Title
Is Insight Always the Same? An fMRI Study of Insight

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 33(33)

ISSN
1069-7977

Authors
Cranford, Edward
Moss, Jarrod

Publication Date
2011

Peer reviewed

An fMRI Study of Insight Using Compound Remote Associate Problems

Edward A. Cranford (eac53@msstate.edu)
Jarrod Moss (jarrod.moss@msstate.edu)
Department of Psychology, Mississippi State University
Mississippi State, MS 39762 USA

Abstract

Prior neuroimaging studies of insight problem solving using Compound Remote Associate (CRA) problems provide consistent results. However, in a prior study (Cranford & Moss, 2010), we found that participants derive solutions by insight in at least two different ways. In the present study, we attempted to extend upon prior studies by dividing insight solutions into two categories: immediate and delayed. The results show a large difference between the pattern of activation for immediate-insight solutions and delayed-insight solutions. Future research may benefit from distinguishing between types of insight in CRA problem solving.

Keywords: Insight; Problem Solving; Restructuring; fMRI

Introduction

Problem solving as a process is what enables humans to discover solutions to even the most difficult of problems. One way to obtain a solution is through a 'search' process. The solver engages in a structured exploration of possible solution paths in order to find a solution. Sometimes the solution appears suddenly, without warning, and is termed insightful. Insight solutions often appear from nowhere and solvers experience an affective response of suddenness and surprise (Aha! experience), sometimes resulting after an impasse; insight solutions are obtained through processes known as restructuring, whereby an incorrect representation of the problem is changed, leading to the access of an insightful, correct representation of the problem (Bowden, & Jung-Beeman, 2007; Ohlsson, 1992; Schooler, Fallshore, & Fiore, 1995). Solutions obtained through search do not generate the 'Aha' experience, and the solution idea seems to be a continuation of previously generated ideas. The key components of insight are often described as impasse, restructuring, and 'Aha!'. However, the process is still not fully understood. Neuroimaging techniques and more consistent experimental tasks have the potential to further our understanding of the neural correlates of insight.

Utilizing neuroimaging methods such as functional magnetic resonance imaging (fMRI) limits the type of task that can be performed. Classic insight problems usually take a long time to solve and are often different than noninsight problems. Therefore, the Compound Remote Associate (CRA) task was developed by Bowden and Jung-Beeman (2003). The CRA problem consists of three words presented to the solver. The solver must come up with a single fourth word that can be combined with each of the other three words to form new compound words or common phrases. For example, if three words—tree, sauce, and big—are presented, the solution is apple. CRA problems can be solved quickly so many problems can be presented in one fMRI scanning session. Also, CRA problems can be solved by insight or by noninsight, search processes (i.e., generate-and-test or trial-and-error), and individual problems can be solved with insight regardless of learning effects over multiple trials (Bowden & Jung-Beeman, 2007).

Prior fMRI studies using the CRA problem to investigate the neural correlates of insight yield somewhat consistent results and offer compelling theoretical arguments for the activation seen in these brain regions. Jung-Beeman and colleagues (2004) expanded on their theory that course semantic coding occurs in the right hemisphere (RH) while fine coding occurs in the left hemisphere (LH) and found that an area in the RH anterior superior temporal gyrus (RH-aSTG) was more active at solution for insight than noninsight solutions. Activity in this region creates broad associations in memory so that seemingly disparate concepts converge on a solution and suddenly emerge into consciousness as an insight. Subramaniam, Kounios, Parrish, and Jung-Beeman (2008) reported a similar region in their results (RH middle temporal gyrus; MTG). Both studies reported activity in the anterior cingulate cortex (ACC). The function of the ACC in insight problem solving is to monitor for competing responses for attention, as is its role in the cognitive control network (i.e., Cole & Schneider, 2007). Other 'insight' areas noted in the two studies are the posterior cingulate cortex (PCC), parahippocampus (PH), right superior frontal gyrus (SFG), and right inferior parietal lobe (IPL).

