Interview

# Interview with Robert G. Shulman

■ Robert G. Shulman is the Sterling Professor of Molecular Biophysics and Biochemistry at Yale University. He received his B.A. and Ph.D. degrees at Columbia in physical chemistry, with an intervening period as Lt(jg) USNR. Soon after graduate studies with C. H. Townes in microwave spectroscopy, he went to Bell Telephone Laboratories, where he started Nuclear Magnetic Resonance research on antiferromagnetics; superconductors; semiconductors; and, eventually, biomolecules. A year working with F. H. C. Crick and S. Brenner on phage genetics solidified his interest in biophysics. He started and led biophysics research at Bell Labs, where he studied a variety of biomolecules by NMR. This led him to *in vivo* NMR, at first in microorganisms; then animals; and, finally, humans. Since 1979 he has been at Yale University, where he has been following metabolism *in vivo* by magnetic resonance in brain and muscle measuring changes during activation. He is a member of the National Academy of Sciences and Institute of Medicine. ■

**JOCN:** Cognitive neuroscience has become fascinated with the role brain imaging will play in understanding how the brain enables mind. You have been one of the primary basic scientists in developing this new technique. Recently, you have voiced some concerns about how cognitive psychologists are interpreting their imaging data. Could you tell us a little about this?

**RGS:** I had been studying cerebral metabolism for a long time by *in vivo* NMR, when functional Magnetic Resonance Imaging (fMRI) was developed and used to investigate the visual cortex. We began to study that region by fMRI because we had been measuring changes in metabolites, particularly lactate, by localized NMR during visual stimulation, and the ability to localize the visual response in each individual would improve data interpretation. The sensory responses reported by PET, and being confirmed rapidly by fMRI, were consistent with a long history of sensory studies of the central and peripheral nervous systems. But the ability to study cognitive stimulations informed us about the mind; and therefore the mind, in addition to the brain, was becoming accessible to study.

Our fMRI visual studies gave such strong signals in individuals that I felt confident to test whether fMRI could observe changes in brain activation during cognitive tasks. Initially, I wanted to see how these signals would compare with the pioneering PET reports of brain activations during such tasks as verb generation, various forms of memory, and other forms of cognition. My colleagues and I decided to repeat tasks whose response had been observed by PET measurements. We tried as accurately as possible to use the same tasks. We decided that Verb Generation, as studied by the Washington University group, was well worth replicating. Following this initial study, our cognitive studies have included an expanded study of verb generation, a study of spatial working memory, two studies on "willed action," and a study of the response of the superior temporal lobe to different rates of auditory word presentation.

With the exception of the working memory tasks, these protocols replicated previous PET studies. In these replicated tasks, there was close but not perfect agreement between the fMRI measurements and the PET reports. The data, obtained by different methods, confirmed the general localization of brain activity observed in the PET experiment. However, because of significant differences in data evaluation, we have not been able to agree with the conclusions drawn from the data by previous studies.

I believe that the origin of these differences is the interpretation of data in terms of the cognitive hierarchical paradigms used, which assume the following:

1. differences between cognitive tasks map to unique regions of the brain or activate the same region differentially;
2. tasks can be designed to have unique (nonshared) cognitive processes;
3. the difference images of two tasks (either by fMRI or PET) will reveal the location of the unique (nonshared) cognitive components;
4. the statistical validity of the difference image can be established quantitatively.

I have no difficulties with points 1 and 4. However, in examining these data, I came to question whether the difference image can be identified with the assumed cognitive differences as stated in 2 and 3. In these specific experiments, the interpretations of the difference images in terms of cognitive differences were based upon secondary features of the data with which our fMRI data did not agree. Furthermore, the PET results themselves often did not support the cognitive interpretation, since features of their own measurements had been neglected by the original experimenters.

The softness of these interpretations where the data were forced to fit the starting assumptions led me to question *more generally* the role of the concepts of cognitive neuroscience in functional brain imaging. I believe that the present high respect for cognitive concepts results in distortions of data in order to support the concepts. Cognitive concepts are the basis of the assumptions of unique, nonshared components between two tasks stated in 2 above. When differences are observed in the images, the assumption in point 3 is often considered proven and the nonshared cognitive component is assigned to the difference images. This field would be better served by using the functional imaging data to question these assumptions, rather than by considering both assumptions proven when a difference image is acquired.

