Rajeev Raizada: Statement of research interests

Overall goal: explore how the structure of neural representations gives rise to behavioural abilities and disabilities

My research seeks to uncover how the human brain's neural representations underlie task performance: how suitably-structured representations allow a task to be performed well, and how poorly-structured representations may lead to impaired performance or to learning disabilities. To work towards this goal, I use fMRI, behavioural psychophysics, and computational analysis. The computational component, grounded in my Ph.D. training in neural networks (Grossberg & Raizada, 2000; Raizada & Grossberg, 2001, 2003), allows me to explore aspects of fMRI data which traditional fMRI analyses are unable to deal with: distinguishing between neural representations which are spatially distributed and overlapping, analysing the similarity-structure of those representations, and relating that structure to behavioural performance.

The caricature of fMRI research, not always unjustified, is that it adds nothing to what we can discover from purely behavioural methods, except for a bunch of pretty pictures showing which parts of the brain lit up. This view, although overly harsh, does raise an important question which neuroimaging studies can address: what can you learn from looking inside the head that can't be learned from outside the head?

A key answer to that question, I believe, is the following: from the outside, using behavioural measures alone, only *performance* can be measured. Looking inside, using neuroimaging, offers the possibility of probing representational *competence*. For example, suppose there are two children who both score equally badly on a school test. From the outside, they look the same: they both get low scores. However, on the inside, one child may have neural representations that are poorly suited for being able to perform the task, whereas the other child may have a perfectly good set of neural representations, but instead may have problems of attention or motivation. The types of help that these two children would benefit from are quite different.

My most recent research provides a potential means for fMRI to distinguish between such cases, and is the first work to relate the structure of people's neural representations to their ability to perform a specific task (Raizada et al., under review). Leading up to this, my work has consistently sought to explore the underlying mechanisms and representations which link brain to behaviour, exploiting methodological advances to address cognitive questions which would otherwise be inaccessible.

Before pattern-based analysis approaches were introduced (by Haxby and colleagues), it was nonetheless possible to use fMRI to study distinct but spatially intermingled neural representations, using the method of adaptation-fMRI (introduced by Grill-Spector and Malach). Both methods were introduced in the field of visual object recognition, and were often couched in terms which made their applicability seem specific to that domain (e.g., using adaptation-fMRI to study viewpoint- and size-invariance in visual object processing). However, building upon the underlying logic of these approaches, I have extended them to address new questions in a broader set of domains. In Raizada & Poldrack (2007b), I conducted the first study to apply adaptation-fMRI to the topic of phonetic perception, based on the reasoning that the neural populations which respond to different phonetic categories are likely to be shared and overlapping, but should be experimentally separable based on their different responses to within- and across-category stimulus pairs. This study was also the first to use fMRI to probe the structure of phonetic categories by deriving and comparing neurometric curves and psychometric curves.

Pattern-based analyses, developed a few years after adaptation-fMRI, are more powerful, as they can examine multivoxel distributed patterns of activation whereas adaptation-fMRI in its standard form still looks at activation one voxel at a time. The broader questions motivating Raizada & Poldrack (2007b) also provided part of the impetus for Raizada et al. (under review). In addition to using more powerful methods, the more recent paper is also able to address the key question of how neural activity relates to behavioural performance, in virtue of studying subjects with widely varying abilities at performing a single fixed task.

A different study of mine also sought to examine parametric connections between brain and behaviour, but this time by looking at fluctuations in performance at at the trial-by-trial level (Raizada & Poldrack, 2007a). This paper also made a conceptual contribution, pointing out how standard subdivisions of attention (goal-driven, stimulus-driven etc.) fail to capture the dimension of how much or how little of the brain's cognitive resources are allocated in response to an attentional challenge. Pursuing further the question of how fMRI can uncover internal aspects of competence that are not revealed behaviourally measured performance, and building upon my interest in the development of language skills, I performed a study of 5-year-old children (Raizada et al., 2008), and found that the hemispheric asymmetry of activation in Broca's area could serve as a marker of language development which remained informative even after behavioural measures of language ability were partialled out.

In order to understand how the brain gives rise to behaviour, images alone are not enough: the underlying mechanisms and representations must be explored. The studies described above are united in the pursuit of that goal. However, standard fMRI analysis approaches are limited in their ability to access some key representational questions, and my most recent work explores new methods, in order to move beyond those limitations. In recent years, interest in these methods and the new questions which they open up to study has steadily grown: at the 2008 Cognitive Neuroscience Society Meeting, I chaired and presented at a symposium on this topic, entitled "Pattern-based fMRI as a route to revealing neural representations".