Even prior to seeing a problem, certain brain states have been found to predict future solution by insight (Kounios et al., 2006). Kounios and colleagues observed activity in the ACC, PCC, and bilateral middle/superior temporal gyrus (M/STG) prior to seeing a problem. These same areas were observed by Subramaniam and colleagues (2008). The ACC may be active due to an increased readiness to apply cognitive control in order to suppress thoughts, initially select a solution space, and, if needed, to switch attention. The M/STG may be active due to a preparation for semantic activation and retrieval of associations. Finally, the PCC may simply reflect differences in attentional demands between preparation periods preceding insight and those preceding noninsight solutions. Although the results from these three studies are informative, they often report many regions of activation that do not overlap across studies using very similar methodology. Clearly some regions such as ACC and right MTG are consistently active, but one possible explanation for areas that do not overlap across studies is that the methodology employed may be averaging over at least two distinct solution processes.

In an initial exploration of insight using CRA problems, we examined whether CRA problems rated by participants as solved by insight actually showed any observable characteristics of insight problem solving as it has traditionally been defined (Cranford & Moss, 2010). In this study, participants attempted 60 CRA problems. For half of the problems, participants verbalized their solution processes. We used the verbal protocols as data which were coded for characteristics of insight (i.e., occurrences of impasses and restructuring elements). Participants rated each solution as insight, search, or other. Initial analyses revealed no differences between insight and search solutions for rates of impasses or restructuring elements. However, we observed that insight solutions were solved in two distinct ways. The first way, termed "immediate-insight" (II), occurred when the first candidate solution verbalized by the participant was the solution and the solution occurred within 15 s of the problem being presented. The second type of insight solution, termed "delayed-insight" (DI), included all other insight solutions not classified as II.

For II solutions, a person may report the quickly solved problems as insight simply because they came to a solution so fast that it seemed sudden and surprising. However, it is unclear whether this should be called insight, traditionally defined. II solutions do not exhibit any observable signs of impasse or restructuring in the verbal protocols and are the first solution candidate reported. Notably, of solutions reported as insight, 77.73% (*SD* = 17.16%) were II and 22.3% (*SD* = 17.16%) were DI. We compared II and DI solutions to search solutions individually. There were significantly fewer impasses and restructuring elements for II than search solutions, but there were more impasses and restructuring elements for DI than search solutions. DI solutions also seem to exhibit the phenomenological characteristics unique to insight as defined here.

The results from Cranford and Moss (2010) tell us either that II solutions do not really employ insight processes but are just sudden or that they are just as insightful as DI solutions but verbal protocols can not adequately capture insight processes in this case. Either way the results of prior insight fMRI studies using CRA problems may only be telling us part of the story. There may be different, or even additional, areas that may be necessary for insight.

The present study tests the hypothesis that the pattern of activation for DI solutions differs from the pattern of activation for II solutions. In the present study, we continued the investigation of insight in an fMRI study of CRA problem solving. We followed the methods of prior fMRI studies as closely as possible (e.g., Subramaniam et al., 2008) examining both solution and preparation intervals. However, after participants categorize their solution processes as insight or search we ask whether the solution was the first word that came to mind. By teasing apart the two types of insight we should gain a better understanding of the network of brain areas associated with insight.

Method

Participants

Participants were 22 right-handed, native-English speaking undergraduates, with normal or corrected to normal vision from Mississippi State University (*M* age = 20.09, *SD* = 2.16, range = 18–26) and were paid for their participation.

Design

The design was a single factor, within-subject, study examining two levels of Problem Type (between II or DI and Search). Imaging was carried out in an event-related design, because individual CRA problems can be solved via Insight or Search depending on the solver, to assess differences in activation for problems reported as involving II or DI versus Search at solution and preparation events.