**JOCN:** Now that is what I call getting an interview started! You touch on a series of crucial and important points. You seem to be saying that while fMRI studies largely replicate the patterns of cerebral activity noted with PET, these patterns of activation when examined more critically are not really telling us much about how the brain enables cognition. What is it in the analysis of fMRI data that finds you so concerned?

**RGS:** I would like to answer your question in three parts. First is your statement that fMRI studies largely replicate the patterns of cerebral activity noted with PET. That is correct, and it is tremendously important. In the past decade, I have been awed by the novelty of PET research, actually measuring brain regions activated by a variety of tasks. In the PET studies of cognition that have been replicated the fMRI activation, images generally confirm their major findings. There are differences, some of which are significant; but the overall agreement leaves no doubt that the same phenomena are being observed. These accurate localizations of brain activations provide a strong database for seeking the neurophysiological basis of mental processes.

Towards the second point of your question, the consistent, replicated data provide the starting point for a considerable understanding of brain activity. However, in my experience, the concepts of cognitive neuroscience have interfered with reaching this understanding. If suggesting that it is wasteful to force data into a framework of cognitive concepts means that we are not learning about how the brain enables cognition, I would have to answer affirmatively. But I am suggesting we could learn more about cognition by giving these experiments more emphasis than the cognitive concepts have allowed.

This can be shown more directly by using as an example a PET study with which our fMRI data completely agree. However, this agreement, by strengthening belief in the data, suggests a very different explanation of the observation than was offered by the generating cognitive hypothesis.

The PET experiment is the classic "Willed action and the prefrontal cortex in man: a study with PET" by Frith, Friston, Liddle, and Frackowiack (1991). In verbal and sensorimotor tasks, they contrasted novel and routine tasks in which each modality used the same stimulus and response mechanisms. The novel verbal task was verbal fluency, to generate words beginning with a specific letter, while the routine task was to repeat a word. The differences between the tasks were taken to reflect internal *willed action* vs. external response determination. Analogously, the novel sensorimotor task was to move either of two fingers when one was touched while the routine task required that the finger touched be moved. A sketch of the difference maps showed the activation to be quite delocalized in the prefrontal cortex (PFC); however, detailed coordinates for the center of activation are given. In the verbal task, coordinates locate the activation in the left prefrontal cortex and the anterior cingulate, while the sensorimotor coordinates are bilaterally located in the PFC and in the anterior cingulate. Frith et al. (1991) assign a common volume of activation in the dorsolateral PFC (DLPFC) to the common willed action component assumed in the different tasks. They conclude: "Thus an association between response generation (willed action) and activity in DLPFC has been consistently observed in different PET studies."

However, while our fMRI data confirm that both verbal and motor difference images are in the DLPFC, they show unambiguously that the two modalities activate *different* regions of the DLPFC. The difference between motor and verbal localizations is 21 millimeters—which clearly separates the responses to the two modalities. Furthermore, a careful examination of the PET results as reported by Frith et al. show virtually identical results to those we obtained with fMRI. In addition, there is a strong left hemisphere response to the verbal tasks and a bilateral response to motor tasks, also confirmed by fMRI, which distinguish the two modalities! So the fMRI and PET results are just about the same—but we came to different conclusions about the two kinds of tasks.

I suggest that a more accurate description of the experimental results is that there are two different sets of activated regions in the DLPFC that are modality linked. If one follows Frith et al. and lumps these DLPFC activations together, one can claim that there is a common activation element that can be considered to local-

ize willed action. However, the better-localized fMRI results have substantiated the PET localization and have made it reasonable to consider that the data show modality-linked activations of different regions of the DLPFC for verbal and motor tasks. The experimental paradigm designed to localize willed action in the difference images has led to the conclusion of a broadly defined "DLPFC activation" being caused by both modalities. The data, however, distinguish between the localization of verbal and motor activations caused by the increased task complexities, although both are in the DLPFC. One activation could be called *motor willed action* and the other *verbal willed action;* but this alternative raises serious consequences, because it would lead to different subsequent experiments.

The above is not an isolated example. Our other fMRI replications of PET cognitive tasks all raise questions about the ways in which existing cognitive concepts improperly identify the measured difference images with the differences assumed in the hierarchical paradigms.