I now describe how I have used novel pattern-based analysis methods to cast new light on how the structure of neural representations can either help or hinder behaviour, and how they differ from standard approaches.

How new analysis methods open up previously inaccessible questions about cognition and behaviour

The brain has rarely been obliging enough to make its neural representations discrete and spatially compartmentalised; instead they are often distributed and overlapping. This makes them inaccessible to standard methods of fMRI analysis, as is schematically illustrated in Fig. 1A. In Figure 1B, the novel hypothesis proposed and tested in Raizada et al. (under review) is illustrated: the more separable the neural patterns elicited by two stimuli are in a person's brain, the better that person should be behaviorally at telling those two stimuli apart. I tested this hypothesis using a cross-linguistic perception task: English and Japanese native speakers listening to the phonemes /r/ and /l/.

In a standard voxel-by-voxel analysis, the only question that can be asked is which voxels have greater signal intensity. When looking at multi-voxel patterns, one can ask a more general question: to what degree are the spatial patterns of fMRI activation statistically separable? In native English speakers, who have no difficulty hearing the difference between /r/ and /l/, the neural patterns should be highly separable, but in Japanese speakers the patterns should be less separable, correspondingly the fact that they find that phonetic contrast much harder to perceive. This is indeed what we found. Crucially, the neural pattern-separability predicted not only group-differences (English vs. Japanese), but also individual differences in perceptual discrimination ability. The more separable a persons neural representations for /r/ and /l/ were, the better they were at hearing the difference between the two sounds. This is the first study to have shown a relationship between multivoxel patterns of fMRI activity and people's behavioural performance.

The fundamental question explored by my /r/-/l/ study is this: what makes a representation suitable for performing a given task? Although this is clearly a significant question, it has remained quite underexplored. A possible reason for this is that many tasks show little variability across individuals: apart from lesion patients, everybody can identify faces, perceive the orientation of gratings and recognise visual object categories without too much difficulty. Cross-linguistic studies such as this one of English and Japanese speakers are scientifically useful exceptions, but more general and also more pressing are learning disabilities, such as dyslexia and dyscalculia.