Procedure

The study took place over two sessions with fMRI data acquired only during the second session. Participants underwent training in a mock MRI (a replica of the shell of the actual MRI). During this training session, participants were familiarized with the MR environment and given feedback if they moved their head too much using a magnetic head motion tracker. First, instructions for the CRA problems were presented. Participants then read instructions for rating problems by insight or noninsight as in prior studies (e.g., Jung-Beeman et al., 2004). However, we used the label 'search' rather than 'noninsight' so participants could make an easier distinction between the two processes, one involving insight and one involving analytic, or methodological search processes. The search label was judged to be more descriptive of the instructions than 'noninsight'. Additional instructions were given for deciding whether an answer was the first thing that came to their mind (immediate solutions) or if the solution came after thinking of other things first (delayed solutions).

For each CRA problem, a fixation cross was first presented in the center of the screen for two, four, or six seconds (preparation interval). The three problem words were then presented in the center of the screen. Participants were instructed to press any button on the response pads as quickly as they could as soon as they thought they had the correct solution. If no solution was reached within a 30s time limit, the next problem was presented. Upon solution, after a second jitter interval of 2-6 seconds, participants were prompted to say their solution aloud and then press a button to continue to the next screen. Participants were then asked whether the solution was obtained via insight, search, or by some other means. After the response, to determine whether a problem was labeled Delayed or Immediate, participants were asked if the solution was the first word they thought of or if they thought of other words before the final solution. After pressing the "yes" or "no" button the next problem began, starting with the preparation interval. The task steps, from fixation cross to the last rating question, repeated for each CRA problem. Participants completed as many runs as necessary to complete all 80 problems, or as many as they could in the 90 minutes.

The MRI session consisted of 30 minutes of practice on additional CRA problems and then 90 minutes of scanning. An active noise canceling microphone system (Psychology Software Tools, Inc., Pittsburgh, PA) was attached to the head coil and used to collect verbalized solution responses. Participants attempted a total of 80 CRA problems taken from the set published by Bowden and Jung-Beeman (2003). CRA problems were projected onto a screen located behind the bore of the MRI. Structural images were acquired first and then participants completed 5-7 runs of CRA problem solving. For each run, participants solved as many problems as they could in seven minutes. The last run was of varying length depending on the number of problems left to be solved. For every two runs of CRA problem solving, participants completed a lexical decision filler task resulting in a total of 7-10 runs. The lexical decision runs were part of another study, and will not be discussed further.

Image Acquisition and Analysis

Imaging was performed on a 3T GE scanner with an 8-channel volume head coil. Functional images were acquired in an axial orientation using a T2*-weighted gradient echo-planar imaging pulse sequence (28 slices, 4 mm thick; 3.75-mm x 3.75-mm in-plane resolution; TR = 2000 ms; TE = 30 ms; Flip Angle = 79; FOV = 24).

The raw neuroimaging data were preprocessed and analyzed using the AFNI software package (Cox, 1996). Images were corrected for slice scan time, motion, spatially smoothed to 7 mm FWHM Gaussian kernel, and were aligned to the anatomical images. Images were transformed to a standard Talairach space (Talairach & Tournoux 1988).

Data were analyzed using general linear model analysis that extracted average responses to each problem type at the solution event, correcting for linear drift and excluding signal changes correlated with excessive head motion. The solution event began two seconds prior to the button press indicating a solution was found. In a replication of prior studies, only the solution event was modeled and not the entire problem solving process because, theoretically, the distinction between insight and search processes lies in the events temporally adjacent to the solution event. Utilizing a deconvolution method (9 regressors, 16 seconds per solution event), the hemodynamic response function (HRF) shape was estimated. Each problem-type (i.e., insight, search, etc.) was examined for both the solution response and the preparation response (beginning at 2 seconds prior to problem onset). The beta values from the regressors were extracted and used in a group analysis. A linear mixed-effects (LME) model analysis was used for each of two separate comparisons: Delayed-Insight vs. Search (DI-S) and Immediate-Insight vs. Search (II-S). Subject was treated as a random factor. Multiple comparison correction was performed using family-wise error (FWE) cluster size thresholding (Forman et al., 1995) to eliminate small clusters likely to be false-positives. Clusters of voxels which showed a significant difference in the shape of the estimated hemodynamic response function (HRF) by condition at a FWE corrected level of $p < .05$ are reported here.