**JOCN:** But is the basic point that fMRI data allowed for greater specificity in making claims of localization; and, because of that, the modularization of brain processes is even greater than the PET data allowed us to infer? If one technique can only tell us that something is happening in New York City whereas another can tell us that in fact it only happens on 42nd Street, doesn't that spell the end of PET work when used in the context of activation studies?

**RGS:** In both of the willed action tasks, verbal and motor, the fMRI data agreed with the previous PET localization within a few millimeters. Both methods showed well-resolved separations of verbal [in left Brodmann's Area (BA)45] and motor tasks (bilateral in BA 46). The PET data, designed to reveal the effects of willed action, assumed to be identical in both tasks, were described as showing activation of the DLPFC. This interpretation, showing a nominally identical response in accord with the starting hypothesis of an identical cognitive component, disregarded the different locations of activation during the two tasks.

The fMRI results, by reproducing the previous localization, confirmed the reality of the localizations and then strengthened the difference of locations within the DLPFC of the two activations. In this way, the fMRI results made it easier to conclude that the two activations were distinct, but the experimental basis of this conclusion had been clearly presented in the PET study.

To extend your analogy: If one assumed that New York City was a den of sin given to idle pursuits and then a survey found topless bars on 42nd Street and extravagant opera productions one mile uptown, one could conclude that the assumption had been proven. On the other hand, another observer might conclude, without disagreeing with these data, that other, more formative explanations were possible and that the results could support other hypotheses.

All four of our cognitive fMRI studies agree with the previous PET data to a first approximation. However, small differences of findings, complemented by a freedom from cognitive preconceptions, have led us to significantly different interpretations of tasks. In addition to the willed action study, our studies have been of verb generation, working memory, and auditory perception. I would be happy to discuss how the flexibility of cognitive concepts has led to questionable interpretations of these previous results showing that the willed action example is not unique. These results raise, in my mind, two practical questions:

1. How can cognitive concepts help to derive neurophysiological information from functional imaging experiments, both PET and fMRI?
2. What kinds of experiments should be attempted by functional imaging to explore cognition empirically?

**JOCN:** You took the words right out of my mouth. Tell us your view.

**RGS:** How can cognitive concepts help to derive neurophysiological information?

I start with the question asked Michael I. Posner in a recent JOCN interview (1996). "What are the significant questions that attention work has to answer?" His answer started "If you think there are specific brain methods that subserve the operations of attention . . ." Then he listed anatomy, real time, development, and lesions as aspects of attention that should be investigated. I call this the "subservience" theory.

I do not think we are ready to search for brain networks that subserve cognitive concepts such as attention, willed action, automatic word processing, or working memory. Given the primitive state of cognitive concepts and the present strengths of functional imaging, this is the time for revising concepts in response to and in support of the experiments, rather than asking experiments to subserve the concepts. My position is based upon three reasons.

First, the concepts introduce errors in the interpretation of experiments, as illustrated by the experimental examples I have presented to you, where cognitive concepts forced the imaging data to fit preconceptions. Another such example is our published fMRI experiments investigating spatial working memory in which the images in a control task, where attention but not memory was required, were subtracted from the images of the memory task. The difference image showed an activation of the right (PFC). However, when the difficulty of the attention control was increased, its degree of activation also increased, so that the difference signal decreased. Obviously, attention and working memory, as defined by these simple tasks, overlapped in

neuronal activation; and only the small difference signal, smaller than the signals from either task, enabled us to publish this as a localization of working memory. This experiment found a specific brain activation that only within narrow limits could be considered to subserve working memory—at the cost, however, of neglecting the finding that attention activated the same localized region. Once a localized activation is not uniquely identified with one conceptual activity, it is time to revise the concepts. Phenomenological theories should try to discover common features in data. Theories of the neuronal correlates of cognition should look for enrichment by the recent localization results rather than for confirmation.

