WHAT HAS FUNCTIONAL NEUROIMAGING TOLD US ABOUT THE MIND? SO MANY EXAMPLES, SO LITTLE SPACE

John Jonides, Derek E. Nee and Marc G. Berman

(Department of Psychology, University of Michigan, Ann Arbor, MI, USA)

How is it that one distinguishes between psychological theories? The dominant strategy is to test for dissociations. Let us say that theory a (T_a) is a proposal about a mechanism that underlies some behavior. Theory b (T_b) is an alternative proposal about a mechanism to account for the same behavior. Put differently, according to T_a, there is some system of psychological processes that underlies performance in some task. According to T_b, performance in this same task is mediated by a different system of processes. Both theories account for the same set of data. Of course, were it the case that T_a and T_b always made the very same predictions about all behaviors, distinguishing between the theories would be a matter of taste, not a matter of science. T_a and T_b are interestingly different only if there is at least one circumstance in which they predict different behaviors. It is this circumstance that allows the theories to be distinguished.

Let us consider one example raised by Coltheart (2006, this issue) to see what impact neuroimaging data have on distinguishing between alternative theories. Smith and Jonides (1997) reviewed what we knew at the time about the cognitive (not neural) architecture of working memory. That review featured the model of Baddeley, which theorizes that working memory for verbal information is mediated by a different cognitive system than working memory for spatial information. Let us call this T_a. The alternative theory not recognized by Coltheart (2006, this issue), but an obvious contrast to T_a, is that the architecture of working memory is singular in form, no matter the type of information stored. According to this theory, T_b, there is a single system of storage and rehearsal processes that operates on both verbal and spatial information.

These two opposing theories can be resolved through behavioral evidence, neuropsychological evidence (lesion data), and neuroimaging evidence, all of which support Baddeley's (1992) theory of separable cognitive systems mediating working memory (T_a). Dual-task paradigms have been used successfully to demonstrate the separability of spatial and verbal aspects of working memory. For example, den Heyer and Barrett (1971) conducted a dual-task experiment in which subjects had to perform a primary task of remembering either the identity or the spatial location of letters on a grid. During the retention-interval, subjects were given a secondary task that was either spatial or verbal. Subjects' performance was substantially worse when the primary and secondary tasks were either both spatial or both verbal compared to when they were mixed. This example illustrates a dissociation between spatial and verbal working memory in that subjects are worse at performing two verbal or two spatial tasks concurrently compared to performing a mixture of verbal and spatial tasks.

Neuropsychology has also provided evidence showing the separability of verbal and spatial working memory. Basso et al. (1982) studied patient P.V. who had a large lesion in her left hemisphere. P.V. had intact language processes but was severely impaired in verbal working memory. Her spatial working memory span, by contrast, was actually better than normal. Hanley et al. (1991) studied patient E.L.D. who had a large lesion in her right hemisphere and showed impaired spatial working memory, but intact verbal working memory. Taken together, patients P.V. and E.L.D. exhibit a double dissociation between verbal and spatial working memory, as the different lesions affected specific aspects of working memory, and not working memory in general.

Smith and Jonides (1997) have been able to show the same dissociation between verbal and spatial working memory using neuroimaging data. Smith and Jonides (1997) found separable/non-overlapping brain areas that were active for spatial working memory tasks (right anterior and posterior areas) and verbal working memory tasks (left anterior and posterior areas). The fact that the two different working memory tasks were mediated by different, non-overlapping brain regions suggests that working memory is not a unitary system, as Baddeley (1992) theorized.

Taken together, all of these techniques converge on the same theory that spatial and verbal working memory are mediated by two distinguishable systems, one for spatial and one for verbal information. What this example illustrates is a case in which neuroimaging evidence does distinguish

between alternative psychological theories, thereby contradicting the thesis presented by Coltheart (2006, this issue). Of course, one might argue that although neuroimaging evidence has probative value in this case, it is not necessary because other behavioral evidence is sufficient to distinguish between the alternative theories. Put somewhat crassly, why spend all the money it takes to do an imaging experiment when you can get the same information from much less expensive behavioral experiments?