Only 19 subjects were included in the analyses. Three participants were excluded from analysis because they solved one or fewer problems with DI and/or II.

Results

Behavioral Results

Participants attempted an average of 78.62 (*SD* = 4.02) problems out of the 80 available for the fMRI session. Participants correctly solved 49.3% (*SD* = 8.05%) of the attempted problems. For the correctly solved problems, 56.51% (*SD* = 17.19%) of their solutions were solved with insight (mean RT = 7.79 sec, *SD* = 5.46) and 34.89% (*SD* = 17.57%) were solved with search (mean RT = 12.36 sec, *SD* = 7.21). Participants gave incorrect responses to 10.54% (*SD* = 5.70%) of the solved problems, leaving 40.16% (*SD* = 12.30%) of the problems unsolved.

Insight and Search responses were labeled as either "delayed" or "immediate" solutions. Solutions were delayed if more than 10 seconds elapsed before solution and/or the participant responded "no" to the question asking if the first word they thought of was the solution. Solutions were labeled immediate if the problem was solved in less than 10 seconds and the participant responded "yes" to the question asking if the first word they thought of was the solution. This is comparable to prior distinctions (Cranford & Moss, 2010). If the solution is the first word they thought of and it occurred very quickly then there was very little evidence of restructuring or impasse. However, if the solution occurred after some delay and was not the first word they thought of then some kind of restructuring and/or impasse must have occurred. Of all the insight solutions, 34.13% (*SD* = 17.47%) were delayed and 65.87% (*SD* = 17.47%) were immediate. Of all the search solutions, 77.82% (*SD* = 16.66%) were delayed and 22.18% (*SD* = 16.66%) were immediate. Of the correctly solved problems, 19.29% (*SD* = 11.47%) were DI (mean RT = 12.28 sec, *SD* = 6.75), 37.22 % (*SD* = 14.62%) were II (mean RT = 6.01 sec, *SD* = 3.49), 27.15% (SD = 13.37%) were DS (mean RT = 13.92 sec, *SD* = 6.95), and 7.74% (SD = 7.95%) were IS (mean RT = 8.57 sec, *SD* = 6.26). Because so few problems were solved immediately by search, search solutions were not separated into immediate and delayed categories.

Imaging Results

Solution Analyses We performed a LME analysis to explore the temporal dynamics of activation when solving a problem with II versus search and DI versus search within a 2-14 second interval after the solution event. All active clusters are significant at a threshold of $p = .01$, $F > 2.808$, and a cluster size of 54 contiguous voxels. In the DI-S analysis the cluster threshold was set to 40 to capture a cluster that was just under threshold. This area, however, was significant at a comparative threshold of $p = .001$, $F > 3.781$, and a cluster size of 16 voxels.

The II-S analysis revealed many significant clusters of activation (see Figure 2) including the right M/SFG extending to the ACC, right M/STG, right thalamus and PH, middle cingulate gyrus extending to the PCC, right lingual gyrus, right precentral gyrus extending to the inferior frontal gyrus (IFG), paracentral lobe, precuneus, and left MFG. However, all of these areas were less active for II than search solutions. Figure 3 shows a representative waveform graph of the active areas seen in the II-S analysis.

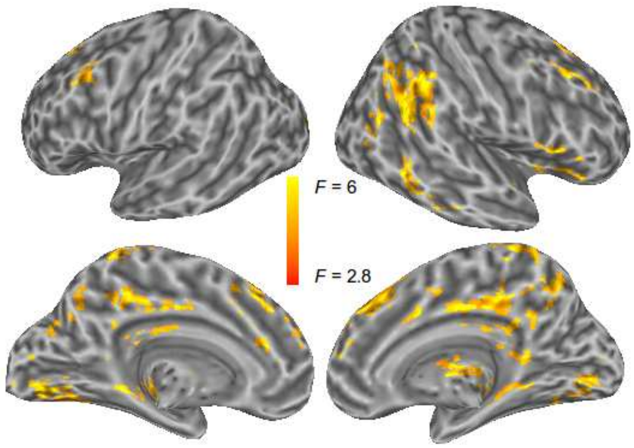

Figure 2: II-S contrast for the solution analysis.