Second, cognitive theories are not reducible to more established theories, unlike in other life sciences. For example, biochemistry has developed many physically based theories and models to handle its complex data. Biochemical models for enzyme kinetics, energy transduction, or control of fluxes, etc., have been based upon chemistry, with its guiding laws such as quantum mechanics and thermodynamics. Cognitive concepts have no such basis. The absence of strong underlying theories supporting cognitive science is partially responsible for the lack of rigor in interpretation noted above. The arbitrariness of cognitive control experiments illustrates the difference between cognitive theories and physical theories, with their well-established understanding and predictions. If force does not equal mass times acceleration, something unexpected is happening. On the other hand, if attention activates BA46 almost as strongly as memory, or activates it not at all, neither is unexpected, showing the ad hoc nature of cognitive theory.

A third concern with present cognitive theories is their avoidance of consciousness. How can one expect that a cognitive theory or concept be complete when it starts by ignoring the subjective brain activities that fill human lives? As the philosopher John Searle has premised, we are seeking the neurophysiological basis of brain activity. The functioning brain, he continues, is like any organ of the body and must not be broken up arbitrarily on the basis of epistomological accessibility. Subjective responses of volition, emotion, pain, etc., while difficult to quantitate and understand, are caused by brain activity just as are the responses currently postulated in cognitive neuroscience. An opportunity for treating the brain organically is offered by functional imaging, whose response does not reflect the observer's understanding but directly shows brain activity. Hence, this concern about cognitive neuroscience is often linked with the question as to whether there really are neural substrates for cognitive concepts, or are the concepts an arbitrary and incomplete level of abstraction, an interference between the working brain and a neural physiological explanation? My answer is that their usefulness for describing the neurophysiological correlates of cognition has not yet been demonstrated.

In brain studies, scientists often feel the excitement of the future. With this strong sense of the future, of the new perspectives needed, there should be more tentativeness in our espousal of present cognitive hypotheses and more responsiveness of theories to experiment than I find in the field. The times never call for subservient experiments, but they are particularly unwelcome when the theories display so little mastery of the subject.

**JOCN:** If the current theories about attention or memory or whatever are inadequate, there is no hope whatsoever for the problem of consciousness. It is likely the case that little if anything that illuminates the issue from a scientific point of view has ever been written on the topic, especially if you mean the aspect of consciousness that deals with sentience. Most people put the problem of consciousness aside and study things that yield reliable behavioral data. One can easily design experiments that predictably assault working memory capacity by changing the attentional demands of a particular task. To find a fMRI change in the way you described, it would seem, would not be that surprising.

But you are clearly after much more. You appear hopeful that brain imaging allows for some kind of fundamental overhaul of that which will constitute a theory in cognitive neuroscience. Could you expand on the implications of your concerns?

**RGS:** You are correct that I hope functional imaging will provide a fruitful path towards an overhauled theory of cognition. In many experimental imaging studies, such a start has been made—so let me talk about these ways first.

Neurophysiological correlates of location, cellular composition, and temporal response are increasingly well identified in the sensory systems of human and animal brains. A sizable fraction of functional imaging studies are confirming and extending our knowledge of these regions. Recent studies of the visual cortex in a few short years are mapping the extrastriate cortex. Flattening the cortex so as to plot activations, and studies activating extrastriate regions responsive to color, shape, and motion in the human brain, have brought its map up to the gross level of the animal brain. With this base, the fMRI and PET studies have moved from the primary visual cortex through the extrastriate areas to occipital regions that respond to objects rather than textures. Numerous studies of the color center of facial recognition, etc., are striving to localize visual excitation beyond the primary center. These studies are in accord with the well-established approach of following the visual system towards an understanding of higher brain function because so much is already known. I have nothing but praise for these studies. Progress will be served by comparisons between groups so that each experimental ad-

vance will be multiply tested. So too should other sensory cortices be studied and cross-sensory activations evaluated. Existing experimental findings from animals, from neurochemistry, electrophysiology, and other data, all contribute to these directions, and the information obtained will continue moving towards the more complex. How far, how soon, are not known—but the rapid progress generates considerable optimism that suitable theories will be found.

The second kind of experiment that moves us toward a neural physiological correlate takes advantage of the activation during stimulation, independent of the output. This ability to measure brain activation instead of brain output has liberated the functional imaging experiments and allowed them to measure covert, as opposed to overt, tasks. More importantly, it enables brain activation to be measured during undirected activities, from stimuli where the difference image is not anticipated to be correlated with a preordained cognitive activity. In these experiments, we are curious to see the activation and motivated to correlate them with more results, including other subjective experiments. Investigations are being reported with music, emotional stimuli, fear, hallucinations, judgments, recognition, as stimuli. In these cases, unexplained brain activations have been found. At this juncture, the experimentalist should be emboldened not to claim that some existing hypothesis completely explains the novel, complex pattern of brain activation; not to consider the problem solved—by claiming closure of theory and experiment—but to go on to more experiments testing proposed explanations. If brain imaging studies only confirm what we already believe, what, in the end, will have been gained?

