordinates)		
				X	Y	Z
Right	Superior frontal gyrus	183	2.49	35	79	56
Right	Frontal pole	92	3.24	21	82	38
Right	Superior frontal gyrus	49	2.87	43	84	58
Right	Frontal pole	41	3.47	28	86	29
Right	Frontal pole	38	2.27	22	83	43
Left	Frontal pole	38	3.18	64	89	38

This table summarizes the interaction effects of group comparisons (novice > expert) and case difficulty comparison (hard > easy). Results reveal novices demonstrated significantly more neural activations than experts on hard questions in the listed anatomical areas

cases; and (3) There was increased prefrontal activity for hard questions in novices relative to experts, and this increase in activation is also relative to easy questions. Stated differently, significantly more prefrontal activity (a region implicated with WM as outlined below) was required for novices than for experts, especially as they reasoned through hard clinical cases.

As hypothesized, there exist common neural areas of activations in novice and expert clinicians associated with WM when reasoning through clinical cases. Common prefrontal activations found in the present study are consistent with areas identified in a meta-analysis of 189 neuroimaging studies related to WM (Rottschy et al. 2012). As such, this reinforces our initial expectations of WM being critical to clinical reasoning. As a result of finding shared neural areas of activation across all levels of expertise and task difficulties, one

could infer that activation of the PFC and use of WM is an index, or measure of, clinical reasoning.

Differences in PFC activity between novice and experts, in which novices demonstrated increased neural activation for both easy and hard clinical cases, could also suggest novices utilize WM more so than experts. More significant prefrontal activations in novices could be explained by the importance of this region in WM in guiding semantic memory retrieval to reason through scenarios, especially during complex clinical cases in which there are more competing clinical distractors (Cabeza and Kingstone 2006; Collette and Van der Linden 2002). When specifically considering increased right frontal polar areas of activations in novices, other work has suggested this area is important in episodic memory tasks requiring ongoing monitoring during retrieval (Velanova et al. 2003). As well, while some have suggested decreased prefrontal activity as a hallmark of expertise (Durning et al. 2015b), an alternative view is that neural areas of activation shift or functionally reorganize as expertise and knowledge structures develop (Guida et al. 2012; Haller and Radue 2005). Our findings of significantly increased prefrontal activations in novices could, therefore, be related to WM guiding semantic retrieval, ongoing monitoring during episodic retrieval, or due to functional reorganization in expertise.

We have demonstrated instances when neural areas diverge based on task complexity or clinician level of expertise. By using more drastic contrasts in level of expertise and question difficulty than in previous work, our research findings support the comment that in the trajectory of development to expertise, there may be phases of neural patterns exhibited in clinical reasoning (Durning et al. 2015b). We have also identified that reading clinical cases, considered the reasoning phase in our work, produces distinct demands on neural activity, and support previously noted concerns that reading might not serve as the best baseline contrast for fMRI research in medical education (Durning et al. 2012).

Other areas of activation

Activation of the left anterior temporal lobe in novices for both easy and hard clinical cases suggests this group relies more heavily on LTM during clinical reasoning. The anterior temporal lobe is associated with human conceptual knowledge, and more specifically the meaning of words and objects across many domains (Rogers et al. 2004, 2006). This finding is less frequent in fMRI research, which more commonly attributes semantic processing to the medial temporal lobe, but supports the notion that anterior temporal lobes have greater recruitment when more precise recall of information is required in semantic tasks (Rogers et al. 2006). This point has been explored in language network neuroimaging studies, where the anterior temporal lobe is noted to be important in text comprehension and for creating coherent representations of dialogue or information (Ferstl et al. 2008; McRae and Jones 2012; Moss et al. 2011). Consequently, semantic processing can be understood as important for language processing and as well for accessing knowledge in clinical reasoning processes (Binder et al. 2009). There have also been accounts of the anterior temporal lobe being activated during the retrieval of abstract concepts, specifically related to words, which are required for judgments (Cabeza and Kingstone 2006). Given the type of information presented in clinical cases is more familiar and less abstract to experts, the anterior temporal lobes may not have shown significant recruitment as less effort for creating coherent representations of information is needed because the knowledge presented is more general or common place to that level of clinician.

Limitations

There are limitations to this study. First, focusing specifically on clinical cases related to gastrointestinal illness tests only one area of clinical knowledge, which may not be representative of how one reasons through other clinical aspects of knowledge related to other body systems or clinical topics. Second, information presented had no associated images, and so results cannot be generalized to clinical reasoning tasks that are visuo-spatial in nature. Third, multiple cognitive processes could elicit similar brain activation; in this study we infer that WM is associated with PFC activation as a result of the research design which engaged the participants in a task representative of clinical reasoning and analyses performed.