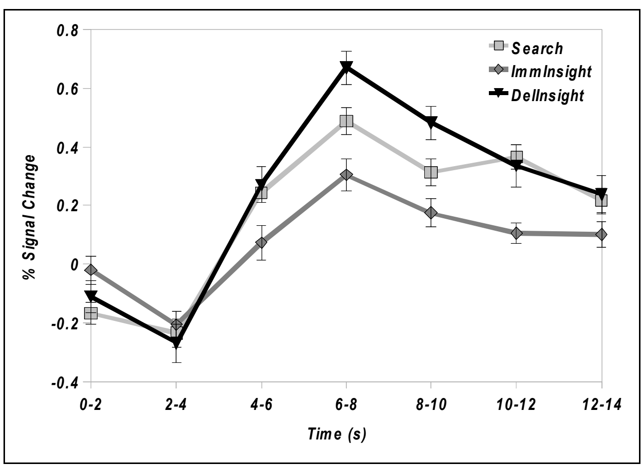

Figure 3: Representative waveform graph of II-S areas of activation in the solution analysis (right M/STG).

The DI-S analysis revealed only a few significant clusters of activation (see figure 4). Unlike the II-S analysis, the DI-S analysis does reveal areas greater for DI than search. The areas include the medial frontal gyrus/ACC, and the right MTG. Both areas have greater signal for DI solutions than search solutions and both areas peak at the solution event. Figures 5 and 6 show the estimated time course for the two areas in the DI-S analysis. These two areas map onto results from prior studies (i.e., Subramaniam et al., 2008).

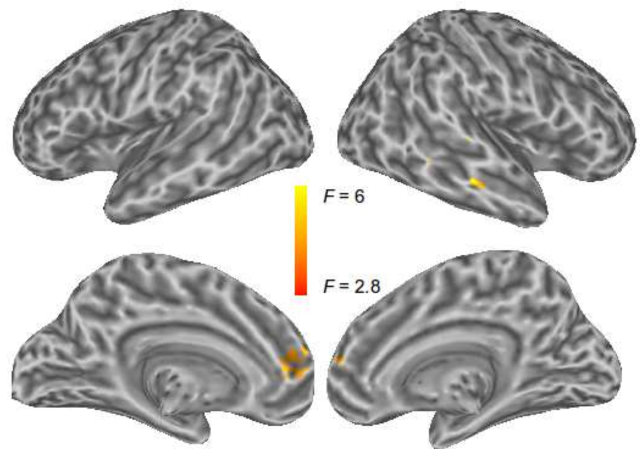

Figure 4: DI-S contrast for the solution analysis.

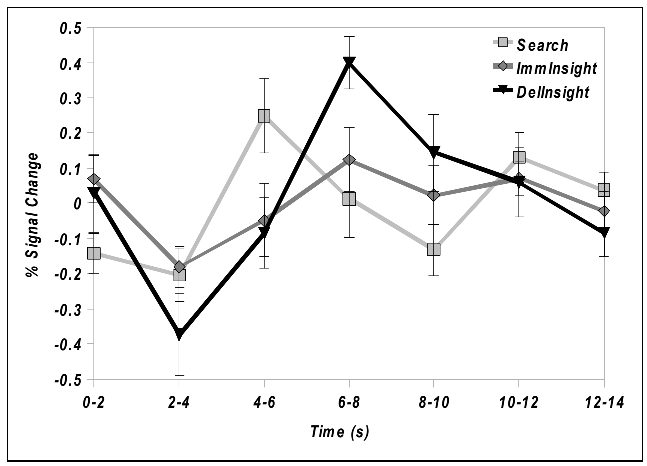

Figure 5: Waveform for DI-S contrast for the solution analysis at the medial frontal gyrus/ACC region.