A third recommended experimental direction would be to take advantage of theories or hypotheses from fields close to cognitive psychology but not included in the mainstream. For example, neural nets provide a theoretical framework for recognition. When extended to the olefactory bulb, they propose a feedback-forced narrowing of glomerular excitation. Certainly the activation can be measured—can the validity of the model be tested by functional imaging?

From another direction, possibly unattractive to some: can activation during consciousness be studied by comparing brain activity with and without some acknowledged consciousness? Francis Crick suggests comparing the awake and asleep states. My colleague Dr. Morton Reiser, a psychoanalyst, suggests a more subjective comparison, presumably easier to arrange. He refers to moving brain activities into consciousness from the preconscious, essentially a state of putative consciousness. At a time when we are not thinking of a relative, our home, or a familiar piece of music, complex associations can be created in our conscious mind by a name, a word, or a musical phrase. What happens to the brain after that stimulus? Certainly, our conscious state changes. Our experimental signal to noise is approaching the level needed to register the response to such a single association in a single subject. It should be possible to combine this response with an understanding of the subject's description of his state of mind. It won't answer the question, but it would be a step—and the step would be aided or hindered by a different set of preconceptions.

Finally, consciousness is an emergent property, greater than the sum of its neural correlates, but within our present paradigm, to be explained by them. It is clear that recent advances in function brain imaging measure localized aspects of neuronal activation, without requiring reliable behavioral data. Cognitive theories should strive to explain these expanded brain activations rather than to seek confirmation from subservient experiments.

**JOCN:** Cognitive processes map to unique regions of the brain. Your willed action example seems to support this idea very strongly; which examples show it is *not* right?

**RGS:** As you say, during cognitive tasks, unique brain regions are activated. But which cognitive processes are responsible? The willed action results are not a good example of this because, as I've detailed above, they show how searching for the location of a preconceived cognitive activation has led to incomplete and therefore inaccurate attributions of the observed activations. However, even if the two willed action tasks reported activated identical regions of the brain, it would not prove that the experiment had identified the location of willed action. Since the tasks don't give identical results, let's consider one case where an identical activation was obtained during different but related tasks, and ask what kind of planning would help to define the cognitive processes responsible by distinguishing the relevant and nonrelevant components. One identical location, where one might refine the cognitive components of the activation, is right prefrontal BA46 activated in several PET and fMRI studies during spatial attention and "working memory" tasks. To what extent would this activation depend upon input or output modalities? What would be the effect of changing this input to 3-dimensional spatial forms or changing the output from finger-tapping to toe-tapping or to verbal output. Or changing the delay to see if the response weakens (or strengthens) at longer memory times?

These are possible kinds of experiments that might disprove or modify the hypotheses by giving unexpected results. That is the kind of significant advance needed and would not be a "subservient" experiment.

**JOCN:** Tasks designed have unique, nonshared cognitive processes. (Note that this need not be assumed a priority, but it seems clear that many have.) Are you arguing that

many of the tasks carry out the same operations in the same order but have different outcomes? What is the problem here?

**RGS:** I'm arguing that it is wrong to assume that tasks have a "unique," nonshared cognitive process. That assumption relegates the experimental results to the secondary role of localizing the assumed nonshared cognitive process. These are subservient experiments.

**JOCN:** Differences will reveal location of nonshared components. Obviously there can be problems here, but is not the activation of entirely different brain areas relevant to the issue? There is clearly evidence that widely different operations can activate the same general area (e.g., many domains activate the anterior cingulate); but operations can be at many levels, and different domains do interfere, so what evidence leads to rejection of this idea?

