Functional brain imaging

Fishy Findings

By the time the eccentric ‘dead salmon’ experiment began spurring a lot of doubt, a few social neuroscientists had already started probing into the reliability of the interpretations of modern day Magnetic Resonance Imaging. Subsequent investigations revealed obscure pitfalls that blew up neural correlations to behaviour...

‘Gambler’s ruin’ is a famous example of a stochastic process, where, in theory, if you are playing a game in which the odds are in your favour and you play long enough, you will come out a winner. But in practice, you have a finite stake and you retire quite early in your game as you run out of money. You have lost a game that was actually in your favour.

MR Imaging and its different flavours

Probability distributions surface in all aspects of life from bettings to stocks and even scientific derivations. Any observation, unless appropriately compensated for randomness, could lead to misinterpretations that can, in turn, set the stage for distorted theories. The notorious ‘dead salmon’ experiment lured a lot of attention among the social neuroscientists in the last decade and provoked a revisit to the fundamentals of statistical analyses in neuroimaging. The experiment did not have a logical design and was not intended to answer any major scientific question for the betterment of mankind but was performed by a curious student with a funny bone. However, what he observed serendipitously, turned out to impugn the reliability of one of the most powerful diagnostic tools in medicine – the functional magnetic resonance imaging (fMRI). Contemporary studies that sought to address the flaws associated with the conventional statistical methods spurred the need for more rigorous corrections and compensations in imaging analyses. But before we go fishing, let’s have a look at the whole technique in more detail.

Water is vital not only for all life on earth, two-thirds of our bodies consist of H_2O, it is essential for magnetic resonance imaging, too. This holds especially true for the proton particles located in the nucleus of the hydrogen atoms of the water molecule. These protons feature certain magnetic properties and so, when an MRI machine produces a magnetic field, the magnetic moments of these protons align with the direction of that field. Then a radio frequency transmitter comes into play and generates an electromagnetic field, through which the alignment of the protons is systematically altered. The protons thus, switch from a low-energy spin-down state to a high energy state and go back to the former when the magnetic field is turned off. What you get as a result is a release of energy, in the form of photons. These are detected by the MRI scanner as an electromagnetic signal. Protons in different tissues return to their equilibrium states at different rates, which is unique to the tissue itself. This difference can be used to construct images of the tissue. Tissue contrast is provided by distinct tissue variables that depend on local water and fat content.

Of peppers and clams

The imaging world saw a breakthrough soon after the discovery of MRI in the 1970s as the technique had more to offer than other contemporary scanning tools viz. computerized tomography (CT) or positron emission tomography (PET). The main highlights of MRI are its high spatial and contrast resolution and the fact that the technique is non-invasive and works *sans* radiation hazards. Thus, it did not come as a surprise, when Paul Lauterbur of the University of Illinois and Sir Peter Mansfield of the University of Nottingham shared the Nobel Prize in Physiology or Medicine in 2003 for “discoveries concerning magnetic resonance imaging”. Interesting to note also that the first ever MRI images taken by Lauterbur were those of green peppers and clams (*The Economist,* 4/07).

fMRI bolsters cognitive psychology

Tweaking a bit on the principles of MRI, researchers soon came up with a more sophisticated technique, the fMRI. Functional imaging primarily serves to image brain activity; this happens through evaluating the differential magnetic properties of oxygenated (diamagnetic) and deoxygenated (paramagnetic) haemoglobin. Neuronal activity is bound to increase blood flow to ‘active zones’ and is associated with a high ratio of oxygenated to deoxygenated haemoglobin. This blood-oxygen-level dependence (BOLD) is the MRI contrast in functional imaging and the readout on the scanner reflects areas of brain activity. Though Seiji Ogawa at AT&T Bell labs was originally lauded for recognising BOLD as a contrast for fMRI in 1990, his work was closely followed by crucial contributions from other scientists who performed the first *in vivo* fMRI studies using externally administered and endogenous contrast agents.

The launch of MRI/fMRI opened new portals in neurology – diffusion MRI revolutionised diagnosis of demyelinating neurodisorders like multiple sclerosis and fMRI enabled ‘neural activity’ studies on even the deeper regions of the brain unlike EEG and MEG, which measure electrical activity restricted to the cortical surface.