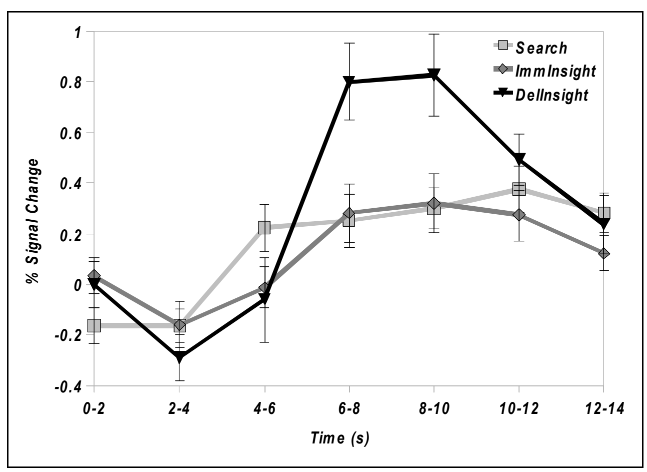

Figure 6: DI-S solution contrast at the right MTG.

An analysis of all insight solutions combined versus search solutions (I-S) was also run. Due to the large proportion of II solutions compared to DI solutions, the I-S analysis resembled the II-S analysis. Also, a contrast of DI versus II solutions closely resembled the results of the II-S contrast because DI and Search solutions have similar patterns of activity except for the areas identified in Figure 4. Due to space constraints, these contrasts are not shown.

Preparation Analyses We performed a LME analysis to explore the temporal dynamics of activation when solving a problem with II versus search and DI versus search within a 2-14 second interval after the preparation event. The DI-S preparation analysis did not reveal any significant clusters.

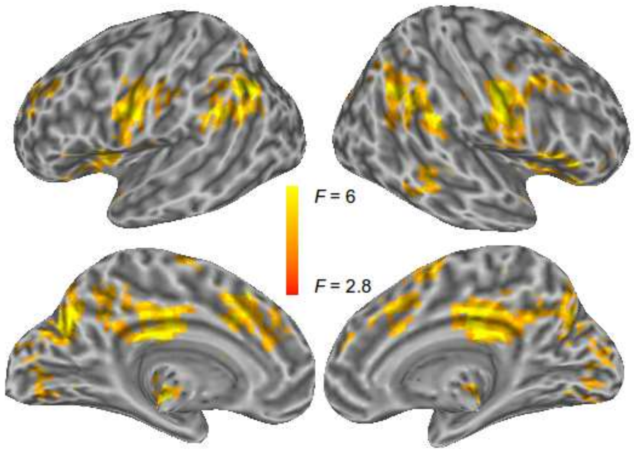

Figure 7: II-S contrast for the preparation analysis.

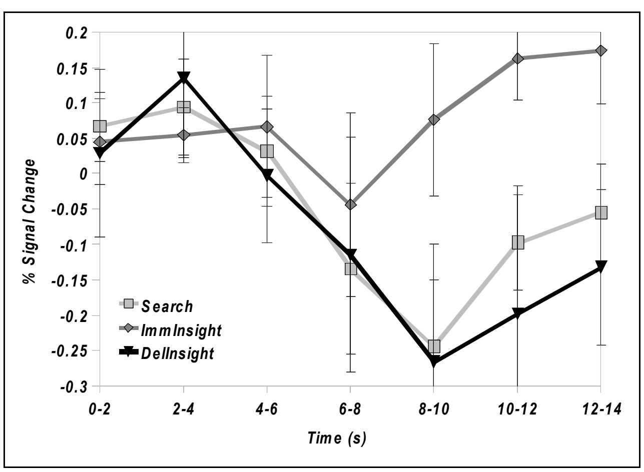

Figure 8: Representative waveform graph of II-S areas of activation in the preparation analysis. Waveforms are only different after problem onset (10 seconds after preparation interval). The actual region represented here is right IFG.