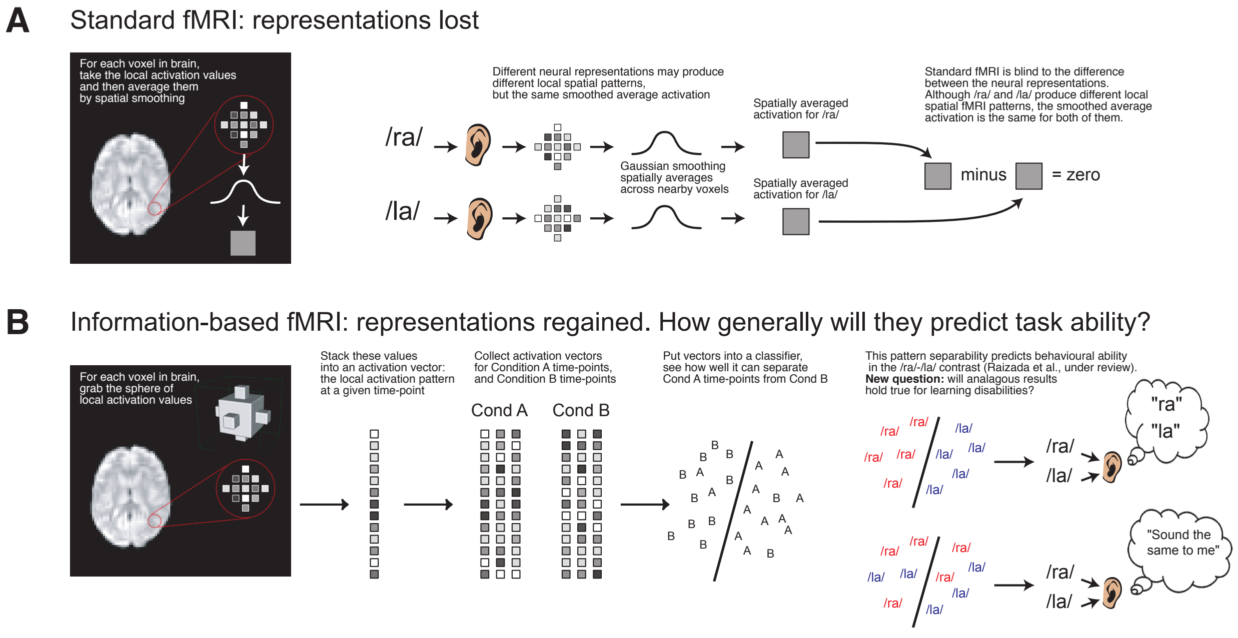

Figure 1: *How my work differs from more traditional fMRI approaches.*

Medium-term research plans: neural representations in dyslexia and dyscalculia

In dyslexia and dyscalculia, children who are otherwise cognitively normal are somehow unable to do tasks which most children can perform well. The study of /r/ and /l/ sketched out above suggests a simple and directly testable hypothesis, namely that the neural representations needed for learning specific linguistic or numerical skills are poorly structured in these children. This raises the question: what are the subtasks which are fundamental for learning reading and arithmetic, and what would poorly-structured and well-structured representations for those tasks look like? I very recently submitted an NICHD R21 grant proposal addressing that topic.

For a child to learn reading and arithmetic, there are a great many tasks and subtasks involved. My recently submitted grant proposal focuses on two which are fundamental: (i) categorical perception of speech sounds, which has been shown to be closely related to the key pre-reading skill of phonological awareness, and (ii) mental representation of the "number line" of integers, which is predictive of success in learning basic arithmetic.

Consider first the task of discriminating speech sounds. For a child to be able to hear the difference between two phonemes, the neural patterns evoked by those sounds in the child's brain must be distinct. On this hypothesis, the less distinct the neural patterns are, the worse the behavioural performance will be. This clearly parallels the hypothesis tested and verified in Raizada et al. (under review) about the representational differences underlying Japanese and English speakers' ability to perceive the /r/-/l/ distinction, and it can be tested the same way.

As well as looking at how distinct different patterns of fMRI activity are, it is also possible to look at how the similarity and dissimilarity of a set of neural patterns is structured. For the task of learning basic arithmetic, a crucial internal representation for a child to possess is that of a "mental number line". My grant submission proposes to test the hypothesis that in control children who do not have math difficulties, the neural representations of numbers of differing magnitudes will reflect the structure of a number line, with numbers which are closer together or further apart eliciting neural responses which are more or less similar, respectively. Conversely, in children with dyscalculia we hypothesise that this representation will be disordered, with a disrupted mapping between numerical and neural similarity.

A neural representational structure which separates small numbers medium from large would be very helpful for performing number-related tasks. An example of the similarity space of such a set of representations is shown in Fig. 2A. If the fMRI signal-to-noise ratio within and across subjects is good enough, then it may even be possible to see meaningful structure in a non-lumped representation, looking at individual numbers (Fig. 2C). A neural representational structure which mixes numbers of all different size together would, in contrast, be a hindrance. Examples of such structures at coarse and fine grains are shown in Figs. 2B and D.

Thus, a hypothesis about the neural representations underlying dyscalculia can be framed in almost exactly the same terms, and tested using almost the same methods, as the hypothesis stated for dyslexia above: the similarity space of numerical representations in normal children will reflect the actual magnitudes of the numbers themselves, whereas in dyscalculic children the neural similarity space will not. Moreover, a more specific hypothesis can also be tested: the degree to which the neural similarity space fails to reflect the actual magnitudes of the numbers will be predictive of the degree of behavioural impairment in the dyscalculic subjects.

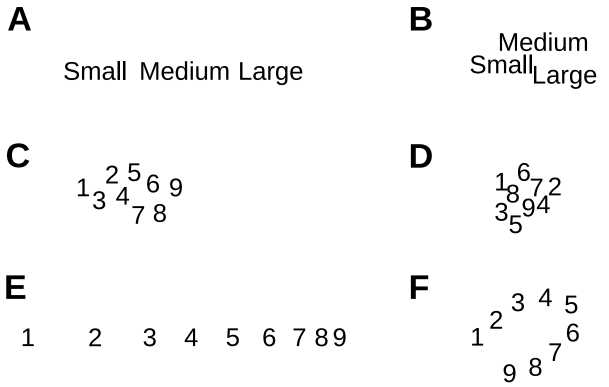

Figure 2: *Hypothetical sketches of possible neural similarity spaces for representing numerical magnitudes. Spaces B and D would be poorly suited for supporting numerical and arithmetical tasks, and are hypothesised to be the kinds of representational structures that will be found in the brains of dyscalculic children, most likely in parietal cortex. Note that these figures depict an abstract similarity space, not physical space in the brain.*

Longer-term research plans: computational tools, cognitive questions and applied goals

Clearly it will take many more experiments to do justice to the question of how neuroscience can improve education, and a research program addressing that has little danger of running out of problems. What is likely, however, is that we will encounter problems which our existing methods are unable to address. I believe that my background in computational modeling provides a very useful foundation for exploring new approaches for finding information in fMRI data. New pattern-based analysis approaches are no mere mathematical curiosities, they are vital for addressing longstanding questions about how the brain gives rise to behaviour, questions which traditional analysis methods have been unable to touch. However, there are no widely agreed-upon or standardised methods in this new and often uncertain area of research. The best-grounded and most effective sets of methods are yet to be discovered, and the ability to exploit and adapt findings from statistical pattern recognition and machine-learning will be essential for such work.