More importantly, the advent of fMRI saw a surge in social and cognitive neuroscience research since the technique offers ample scope for a neuroscientist to delve into local brain activity that 'correlates' with cognitive tasks undertaken by the subject in the scanner. Using such 'reverse inferences', the scientist is able to predict, based on behavioural variables of the subject, whether a particular cognitive process is engaged. This is then correlated to 'active' regions on the brain scans. But are such 'speculations' really reliable? We're about to find out.

fMRI receives criticism

Despite all its advantages, fMRI has been receiving sceptical remarks from the scientific community owing to several loopholes in the statistical methods in imaging analyses.

A couple of years ago, Ed Vul, then a grad student at the Department of Brain and Cognitive Sciences at the Massachusetts Institute of Technology (MIT), wrote the flagrant paper "Voodoo Correlations in Social Neuroscience" retitled "Puzzlingly High Correlations in fMRI Studies of Emotion, Personality and Social Cognition" (*Perspect Psychol Sci* 4(3):274-90). His observations came as an eye-opener, though not without a flurry of controversy, to social neuroscientists. What exactly are those slip-ups in today's imaging analyses? Before getting there, we must quickly recall how neuroimaging analyses work...

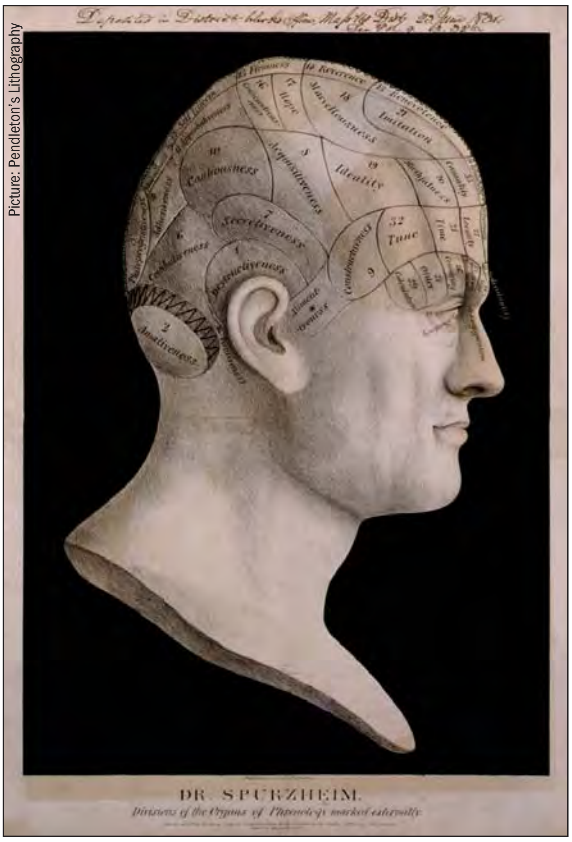
Inferences in fMRI are drawn from the following stereotypic drill (*Trends Cogn Sci*, 10(2):59-63):

- In the present study, when task A was presented, brain area Z was active
- In previous studies, when cognitive process X was putatively engaged, brain area Z was active
- Thus, brain activity in area Z in the present study demonstrates engagement of cognitive process X during execution of task A.

To elaborate on the derivative process, let us take an example. An fMRI study on rats (*J Neurosci*, 25:149-156) compared activity in the striatum during pup-suckling versus cocaine administration. Suckling as opposed to cocaine triggered greater activity in the dorsal and ventral striatum of the pups and led the authors to conclude that the former is more 'rewarding'. The cognitive process 'reward' was thus, *inferred* from activation of the striatum. It is hence, evident that 'reverse inference' in most cases is only an *implied* correlation of a cognitive process to an area of brain activity. The assumption can be deductively valid only if statement 2 is exclusive, i.e. in the case that area Z is activated *if and only if* cognitive process X is engaged.