Conclusion

Our work demonstrates the role and importance of WM to clinical reasoning. Clinician level of expertise elicits differences in neural areas activated in clinical reasoning tasks, and demonstrate that novice clinicians rely more heavily on WM than experts, especially during hard tasks. Importantly, our research provides a more direct understanding of neural areas activated in clinical reasoning. Though our research design explored a specific type and context of medical knowledge, these contributions can improve our understanding of how novices and experts access and use WM during clinical reasoning. Continuing to partition phases of clinical reasoning and decision making in future fMRI research may also be of importance within medical education, as results to date offer different perspectives on what constitutes a reasoning phase, and when reasoning and decision making are actually occurring. It is plausible reasoning and decision-making are happening at different points of time for novices and experts, and disentangling each of these processes as well as clarifying what constitutes a reasoning task in clinical education might allow for more thoughtful future investigations.

Appendix 1: example clinical case: easy (concordant)

A 38 year-old man, diagnosed 18 years ago with ulcerative colitis, is referred to your outpatient clinic with itching and abnormal liver enzymes.

Past medical history

Non-smoker, drinks 1–2 beers/day. No history of blood transfusions, IV drug use or high-risk sexual behaviour. No history of psychiatric illness. No family history of liver disease. Has had 3 courses of prednisone for ulcerative colitis flares approximately every 6 years.

Recent history

Just finished a course of corticosteroids (prednisone) 6 months ago. Complaining of a mild, generalized itching in the last 3 months without demonstrable skin rash.

Medications

Asacol (mesalamine) 3 g/d for ulcerative colitis maintenance.

Assessment

No diabetes, no joint pains, no lung disease. Physical exam normal. Normal body mass index. Has one soft non-bloody bowel movement/day. Eating well. No abdominal pain.

Additional tests

Ultrasound results: normal gallbladder with no biliary dilation.

Lab findings

ALT 45 (7–40)
AST 32 (5–35)
Alk Phos 536 (30–145)
GGT 540 (20–35)
Total Bili 12 (5–22)
Hgb 155 (140–180)
WBC 9.1 (3.5–12)
PLT 180 (150–400)
INR 1.0 (0.9–1.1)

Appendix 2: example clinical case: hard (discordant)

A 38-year-old man diagnosed with ulcerative colitis 18 years ago, is referred to your outpatient clinic with itching and abnormal liver enzymes.

Previous medical history

Had one initial course of prednisone for first ulcerative colitis flare, but since then requires no medications. Was a daily IV heroin user from the ages of 18–21. Non-smoker. Emphatic that he does not drink more than 1–2 beers/day. No family history of liver disease. No diabetes, no joint pain, no lung disease, no psychiatric history.

Recent history

States eating well but feels quite nauseated lately with occasional abdominal cramping. Has one soft, non-bloody bowel movement/day. Is itchy over the last 3 months, but usually gets eczema around this time of year.

Medications

Tylenol 2 tabs/day for abdominal cramping.

Assessment

Normal BMI. Patches of eczema.

Additional tests

Ultrasound showed normal gallbladder and no biliary dilatation.

Lab findings

ALT 465 (7–40)
AST 432 (5–35)
Alk Phos 106 (30–145)
GGT 34 (20–35)
Total Bili 22 (5–22)
Conj Bili 16
Hgb 145 (140–180)
WBC 8.1 (3.5–12)
PLT 130 (150–400)
INR 1.2 (0.9–1.1)

References

Amaro, E, Jr, & Barker, G. J. (2006). Study design in fMRI: Basic principles. *Brain and Cognition, 60*(3), 220–232.

Ashby, F. G. (2011). *Statistical analysis of FMRI data*. Cambridge: MIT Press.

Attwell, D., & Iadecola, C. (2002). The neural basis of functional brain imaging signals. *Trends in Neurosciences, 25*(12), 621–625.

Baddeley, A. (1992). Working memory. *Science (New York, N.Y.), 255*(5044), 556–559.

Baddeley, A. (1996). Exploring the central executive. *The Quarterly Journal of Experimental Psychology, 49*(1), 5–28.

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 4*(11), 417–423.

Baddeley, A., Eysenck, M. W., & Anderson, M. C. (2014). *Memory*. Hove: Psychology Press.

Bahrami, P., Graham, S. J., Grantcharov, T. P., Cusimano, M. D., Rotstein, O. D., Mansur, A., & Schweizer, T. A. (2014). Neuroanatomical correlates of laparoscopic surgery training. *Surgical Endoscopy, 28*(7), 2189–2198.

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex, 19*(12), 2767–2796.

Cabeza, R., & Kingstone, A. (2006). *Handbook of functional neuroimaging of cognition*. Cambridge: MIT Press.

Carpenter, P. A., Just, M. A., & Reichle, E. D. (2000). Working memory and executive function: Evidence from neuroimaging. *Current Opinion in Neurobiology, 10*(2), 195–199.