A key question of this sort is how best to extract the information which is present in distributed patterns fMRI activation. Should the activation from a whole brain full of voxels be considered all at once, or should local spatial neighbourhoods of voxels be analysed separately, or should some other criterion for chunking together voxels be used? This is the long-standing question of feature-selection, which has received much attention in the machine-learning community. Currently used methods in fMRI, such as "recursive feature elimination" have intuitive and practical appeal, but do not fully exploit the rigorous theoretical advances which machine-learning researchers have already made. One key problem is the fact that selecting from the huge number of possible voxel-combinations can become computationally intractable. I am currently investigating approaches derived from information-theory, such as the "minimum Redundancy Maximum Relevance Feature Selection" (mRMR) method developed by Peng and colleagues, which are able to select informative subsets of voxels at low computational cost.

Although these computational and methodological questions are important in their own right, my interest in them is rooted in the opportunities that they offer for addressing fundamental cognitive questions. Specifically, they allow new lines of attack on longstanding questions about how neural representations are structured, and how those representations enable behavioural performance.

One avenue which I am particularly interested in exploring is using training-based approaches to enhance learning. Such an approach has several potential advantages, but is not without its difficulties. One advantage is that intervention is the ideal causal probe: the best way to test a hypothesised causal link between A and B is to kick A and see if B moves. If process B is one that benefits learning or education, then the experiment not only probes a basic science question but also may have practical utility.

Because the work described above provides specific hypotheses about how a given representational structure either helps or hinders task performance, it also makes testable predictions about how training-induced improvements in performance should be underpinned by a restructuring of the underlying representations. For example, consider a remediation for dyscalculia which seeks to improve children's representation of a "mental number line". Pre-versus-post improvements in performance may or may not be a result of better-structured neural representations of number. Using the approaches described here, that structure can be measured by neuroimaging, probing representational competence directly.

Our field too often tends to split between the methodological, the cognitive and the applied: people who work on statistical algorithms do not always ask how their methods can be used to reveal cognitive and neural mechanisms, and researchers into learning and cognition do not always ask how their insights might be able to impact upon those whose learning and cognition is impaired. My work unites all these areas, giving me a powerful "scientific toolkit" for continuing to develop a research program which confronts difficult theoretical issues, while also addressing problems that impact upon the real world. Building such a research program is long, hard and often challenging work; I can't think of anything that I'd rather do.

References

Grossberg, S. & Raizada, R. D. S. (2000). Contrast-sensitive perceptual grouping and object-based attention in the laminar circuits of primary visual cortex. *Vision Research*, *40*(10-12), 1413–1432.

Raizada, R. D. S. & Grossberg, S. (2001). Context-sensitive binding by the laminar circuits of V1 and V2: A unified model of perceptual grouping, attention, and orientation contrast. *Visual Cognition*, *8*(3-5), 431–466.

Raizada, R. D. S. & Grossberg, S. (2003). Towards a theory of the laminar architecture of cerebral cortex: computational clues from the visual system. *Cerebral Cortex*, *13*(1), 100–113.

Raizada, R. D. S. & Poldrack, R. A. (2007a). Challenge-driven attention: interacting frontal and brainstem systems. *Frontiers in Human Neuroscience*, *1*, 3.

Raizada, R. D. S. & Poldrack, R. A. (2007b). Selective amplification of stimulus differences during categorical processing of speech. *Neuron*, *56*(4), 726–740.

Raizada, R. D. S., Richards, T. L., Meltzoff, A., & Kuhl, P. K. (2008). Socioeconomic status predicts hemispheric specialisation of the left inferior frontal gyrus in young children. *Neuroimage*, *40*(3), 1392–1401.

Raizada, R. D. S., Tsao, F. M., Liu, H. M., & Kuhl, P. K. (under review). Quantifying the adequacy of neural representations for a cross-language phonetic discrimination task: prediction of individual differences. Under review at Cerebral Cortex, at secondary stage of revisions. Manuscript number CerCor-2008-00712.