Overestimated correlations

The major challenge in reverse inferences lies in ascertaining the selective activation of a brain area by a certain cognitive process. Since several discrete cognitive processes can account for neural activity in a single brain area, the neural correlations to behaviour are weak derivations. For this reason, Russell Poldrack (*Trends Cogn Sci*, 10(2):59-63) emphasises the need to estimate selectivity by using one of the online databases viz. the BrainMap database, which compile several thousand comparisons to provide a useful proof of concept. Nevertheless, as Ed puts it, 'voodoo correlations' do exist. Here is a gist of the arguments that he came up with when interviewed by Jonah Lehrer (*Scientific American*, 2009).



Picture: Pendleton's Lithography

For phrenologists like Gall or Spurzheim, neuroimaging would have been an attractive tool to visualise the 'trigger' that controls personality or emotionality. Too bad that, nowadays, no-one believes in pseudo-science anymore.

Ed was a grad student in the lab of Nancy Kanwisher at MIT working on fMRI data, when reports of high correlations (of greater than 0.8) between behaviour or self-report measures of personality or emotion and measures of brain activation caught his attention. He was strangely puzzled by certain inferences viz. "increased activity in a specific brain area that highly correlated with the speed at which a subject walked out of a room... the fMRI activity was, in fact, measured some two hours before the walking happened". After several such instances, Ed was convinced that there was something fishy going on with the statistics and so, set out to investigate.

A typical fMRI scan produces a lot of data on the brain, which is divided into many little cubic units called voxels. In a high-resolution fMRI scan, there are about 40-60,000 voxels in the 3D picture. Now, when one has to pick some of these innumerous voxels to find a correlation to a certain aspect of behaviour, the easiest thing to do is to choose voxels in a cluster that show a high correlation with this behaviour. What distorts results is that the researchers then go on to report a quantitative measure of the correlation magnitude obtained just from within the voxels that they have 'pre-selected' for having a high correlation i.e. average of voxels that pass a certain threshold within each cluster and *not* of *all* the voxels in the cluster. Ed's argument is, "If researchers select only highly correlated voxels, they select voxels that 'got lucky', as well as having some underlying correlation". This bumps up correlation numbers enormously.

A game of luck

To find a day-to-day analogy to blowing-up correlation numbers by virtue of chance, let us take the 'stock investments' scenario. When a group of investment analysts do exceeding well in their stock picks for a certain month, it is a combination of both talent and luck (with the latter doing more) that favours them. If we were to correlate their 'talent' to doing well in stock picks and ask the same group to predict again the next month, they cannot duplicate their performance in the last month. The reason is that, this time, luck favours a few and hurts the others but, earlier, the whole group had luck on their side. So, their average performance in the previous month is an overestimate of their true talent viz. the performance that they can duplicate in the average month.

In essence, Ed picked on the non-independence of imaging analysis, since this way researchers pre-define voxels in a cluster, which have a high correlation and run additional analyses of the same clusters in the same dataset and estimate a high average correlation. "This practice", as Ed points out, "inflates the correlation measurement because it selects those voxels that have benefited from chance, as well as any real underlying correlation, pushing up the numbers."

Ed's claims were well-supported by his estimates on the reliability of fMRI and personality or emotional measures. His data showed evidently limited reliability of the two measures, to which correlation claims greater than 0.8 do not conform.