In this case, perhaps that argument has some sway. However, there are other cases where we have shown that alternative psychological theories can be distinguished *only* with neural data and not with behavioral data. Let us examine one such example, taken from work published by Nelson et al. (2003). Nelson and colleagues were interested in interference-resolution: the ability to resolve conflict from irrelevant competing information. These authors sought to identify whether separable forms of conflict exist. For this purpose, they focused on conflict arising from previous information in working memory (proactive interference) *versus* that resulting from previous responses (response-conflict).

To distinguish between these possibilities, Nelson et al. (2003) adapted an item-recognition paradigm (Sternberg, 1966). In this task, subjects are given a target-set of letters to hold in working memory and are later given a probe and are asked to determine whether or not the probe is a member of the target-set. To introduce proactive interference, some probes were members of the target-set of the preceding trial, but not members of the target-set of the current trial, therefore demanding a negative response (recent-negatives). Contrasting responses for recent-negatives with items that had not been members of a recent target-set (non-recent negatives) yielded an interference effect measured in both reaction-time and error-rate. This interference effect was presumably attributable to the familiarity of recent-negatives, making them difficult to reject. To introduce response-conflict, some probes were not only in the previous target-set, but they were also the previous probe-item, therefore demanding a positive response on the previous trial. However, on the current trial, these probes were not members of the current target-set, therefore requiring a negative response on the current trial. Thus, these probes required a response that conflicted with the previous response.

Looking at the behavioral data, interference effects from proactive interference were indistinguishable from those resulting from response-conflict. From these data alone, one would conclude that interference-resolution from working memory and from responses occur via the same processes. However, a look at the neuroimaging data yields a very different picture. Proactive interference yielded activation in the left inferior frontal gyrus, whereas response-conflict yielded activation in the anterior cingulate. The dissociation of these two regions suggests that interference-resolution from conflicting working memory representations is mediated by different neural structures than those resulting from response-conflict. This supports the idea that interference-resolution is not a unitary construct, as the behavioral data might imply, but rather separate forms of resolution may act upon different forms of conflict.

The examples we have discussed thus far make use of the logic of double dissociations. This logic goes as follows. Assume that performance on some Task A is mediated by a psychological process A^I, which in turn is implemented in some brain network A^{II}. Similarly, assume that behavior B is mediated by process B^I and by brain network B^{II}. These assumptions are informatively tested by several experimental programs. In a behavioral task, one might find some variable that modulates performance of behavior A but not behavior B and some other variable that modulates performance of behavior B, but not A. That pattern of results would provide a test of the underlying assumptions. Alternatively, one might find a patient who has damage to brain network A^{II} but not B^{II} and another patient who has damage to B^{II} but not A^{II}. If the former patient's performance on task A suffers but not on task B, and the latter's performance on task B suffers but not task A, this is a test of the underlying assumptions. And finally, one might conduct an experiment in which subjects perform task A and task B, finding that network A^{II} is activated by task A and that network B^{II} is activated by task B. All these sources of evidence have equivalent information value for testing the underlying psychological theory that there are two separable psychological processes underlying the two behaviors in question.

But imaging data have value beyond this as well. One use of imaging data that goes beyond the logic of double dissociations hinges on making use of commonly activated brain regions across multiple tasks to inform theories of the processes underlying these tasks (see also Seron and Fias, 2006, this issue). An example of this comes from recent research in our laboratory having to do with the mechanisms mediated by the left inferior frontal gyrus. In a long line of work using the item-recognition task that we described above, we have found activation in left inferior frontal gyrus due to the resolution of interference on trials in which a negative response is demanded but there is seduction from a high familiarity value to make a positive response. Thompson-Schill et al. (1997) and others have documented activation in a similar region in a very different task, the Verb-generation task. In this task, subjects are presented a series of nouns, and they must respond with an associated verb. Some of the nouns have few verbs that are associated (e.g., 'scissors'), while others have many

(e.g., 'ball'). Thompson-Schill et al. (1997) found that left inferior frontal gyrus is activated as a function of the number of alternative verbs that might be generated (i.e., ball activates it more than scissors). In follow-up work, Nelson et al. (submitted) tested a single group of subjects on both the adapted item-recognition and Verb-generation tasks, and they confirmed a tight overlap in activation between these two tasks in inferior frontal gyrus. Going beyond the obvious value of these data for localization, this result provides support for the theoretical stance that both tasks have an important processing component in common. Indeed, Jonides and Nee (2006) discuss various alternative psychological accounts that might be consistent with these facts, such as a biased-competition model. What is important to note is that the accounts we consider are not neural ones; they are psychological ones.

