Functional MRI: A Phrenology for the 1990’s?

Mark S. Cohen, Ph.D.
Director of Magnetic Resonance Imaging
UCLA Brain Mapping Division
710 Westwood Plaza
Los Angeles, CA 90095

“...In single individuals, fMRI (functional MRI) responses have been reported to visual stimuli [1-7], somatosensory/motor activity [5, 8-10], and acoustic stimuli [11].”

“...blood flow changes were reported using non-tomographic techniques more than fifteen years ago [12] and through the use of PET [13], but with fMRI it becomes possible to interrogate the locus of such activity on single subjects with a high degree of reliability [14].”

These are my own words, and though I would stand by them, I may well live to regret writing them. Functional MRI is touted widely, by no means solely by this author, as a method to detect regional brain activation patterns on individuals, and it reasonable to expect that enthusiastic readers might take this rather literally. In fact, however, it is only in the rare case that the validity of single subject work has been tested critically (e.g., in presurgical mapping studies, or multiple subject studies with identical outcomes). Taken at face value, the claim for single subject sensitivity, the failure to replicate across individuals, or even across trials for single subjects, have all been interpreted in print as demonstrating “differences in processing strategy” or “learning effects.”

Consider a recently presented paper asking, “are the names of living and extinct organisms represented at different loci in the human brain?” (To avoid criticizing the work of an individual author, the experimental details have been altered...). While being imaged in an MR unit, the subjects are presented a series of pictures of obviously living (condition A) and obviously extinct (condition B) animals and asked silently to name them. A set of difference images between conditions A and B is then calculated for each subject. The pictures show a variety of locations that have larger signal in one condition than the other over three repeats of each condition. Our authors state that this is effectively conclusive proof that, in fact, the cerebral representation of these life forms is different. Needless to say, we are then presented with an exhaustive post-hoc review of the linguistic theories that might be cited in support of (or at least consistent with) this result.

The data themselves are quite interesting showing widely different areas of “activation” in each subject, demonstrating, “differences in cognitive processing strategy” and a remarkably uniform distribution of active regions within individuals (perhaps a slight predominance of regions on the cortical surface) including both white matter and ventricular areas, and few common areas across individuals. The authors do not find such a result silly, and have been trying for some time to get these data into print, as I am certain they eventually will. The fact that it makes effectively no sense in the context of one hundred years of neuroscience seems to be irrelevant. Just what is our failure here?
One of the key issues here is a confounding of the restricted use of the term “significance” in the statistical sense, with its colloquial meaning, or with truth. The raw statistic (t values, signal intensity differences, \(\chi^2\) or F ratio is a matter of simple calculation. The probability level calculated from these numbers, however, is based on a series of critical assumptions, most of which are at best poorly understood in brain imaging studies. What, for example, is the magnitude of the effect of motion on these statistics? Is there a closed form solution for the problem of repeated semi-independent measures (e.g., the individual voxels in the image set, or the repeated, yet autocorrelated, sequential images in an MR acquisition? A more important underlying problem is the lack of critical data exclusion criteria, based on an informed understanding of plausible neuroscience. Perhaps the most vexing problem is one of cognitive dissonance, “Why would I have spent so much money on my MR scanner if the data I collect are invalid?”

Whatever the resolution to these issues and controversies, our willingness to come to terms with them will forever affect the perception of the field of fMRI and other activation imaging. A failure to confront bad science, either by our own serious self-criticism, or by more rigid reviews of the proposed literature, may result in the discrediting of an entire new form of brain mapping.

We must therefore, as a team, work to validate the range over which any brain mapping method can be used in single subject work; we must be circumspect in our interpretations of the data in the context of the validations that exist already, particularly when the potential clinical and research value of such a single subject approach is so overwhelming. Perhaps in this way we can prevent fMRI from becoming the phrenology of the late 20th century.
REFERENCES:


