

Rajeev Raizada: Statement of research interests

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

There tends to be a split in the field of fMRI: some researchers focus primarily on asking cognitive and neural questions, whereas others concentrate more on developing methods. An unfortunate side-effect of this split is that research on cognitive questions may fail to use the methods that could most powerfully address them, and the developers of new methods may fail to use them to help us better understand the brain.

My own research attempts to span this divide. I seek 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 have developed and am continuing to explore new computational methods for analysing the multivariate spatial patterns in fMRI data, and for relating these patterns to individual differences in people's behaviour.

The field of pattern-based fMRI analysis (also sometimes referred to as MVPA) is currently growing at an explosive rate (Raizada & Kriegeskorte, 2010). Three aspects of my own work set it apart. The most central of these is the goal of relating fMRI patterns to behaviour. Most research in this area currently focuses on "decoding," namely reasoning backwards from observed patterns of fMRI activation to infer which experimental condition gave rise to them. In contrast, my work goes beyond that by asking whether we can "reason forwards" from distributed brain activation to make inferences about people's behaviour: Raizada et al. (2009) was the first paper to show a relationship between the pattern-separability of people's neural representations and individual differences in their behavioural ability. The second distinctive aspect is that my computational training in neural networks (Grossberg & Raizada, 2000; Raizada & Grossberg, 2001, 2003) has equipped me with the tools to generate new analysis approaches, rather than being forced to use whatever off-the-shelf tools might be available. Some examples of my new approaches, described in more detail below, are studies linking brain-wide pattern separability to people's behavioural performance (Raizada et al., 2010) and developing improved methods for finding distributed neural representations at the multi-subject group-level (Raizada & Lee, 2010a, 2010b). The third differentiating factor in my work is its bigger picture goal of relating these more theoretical questions to real-world problems, specifically the question of how Cognitive Neuroscience can help to diagnose and remediate learning difficulties and educational disadvantage (Raizada et al., 2008; Raizada & Kishiyama, 2010; Raizada et al., in preparation).

A longstanding goal of my research has been to try to move beyond fMRI studies which simply show "what lit up," and instead to investigate how specific information processing operations take place in the brain's distributed networks. In Raizada & Poldrack (2007b), the method of adaptation-fMRI was used to study how the overlapping neural representations of different phonemes are contrasted against each other during the categorical perception of speech. This led directly to exploration of new pattern-based fMRI analysis methods, which use tools from machine learning to access similar questions but in a more powerful way. While at Dartmouth, I have co-supervised a Ph.D. student on a research project building upon that categorical perception work, resulting in a manuscript recently submitted for publication (Lee et al., 2010). 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 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.

Distinguishing behavioural performance from neural representational competence at the level of individuals

Relating the structure of people's neural activation patterns to their behavioural ability offers a route to go beyond behavioural measures of *performance* by probing underlying representational *competence*. An example to illustrate this distinction is as follows: 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 recent research provides a potential means for fMRI to distinguish between such cases: in Raizada et al. (2009), I proposed and tested the hypothesis that the more statistically separable the spatial fMRI patterns elicited by different stimuli are, the more perceptually distinguishable those stimuli should be. This was tested by looking at the perception of the /r/-/l/ contrast by English and Japanese speakers. In native English speakers, who have no difficulty hearing the difference between /r/ and /l/, the elicited 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.

That paper found its brain-behaviour correlation within a specific Region-of-Interest (ROI), which was derived from a group-level analysis. However, in new work, I have found that an equally strong link between pattern separability and behavioural performance also holds true when the entire brain is considered at once, without any ROI or voxel-selection at all (Raizada et al., 2010). The analysis uses a linear classifier known as a Pseudo Fisher Linear Discriminant, applied to the whole brain at once. Although specified by a very simple equation, this type of classifier has been shown in the machine-learning literature to perform well when large numbers of redundant feature dimensions are present. This is precisely the situation in a whole-brain analysis.

The fact that this correlation emerges from a whole-brain test, rather than from having to restrict the analysis to a carefully selected ROI, offers a possible solution to a difficulty which has recently received much attention, namely the problem of selection-bias. If the criterion used to define a given ROI is the same as the one being tested for in the data extracted from that ROI, then the resulting correlation test may be biased or even wholly invalid (Kriegeskorte et al., 2009; Vul et al, 2009). A direct way to avoid any risk of selection bias is simply not to do any voxel or ROI selection. The whole brain analysis does just that.

As part of my long-term goal of seeking to apply Cognitive Neuroscience towards improving edu-

cation, I have recently been testing these ideas in the domain of numerical cognition. The first study to date relates fMRI activation on a non-symbolic numerical “distance effect” task to behavioural scores on standardised tests of arithmetic and language. In a distance-effect task, the subject judges which one of a pair of numbers presented together is larger; the closer together in magnitude the two numbers are, the longer the subject’s reaction time. This task has been found in standard fMRI analyses to produce activation differences in parietal cortex. However, without needing to select any ROI, we found a whole-brain correlation with behavioural performance in this distance-effect data: the separability between neural patterns elicited by large-distance pairs and small-distance pairs was positively correlated with subjects’ scores on the arithmetic tests. Moreover, the correlation with behaviour was specific: fMRI pattern separability did not correlate with language test scores at all (Raizada et al., 2010).

Long-term research vision: key neural questions, and the new tools needed to tackle them

Some of the key problems in Cognitive Neuroscience that I would like to address over the coming years are the following:

- Finding sufficiently strong brain-behaviour links, such that fMRI gains genuine diagnostic power, beyond what behavioral tests on their own can provide
- Identifying neural indicators of developmental cognitive problems, before they manifest themselves in behaviour later in childhood
- Discovering direct mappings between the structure of complex cognitive tasks and the neural representational spaces which underlie those tasks’ performance

In order to tackle these problems, new tools will be needed. Moreover, we must be able to validate those new tools, so that we avoid trying to probe the unknown with the unknown. Developing new approaches to extract information from the brain is an exciting pursuit: we can invent new analysis methods, code them up, and immediately try them out on real neural and behavioural data. However, the profusion of new analyses that can be devised raises a problem: how can we know whether our new analyses truly tell us something about the brain’s cognitive and neural processing, or, more worryingly, whether the results that emerge might merely be the consequences of arbitrary aspects of the analysis methods that we happened to use?

In some of my recent work, I have tackled this problem by using categorical speech perception as a testbed. In a previous study (Raizada & Poldrack, 2007b), I used adaptation-fMRI and a set of stimuli along the /ba/-/da/ continuum to tease-apart the neural representations of phonetic categories which are intermingled at the subvoxel level. Investigating the same phenomenon from a different angle, my student and I more recently used the same stimuli but a different task design, seeking to explore multivoxel distributed representations of those phonetic categories. This approach has two advantages. First, it allowed us to compare different spatial scales of neural representation of the same perceptual phenomena (Lee et al., 2010). Second, these two independent data sets allow candidate new analyses to be subjected to a strong validation test: replicability.

Using this testbed, I have developed a new approach to performing “searchlight” multivoxel pattern analyses, and have shown that using this new approach yields results which replicate very closely across the two independent datasets (Raizada & Lee, 2010a, 2010b). This replication constitutes a strong validation of the new methods, thereby providing a sound starting point for investigating new cognitive and neural processes whose workings are not yet known.

Building up the diagnostic and predictive power of fMRI, and seeking strong mappings between the structure of neural representations and of behaviour, are wide-ranging goals. However, these goals must be tackled in specific domains. Above, I described some studies which have laid the initial groundwork for addressing these problems, in the areas of speech perception and numerical processing. Those studies, however, have involved relatively low-level tasks and stimuli, and have so far been either of adults or of unimpaired children. Over the next five years, a central goal of mine will be to test the hypothesis that measures of neural representational competence can be used to predict behavioural performance not only in normal adults, but across the entire spectrum from the unimpaired to the learning disabled, and throughout the course of development from early childhood to adulthood, and even into old age.

Moreover, I plan to seek neural representational structures underlying not only lower-level tasks but also more complex cognitive phenomena. Two mid-level domains that I believe to be particularly ripe for this sort of approach are the neural similarity spaces underlying the phonological awareness skills that enable learning to read, and the inter-relations between the neural representations of performing basic addition, subtraction, multiplication and division. In a child who has mastered these skills, the neural representational structure should mirror the relations between the tasks themselves. In a child who is struggling to learn these skills, neuroimaging of the child's representational structure could potentially help us to find where the learning is going wrong, and to remediate the learning accordingly.

Our field too often tends to split between the methodological, the cognitive and the applied. My work unites all these areas, and generates a set of hypotheses and lines of attack for 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.
- Lee, Y. S., Granger, R. H., & Raizada, R. D. S. (2010). Different spatial scales of categorical processing in Broca's area and the supramarginal gyrus: multivoxel analysis and adaptation-fMRI compared. *Being revised for resubmission*.
- 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. & Kishiyama, M. (2010). Effects of socioeconomic status on brain development, and how cognitive neuroscience may contribute to leveling the playing field. *Frontiers in Human Neuroscience*, 4, 3.
- Raizada, R. D. S. & Kriegeskorte, N. (2010). Pattern-information fMRI: new questions which it opens up, and challenges which face it. *International Journal of Imaging Systems and Technology*, 20, 31–41.

- Raizada, R. D. S., Lebedeva, G., Bhagat, S., & Kuhl, P. K. (in preparation). Differences in parental language input predict broca's structure in five-year-olds.
- Raizada, R. D. S. & Lee, Y. S. (2010a). Smoothness without smoothing: why Naive Bayes is not naive for multi-subject searchlight studies. *Under review*.
- Raizada, R. D. S. & Lee, Y. S. (2010b). A proposed method to improve the spatial specificity of multi-subject "searchlight" studies, with validation via a multi-site replication. *Under review*.
- Raizada, R. D. S. & Poldrack, R. A. (2007a). Challenge-driven attention: interacting frontal and brain-stem 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., Holloway, I. D., Ansari, D., & Kuhl, P. K. (2010). Linking brain-wide multivoxel activation patterns to behaviour: Examples from language and math. *Neuroimage*, 51(1), 462–71.
- Raizada, R. D. S., Tsao, F.-M., Liu, H.-M., & Kuhl, P. K. (2009). Quantifying the adequacy of neural representations for a cross-language phonetic discrimination task: Prediction of individual differences. *Cerebral Cortex*, 20(1), 1–12. Advance Online Publication: April 22, 2009.