Indeed, we are not the only authors who have used neuroimaging data to propose new theories of psychological functions. Neuroimaging data interrogating activation of the anterior cingulate cortex has done much to change how we think about interference-resolution. It has long been thought that controlled processing enables goal-directed behaviors to function in the face of competition (e.g., James, 1890). However, many of the theories regarding how controlled processing enables interference-resolution remained incomplete until they were informed by neuroimaging data (e.g., Cohen et al., 1990). Specifically, it was a mystery how increased control was activated in task-appropriate situations. Neuroimaging data revealing activation of the anterior cingulate cortex in many experimental tasks galvanized a solution to this problem.

After observing the activity of the anterior cingulate in situations of conflict, several authors proposed a new theory postulating that the anterior cingulate monitors conflict and calls for increased control in high-conflict contexts (e.g., Botvinick et al., 2001). Since its inception, this theory has received a good deal of behavioral (e.g., Ullsperger et al., 2005) and neural support (e.g., Kerns et al., 2004). Although some authors have found that modifications to the original theory make more accurate predictions (e.g., Brown and Braver, 2005), what is clear is that without neuroimaging data, such a theory would never have been proposed. Indeed, behavioral data alone speak against such a theory since correlations among interference effects in these various tasks are vanishingly low, ostensibly supporting separate mechanisms underlying interference-resolution in different contexts (see, e.g., Kramer et al., 1994). It is the neuroimaging data (of the anterior cingulate among other structures) that lead to the view that there are shared mechanisms (Wager et al., 2005; Fan et al., 2003). Moreover, the imaging data have driven the development of precise computational models of the monitoring, detection, and resolution of conflict (e.g., Botvinick et al., 2001).

Yet another important example showing that neuroimaging data can inform psychological theory by revealing common mechanisms underlying different processes comes from the important work of Stephen Kosslyn and his collaborators (see Kosslyn et al., 2001; Kosslyn and Thompson, 2003; Ganis et al., 2004). In a long series of studies, these investigators have examined the competition between two theories of mental imagery. According to one, mental images have no special status; they are just one example of propositional mental descriptions similar to ones that underlie natural language. The alternative view is that mental imagery requires a different class of processes than those used in language, a class that makes heavy use of the same systems used to perceive in the face of sensory stimulation. The evidence from years of behavioral and neuroimaging research about this problem is outlined by Kosslyn et al. (in press). For example, Ganis et al. (2004) show that some 92% of voxels that are activated during visual perception are also activated during visual mental imagery, thereby demonstrating a very close correspondence between these two systems. These same investigators found, however, that there were some important differences in brain activation between perception and imagery as well, leading to refined hypotheses about what is shared in common and what distinguishes these two cognitive functions. What is critical about this line of research with respect to our present review is that it makes heavy use of data about the localization of processes not for purposes of localization, but to refine psychological theory about mental imagery.

As our title indicates, there are so many examples of the usefulness of imaging data for psychological theory, and so little space to elaborate them. Contrary to the title of a recent book by Uttal (2001), neuroimaging is not the new phrenology. To be sure, it does provide very useful information about localization for clients such as behavioral neurologists who seek that information. But it also provides important information about psychological theory, a source that has every bit the status and probative value of behavioral data from normal and brain-injured humans. It is the co-equal status among these sources of evidence that leads us to the view that a balance is needed in the sorts of evidence one adduces about psychological theories. Neither behavioral nor neuroimaging imperialism is the order of the day. Each source of data has its strengths and its weaknesses (see, e.g., Henson, 2006, this issue), and each can contribute importantly to the development of psychological theory.

REFERENCES

Baddeley A. Working memory. *Science, 255:* 556-559, 1992.
Basso A, Spinnler H, Vallar G and Zanobio ME. Left hemisphere damage and selective impairment of auditory

verbal short-term memory: A case study. *Neuropsychologia, 20:* 263-274, 1982.

BOTVINICK M, BRAVER TS, BARCH DM, CARTER CS and COHEN JD. Conflict monitoring and cognitive control. *Psychological Review, 108:* 625-652, 2001.