**RGS:** First, let us distinguish between activation of different locations and of different intensities at the same location. An fMRI measurement of task response measured with respect to its own pretask baseline helps to distinguish these responses. If activated regions in the two tasks overlap, then the difference map is vulnerable to all the errors generated by differences of large numbers. Although it is not necessary that these differences be small, it has often been true and has caused errors. More importantly, similar locations reveal shared components in tasks assumed to be different, so that intensity differences in the same location are ambiguous. Are the responses similar because of location or different because of intensity? The strength of imaging experiments is the existence of nonshared activations. The PET procedures of generally measuring only differences between tasks and not measuring task-minus-baseline has been responsible for many errors, such as the neglect of and subsequent misinterpretation of Broca's area during verb generation, in the task discussed earlier.

I wholeheartedly agree that the observations of anterior cingulate activations in a variety of tasks is fascinating. Since its prominent activation during the Stroop task, cingulate activation has often shown up in several cognitive tasks, all of which were described as involving decisions. More experiments on this activation should be done—more careful lateral locations, more definite elimination of large vessel inflow—are just some of the experimental possibilities. Cingulate activation stands out as an opportunity for extending our understanding by looking for tasks that do activate it and for related tasks that don't.

**JOCN:** Were you implying that difference images can't be analyzed statistically?

**RGS:** Difference images certainly can be analyzed statistically; and, of course, they are. However, there does not appear to be a uniformly accepted method of statistical analysis that is generally applicable in closed form, although recently Gaussian Random field models offer that promise. Statistical analysis that depends upon signal-to-noise (S/N) is different from a proper spectroscopic analysis, which measures signal and noise separately, or at least measures signal when noise is constant. As my colleague Doug Rothman has pointed out, when comparing one group of subjects, say, men with other groups, i.e., women or children, one must be sure that the noise is identical before relying on statistical analysis of difference maps.

**JOCN:** In addition to hierarchical subtractions, there has now been a number of other methods for dealing with imaging data such as additive factors, correlation between areas . . . which of these methods seem to you to be the most promising?

**RGS:** Depending upon correlations between areas to demonstrate the significance of an activation requires assumptions about the uniqueness of the cognitive difference and assumes a model for its interaction with other areas. If both these assumptions are needed to validate the difference map, the argument becomes circular; i.e., I see this effect because I have assumed it.

**JOCN:** Related tasks performed in the same domain may tend to activate areas close together. Has any common theory of how this happens developed from your work?

**RGS:** This is a wonderful question, but I have no answer.

**JOCN:** Do you feel that the individual subject data you are obtaining support the findings of general areas within which individuals differ, or do individuals tend to violate even the general areas?

**RGS:** This is a wonderful question, but I have no answer.

**JOCN:** Have you developed any general notions of when, for example, in language studies, there will be recruitment of the nondominant hemisphere?

**RGS:** This is a wonderful question, but I have no answer.

**JOCN:** The imaging data say that the same sensory area (e.g., V4 or even V1) can be activated either from input or by a voluntary act of the subject. Do you have any ideas of how these top-down and bottom-up sources can be separated in PET or fMRI data? Do you think that the confusion between top-down and bottom-up sources of activation may account for some of the discrepancies that you are worried about?

**RGS:** PET and fMRI have the ability to map brain activation during either top-down or bottom-up stimuli and as such differ from the many protocols that require objective input and output. While true of sensory stimulation, it has also been observed that cognitive tasks, like verb generation, also activate the same region during top-down (overt) or bottom-up (covert) performances of the same task. This junction, where the same area can be excited either by an input or a voluntary act of the subject, is where cognitive tasks can meet the inaccessible regions of mind, where the unreachable, ineluctably subjective can replicate, add, or subtract from input activations. This meeting place is the window to the mind and represents a hope, offered by imaging methods, of healing Descartes' dualism.

**JOCN:** Thank you.

## REFERENCES


Frith, C. D., Friston, K., Liddle, P. F., & Frackowiack, R. S. J. (1991). Willed action and the prefrontal cortex in man: a study with PET. *Proceedings of the Royal Society of London B, 244,* 241-246.

Gazzaniga, M. S. (1996). Interview with Michael I. Posner. *Journal of Cognitive Neuroscience, 7,* 83-87.

Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M., & Raichle, M. E. (1989). *Journal of Cognitive Neuroscience, 1,* 153-170.

Petersen, S. E., Fox, P. T., Snyder, A. Z., & Raichle, M. E. (1990). *Science 249,* 1041-1044.