The II-S analysis revealed many significant areas of activation (see Figure 7). The areas included are the left precuneus, IPL, insula, MFG, and thalamus, the right IFG, supramarginal gyrus, M/SFG, and MTG, and the bilateral precentral gyrus. These areas are greater for II than search. However, the activation in these areas peaks well after the preparation period and overlaps with activity during problem solving periods and may not really be preparation activity but activity from actual problem solving behavior. Figure 8 shows a representative waveform graph of the active areas seen in the II-S preparation analysis.

Discussion

The results of the II-S and DI-S contrasts for the solution event provide strong support that the solution processes for the delayed and immediate insight problems were different. The shape of the estimated HRFs shown in Figures 5 and 6 for the two clusters of activation in the DI-S contrast show that both of these areas had a larger peak for DI around 6-8 s after the solution event. This 6-8 s delay is expected as the delay of the hemodynamic response is well known (Boynton, 1996). In the ACC region, it appears that there is a smaller peak for search that peaks prior to the peak of the DI response. The ACC has been proposed to be involved in conflict monitoring and cognitive control, and in particular in studies with CRA problems, the dorsal ACC has been proposed to aid in selecting between conflicting modes insightful and noninsightful problem solving or in selecting between multiple potential associations in solving CRA problems (Subramaniam et al., 2008). The peak in the ACC region for DI occurs at the same time as a more sustained peak in DI activation in the right MTG. This result could mean that the ACC is activated around the same time that the solution is being activated and retrieved from memory.

It is unclear if the earlier peak for search in the ACC is meaningful. It could be that the timing of ACC activity relative to the timing of retrieving the solution in memory is important for the affective response associated with insight. The fact that II solutions are also accompanied by an affective response argues against this interpretation because there was no difference in this ACC region in the II-S contrast. However, it could be that a similar affective response is generated by two different solution processes. Immediate insight problems may lead to the 'Aha' response due to suddenness while the DI solutions may lead to an 'Aha' because of the retrieval of a distantly related solution word that was different from words that had just previously been retrieved from memory. Retrieving a distantly related word that has little relation to solution words currently being considered might be how impasses are overcome in this type of problem. This interpretation seems to fit with the behavioral finding of more impasses and restructuring to overcome those impasses in the behavioral results of prior work using CRA problems (Cranford & Moss, 2010).

The II-S contrast results show a set of regions including the inferior frontal gyrus, MTG, parahippocampus, PCC, and inferior parietal lobe that have been associated with semantic memory retrieval (e.g., Cabeza & Nyberg, 2000). In fact, as figure 3 shows, search solution activity generally peaked higher than II solution activity in all of these regions. One interpretation of these results is just that search and DI involve more sustained memory retrieval processes than do the II solutions. The larger peaks in these regions for the more sustained problem solving processes may contribute to the larger peaks in these areas around the time of solution. Another possibility is that the II solutions were relatively more active or somehow primed to a greater degree by the problem words so that their retrieval did not require as much retrieval “effort”.

There were no clusters of activation where II solution activity was greater than search activity. This result is in contrast to the DI-S contrast where all significant regions showed a larger peak to DI solutions. These results support the idea that II and DI solution processes are distinct both behaviorally and neurally.

The preparatory activity shown in Figure 7 seems to be related to problem solving activity rather than any kind of preparation prior to problem solving that biases participants in favor of an insight solution process. As can be seen in Figure 8, II seems to show less decrease in activation than to DI and search solutions. Many of the regions significant in this analysis are part of the default network (Buckner, Andrew-Hanna, & Schacter, 2008). The default network is generally more active during rest periods of a task and less active the more a person is engaged in goal-oriented activity, such as problem solving. Also, there was no significant preparatory activity found in the DI-S contrast. These results are obviously in contrast with prior results showing clear preparatory activity prior to insight solutions (Kounios et al., 2006; Subramaniam et al., 2008). This difference in results could be due to the fact that we analyzed the 2 s period just prior to problem onset while other studies (e.g., Subramaniam et al.) have examined the start of the preparation period. Because the preparation period is jittered to be 2, 4, or 6 s due to fMRI experimental design, we may be time-locking our analysis to a different point in the preparation period than did prior work. Another consequence of examining activity 2 s prior to problem solving is that there is no jitter between preparation and start of problem solving making it hard to pull apart activation associated with the preparation period and activation associated with the start of problem solving.