BROWN JW and BRAVER TS. Learned predictions of error likelihood in the anterior cingulate cortex. *Science, 307:* 1118-1121, 2005.

COHEN JD, DUNBAR K and MCCLELLAND JL. On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review, 97:* 332-361, 1990.

COLTHEART M. What has functional neuroimaging told us about the mind (so far)? *Cortex, 42:* 323-331, 2006.

DEN HEYER K and BARRETT B. Selective loss of visual and verbal information in STM by means of visual and verbal interpolated tasks. *Psychonomic Science, 25:* 100-102, 1971.

FAN J, FLOMBAUM JI, MCCANDLISS BD, THOMAS KM and POSNER MI. Cognitive and brain consequences of conflict. *Neuroimage, 18:* 42-57, 2003.

GANIS G, THOMPSON WL and KOSSLYN SM. Brain areas underlying visual mental imagery and visual perception: An fMRI study. *Cognitive Brain Research, 20:* 226-241, 2004.

HANLEY JR, YOUNG AW and PEARSON NA. Impairment of visuo-spatial sketch pad. *Quarterly Journal of Experimental Psychology, 43A:* 101-125, 1991.

HENSON R. What has (neuro)psychology told us about the mind (so far)? A reply to Coltheart (2006). *Cortex, 42:* 387-392, 2006.

JAMES W. *The Principles of Psychology* (vol. 1). New York: Dover Press, 1890.

JONIDES J and NEE DE. Brain mechanisms of proactive interference in working memory. *Neuroscience, 139:* 311-316, 2006.

KERNS JG, COHEN JD, MACDONALD AW, CHO RY, STENGER VA and CARTER CS. Anterior cingulate conflict monitoring and adjustments in control. *Science, 303:* 1023-1026, 2004.

KOSSLYN SM, GANIS G and THOMPSON WL. Neural foundations of imagery. *Nature Reviews Neuroscience, 2:* 635-642, 2001.

KOSSLYN SM and THOMPSON WL. When is early visual cortex activated during visual mental imagery? *Psychological Bulletin, 129:* 723-746, 2003.

KOSSLYN SM, THOMPSON WL and GANIS G. *The Case for Mental Imagery*. New York: Oxford University Press, in press.

KRAMER AF, HUMPHREY DG, LARISH JF, LOGAN GD and STRAYER DL. Aging and inhibition: Beyond a unitary view of inhibitory processing in attention. *Psychology and Aging, 9:* 491-512, 1994.

NELSON JK, REUTER-LORENZ PA, PERSSON J, SYLVESTER C-YC and JONIDES J. Interference resolution in the left inferior frontal gyrus. Submitted.

NELSON JK, REUTER-LORENZ PA, SYLVESTER C-YC, JONIDES J and SMITH EE. Dissociable neural mechanisms underlying response-based and familiarity-based conflict in working memory. *Proceedings of the National Academy of Sciences of the USA, 100:* 11171-11175, 2003.

SERON X and FIAS W. How images of the brain can constrain cognitive theory: The case of numerical cognition. *Cortex, 42:* 406-410, 2006.

SMITH EE and JONIDES J. Working memory: A view from neuroimaging. *Cognitive Psychology, 33:* 5-42, 1997.

STERNBERG S. High-speed scanning in human memory. *Science, 153:* 652-654, 1966.

THOMPSON-SCHILL SL, D'ESPOSITO M, AGUIRRE GK and FARAH MJ. Role of the left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proceedings of the National Academy of Sciences of the USA, 94:* 14792-14797, 1997.

ULLSPERGER M, BYLSMA LM and BOTVINICK MM. The conflict-adaptation effect: It's not just priming. *Cognitive, Affective, and Behavioral Neuroscience, 5:* 467-472, 2005.

UTTAL WR, *The New Phrenology: The Limits of Localizing Cognitive Processes in the Brain*. Cambridge, MA: MIT Press, 2001.

WAGER TD, SYLVESTER C-YC, LACEY SC, NEE DE, FRANKLIN M and JONIDES J. Common and unique components in resolving response conflict revealed by fMRI. *Neuroimage, 27:* 323-340, 2005.

John Jonides, Department of Psychology, University of Michigan, 530 Church Street, Ann Arbor, MI 48109-1043, USA. e-mail: jjonides@umich.edu