Charlin, B., Tardif, J., & Boshuizen, H. P. (2000). Scripts and medical diagnostic knowledge: Theory and applications for clinical reasoning instruction and research. *Academic Medicine, 75*(2), 182–190.

Coderre, S., Jenkins, D., & McLaughlin, K. (2009). Qualitative differences in knowledge structure are associated with diagnostic performance in medical students. *Advances in Health Sciences Education, 14*(5), 677–684.

Coderre, S., Wright, B., & McLaughlin, K. (2010). To think is good: Querying an initial hypothesis reduces diagnostic error in medical students. *Academic Medicine: Journal of the Association of American Medical Colleges, 85*(7), 1125–1129.

Collette, F., & Van der Linden, M. (2002). Brain imaging of the central executive component of working memory. *Neuroscience and Biobehavioral Reviews, 26*(2), 105–125.
Croskerry, P. (2009). A universal model of diagnostic reasoning. *Academic Medicine: Journal of the Association of American Medical Colleges, 84*(8), 1022–1028.
Downar, J., Bhatt, M., & Montague, P. R. (2011). Neural correlates of effective learning in experienced medical decision-makers. *PLoS One, 6*(11), e27768.
Durning, S. J., Capaldi, V. F., II, Artino, A. R., Jr., Graner, J., Van Der Vleuten, C., Beckman, T. J., et al. (2014). A pilot study exploring the relationship between internists' self-reported sleepiness, performance on multiple-choice exam items and prefrontal cortex activity. *Medical Teacher, 36*(5), 434–440.
Durning, S. J., Costanzo, M. E., & Artino, A. R. (2015a). Neural basis of nonanalytical reasoning expertise during clinical evaluation. *Brain and Behavior, 5*(3), e0309.
Durning, S. J., Costanzo, M., Beckman, T., Artino Jr., A., Roy, M., Van Der Vleuten, C. (2015b) Functional neuroimaging correlates of thinking flexibility and knowledge structure in memory: Exploring the relationships between clinical reasoning and diagnostic thinking. *Medical Teacher, 16*, 1–8.
Durning, S. J., Elnicki, M. E., Gruppen, L., Torre, D., & Hemmer, P. A. (2010). AMEE 2009 spotlight on educational research. *Medical Teacher, 32*(4), 340–342.
Durning, S. J., Graner, J., Artino, A. R., Pangaro, L. N., Beckman, T., Holmboe, E., et al. (2012). Using functional neuroimaging combined with a think-aloud protocol to explore clinical reasoning expertise in internal medicine. *Military Medicine, 177*(9 Suppl), 72–78.
Elstein, A., Shulman, L. S., & Sprafka, S. A. (1978). *Medical problem solving: An analysis of clinical reasoning*. Cambridge, MA: Harvard University Press.
Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102*(2), 211–245.
Evans, J., & Stanovich, K. (2013). dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science, 8*(3), 223–241.
Ferstl, E. C., Neumann, J., Bogler, C., & von Cramon, D. Y. (2008). The extended language network: A meta-analysis of neuroimaging studies on text comprehension. *Human Brain Mapping, 29*(5), 581–593.
Guida, A., Gobet, F., Tardieu, H., & Nicolas, S. (2012). How chunks, long-term working memory and templates offer a cognitive explanation for neuroimaging data on expertise acquisition: A two-stage framework. *Brain and Cognition, 79*(3), 221–244.
Haller, S., & Radue, E. W. (2005). What is different about a radiologist's brain? *Radiology, 236*(3), 983–989.
Heeger, D. J., & Ress, D. (2002). What does fMRI tell us about neuronal activity? *Nature Reviews Neuroscience, 3*(2), 142–151.
Heit, E., Rotello, C. M., & Hayes, B. K. (2012). Relations between memory and reasoning. *The Psychology of Learning and Motivation, 57*, 57–101.
Huettel, S. A., Song, A. W., & McCarthy, G. (2009). *Functional magnetic resonance imaging*. Sunderland: Sinauer Associates Incorporated.
Jenkinson, M. (2003). Fast, automated, N-dimensional phase-unwrapping algorithm. *Magnetic Resonance in Medicine, 49*(1), 193–197.
Jenkinson, M., & Smith, S. (2001). A global optimisation method for robust affine registration of brain images. *Medical Image Analysis, 5*(2), 143–156.
Logothetis, N. K. (2003). The underpinnings of the BOLD functional magnetic resonance imaging signal. *The Journal of Neuroscience, 23*(10), 3963–3971.
McCarthy, P. (2015, August 3). *GLM*. Retrieved from http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/GLM.
McLaughlin, K., Coderre, S., Mortis, G., Fick, G., & Mandin, H. (2006). Can concept sorting provide a reliable, valid and sensitive measure of medical knowledge structure? *Advances in Health Sciences Education, 12*(3), 265–278.
McRae, K., & Jones, M. (2012). *Semantic memory. The Oxford handbook of cognitive psychology*. Oxford: Oxford University Press.
Melo, M., Scarpin, D. J., Amaro, E., Passos, R. B. D., Sato, J. R., Friston, K. J., & Price, C. (2011). How doctors generate diagnostic hypotheses: A study of radiological diagnosis with functional magnetic resonance imaging. *PLoS One, 6*(12), e28752.
Moss, J., Schunn, C. D., Schneider, W., McNamara, D. S., & VanLehn, K. (2011). The neural correlates of strategic reading comprehension: Cognitive control and discourse comprehension. *NeuroImage, 58*(2), 675–686. doi:10.1016/j.neuroimage.2011.06.034.
Norman, G. (2005). Research in clinical reasoning: Past history and current trends. *Medical Education, 39*(4), 418–427.