In conclusion, the results of this fMRI study show that there is a difference between insight solutions obtained very soon after presentation of the problem and those obtained after significant search. This distinction was motivated by behavioral work showing that DI solutions were more characteristic of the classic definition of insight problem solving. Prior studies of insight problem solving using CRA problems have not made this distinction. The results of our DI-S contrast are in accord with prior results showing insight-related activity in right MTG and ACC. These results provide further evidence for the role of these areas in insight while at the same time demonstrating the benefit of separating two different solution processes that are both classified as insight by problem solvers.

Acknowledgments

We would like to thank Jennifer Kueven, Devin Busha, and McKensie Martin for help with data collection. We would like to specifically acknowledge the unique contributions of Joshua Liddell to this research during his time with us.

References

Bowden, E. M., & Jung-Beeman, M. (2003). Normative data for 144 compound remote associate problems. *Behavioral Research Methods, Instruments, and Computers, 35*, 634-639.

Bowden, E. M., & Jung-Beeman, M. (2007). Methods for investigating the neural components of insight. *Methods*, *42*(1), 87-99.

Boynton, G. M., Engel, S. A., Glover, G. H., & Heeger, D. J. (1996). Linear systems analysis of functional magnetic resonance imaging in human V1. *Journal of Neuroscience*, *16*(13), 4207-4221.

Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The brain's default network: Anatomy, function, and relevance to disease. *Annals of the New York Academy of Sciences*, *1124*, 1-38.

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

Cole, M., & Schneider, W. (2007). The cognitive control network: Integrated cortical regions with dissociable functions. *Neuroimage*, *37*(1), 343-360.

Cox, R. W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research*, *29*(3), 162-173.

Cranford, E., & Moss, J. (2010). Investigating insight using compound remote associate problems. *Proceedings of the Thirty-second Annual Conference of the Cognitive Science Society.*

Forman, S. D., Cohen, J.D., Fitzgerald, M., Eddy, W., Mintun, M., & Noll, D. (1995). Improved assessment of significant activation in functional Magnetic Resonance Imaging (fMRI): Use of a cluster-size threshold. *Magnetic Resonance in Medicine*, *33*(5), 636-647.

Jung-Beeman, M., Bowden, E. M., Haberman, J., Frymiare, J. L., Arambel-Liu, S., Greenblatt, R., et al. (2004). Neural Activity When People Solve Verbal Problems with Insight. *PLoS Biology*, *2*(4), 500-510.

Kounios, J., Frymiare, J. L., Bowden, E. M., Fleck, J. I., Subramaniam, K., Parrish, T. B., et al. (2006). The Prepared Mind: Neural Activity Prior to Problem Presentation Predicts Subsequent Solution by Sudden Insight. *Psychological Science*, *17*(10), 882-890.

Ohlsson, S. (1992). Information-processing explanations of insight and related phenomena. In M. T. Keane & K. J. Gilhooly (Eds.), Advances in the psychology of thinking (pp. 1-43). London: Harvester Wheatsheaf.

Schooler, J. W., Fallshore, M., & Fiore, S. M. (1995). Epilogue: Putting insight into perspective. *The nature of insight.*, 559-587.

Subramaniam, K., Kounios, J., Parrish, T. B., & Jung-Beeman, M. (2009). A brain mechanism for facilitation of insight by positive affect. *Journal of Cognitive Neuroscience*, *21*(3), 415-432.

Talairach, J., & Tournoux, P. (1988). Co-*planar stereotaxic atlas of the human brain*. New York: Thieme.