Oldfield, R. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*(1), 97–113.

Pelaccia, T., Tardif, J., Triby, E., & Charlin, B. (2011). An analysis of clinical reasoning through a recent and comprehensive approach: The dual-process theory. *Medical Education Online*, *16*. doi:10.3402/meo.v16i0.5890.

Poldrack, R. A. (2007). Region of interest analysis for fMRI. *Social Cognitive and Affective Neuroscience, 2*(1), 67–70.

Poldrack, R. A., Mumford, J. A., & Nichols, T. E. (2011). *Handbook of functional MRI data analysis*. Cambridge: Cambridge University Press.

Ranganath, C., Johnson, M., & D'Esposito, M. (2003). Prefrontal activity associated with working memory and episodic long-term memory. *Neuropsychologia, 41*(3), 378–389.

Rogers, T., Hocking, J., Noppeney, U., Mechelli, A., Gorno-Tempini, M., Patterson, K., & Price, C. (2006). Anterior temporal cortex and semantic memory: Reconciling findings from neuropsychology and functional imaging. *Cognitive, Affective and Behavioral Neuroscience, 6*(3), 201–213.

Rogers, T., Lambon Ralph, M. A., Garrard, P., Bozeat, S., McClelland, J. L., Hodges, J. R., & Patterson, K. (2004). Structure and deterioration of semantic memory: A neuropsychological and computational investigation. *Psychological Review, 111*(1), 205–235.

Rosler, F., Ranganath, C., Röder, B., & Kluwe, R. (2009). *Neuroimaging of human memory*. Oxford: Oxford University Press.

Rottschy, C., Langner, R., Dogan, I., Reetz, K., Laird, A. R., Schulz, J. B., et al. (2012). Modelling neural correlates of working memory: A coordinate-based meta-analysis. *NeuroImage, 60*(1), 830–846.

Ruff, C. C., Knauff, M., Fangmeier, T., & Spreer, J. (2003). Reasoning and working memory: Common and distinct neuronal processes. *Neuropsychologia, 41*(9), 1241–1253.

Sanfey, A. G., & Chang, L. J. (2008). Multiple systems in decision making. *Annals of the New York Academy of Sciences, 1128*(1), 53–62.

Savoy, R. L. (2001). History and future directions of human brain mapping and functional neuroimaging. *Acta Psychologica, 107*(1–3), 9–42.

Savoy, R. L. (2006). Using small numbers of subjects in fMRI-based research. *IEEE Engineering in Medicine and Biology Magazine: The Quarterly Magazine of the Engineering in Medicine and Biology Society, 25*(2), 52–59.

Simmons, B. (2010). Clinical reasoning: Concept analysis. *Journal of Advanced Nursing, 66*(5), 1151–1158.

Simons, J. S., & Spiers, H. J. (2003). Prefrontal and medial temporal lobe interactions in long-term memory. *Nature Reviews Neuroscience, 4*(8), 637–648.

Smith, S. M. (2004). Overview of fMRI analysis. *The British Journal of Radiology, 77*(suppl_2), S167–S175.

Süß, H. M., Oberauer, K., Wittmann, W. W., & Wilhelm, O. (2002). Working-memory capacity explains reasoning ability—and a little bit more. *Intelligence, 30*(3), 261–288.

Tadel, F. (2015, August 10). *Coordinate systems*. Retrieved from http://neuroimage.usc.edu/brainstorm/CoordinateSystems.

Velanova, K., Jacoby, L. L., Wheeler, M. E., McAvoy, M. P., Petersen, S. E., & Buckner, R. L. (2003). Functional–anatomic correlates of sustained and transient processing components engaged during controlled retrieval. *The Journal of Neuroscience, 23*(24), 8460–8470.

Worsley, K. J., Evans, A. C., & Marrett, S. (1992). A three-dimensional statistical analysis for CBF activation studies in human brain. *Journal of Cerebral Blood Flow Metabolism, 12*(6), 900–918.