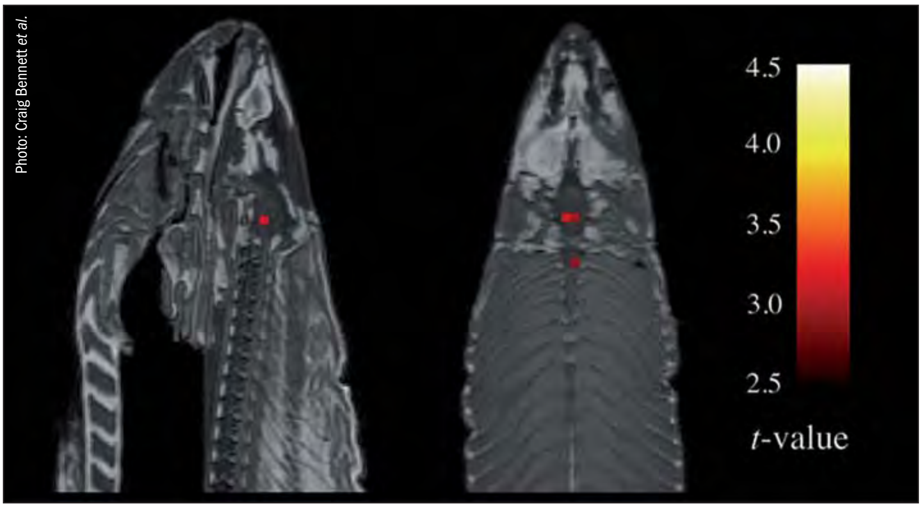


Photo: Craig Bennett *et al.*

From the poster Bennett *et al.* presented at the 15th Annual Meeting of the Organization for Human Brain Mapping, San Francisco, CA, June 2009: One mature Atlantic Salmon (*Salmo salar*) participated in the fMRI study. The salmon was approximately 18 inches long, weighed 3.8 lbs and was not alive at the time of scanning. Several active voxels were discovered in a cluster located within the salmon's brain cavity.

To sum up, at the root of voodoo correlations, as Ed puts it, are "analyses inadvertently capitalized on chance, resulting in inflated measurements of correlation".

Ed's strike at the problem

What then is a sensible alternative to non-independence analyses? A good solution that Ed proposes is to select voxels from different regions of interest objectively. In doing so, one should be "blind" to the correlations of those voxels with the behavioural measure besides being aware that individuals' brains exhibit enormous variability. Once a subset of voxels, say those within an anatomical confine that are significantly active, is defined, one number (e.g. the mean signal change) should be aggregated from these voxels. Only then should the behavioural data be examined, and an unbiased correlation be computed between the brain area of interest and the behavioural measure. His paper also suggests a 'split half' analysis to draw voxelwise correspondences between functional anatomies of two subjects.

It's good news that Ed's paper, besides merely criticising existing pitfalls, also offers tips for more reliable and simpler analyses of fMRI data to eventually expel miscalculations arising from non-independence measurements.

The infamous fish

The 'dead salmon' experiment was another story that turned many heads, urging scientists to pull up their socks and reconsider the analytical strategies in functional imaging.

The story of the 'dead salmon' was sparked off within the premises of Dartmouth College when an eccentric grad student working on fMRI, Craig Bennett, scribbled a list of weird objects to scan. Well, the situation was typical of the crazy things you might do when sticking around in your lab on a weekend, like developing a thumb print on an X-ray film or pulling a strand of your hair to check the grey under the microscope. There was no objective when Craig went on to place a dead salmon into the scanner on a Saturday morning in the spring of 2005. The fish was shown a "series of photographs depicting human individuals in social situations". And, as part of the "ritual", the fish was "asked" to determine what emotion the individual in the photo must have been experiencing. Craig put away the data from the functional runs in his bottom drawer, until 2008. After all, a dead fish was a bit of a nitwit thing to do...

In 2008, when Craig was discussing an issue of false positives associated with the multiple comparisons problem, he was prompted to analyse his data obtained three years previous. And the rest is history... Craig was startled at what he saw – active voxels in the scan, right along the midline of the salmon's brain! A dead fish conceiving human emotions? Or just a red herring?

The result that was completely nuts had a point – a major word of caution. The authors turned it into a paper as a warning about the dangers of false positives in fMRI data. In a set of fMRI scans of a hundred-thousand voxels with inherent noise, multiple statistical tests done simultaneously are bound to produce significant results. But here again, as with Ed's argument, the difference may arise merely by random chance. Neuroscientists can embellish the data in a way that any noise is hidden but, in doing so, rigorous statistical checks have to be included. Craig, in an interview with Alexis Madrigal (*Wired Science*, 2009), said, "We could set our threshold so high that we have no false positives, but we have no legitimate results, or we could also set it low that we end up getting activity in fish's brain. It's a fine line that we walk."

Hopeful for the future

Both Ed Vul and Craig Bennett were making the same point. fMRI data could be deceitful, if neuroscientists remain unaware of the repercussions of shallow statistical analyses. While Ed's work calls for non-independence testing, Craig with his fish has proven the need for multiple comparisons correction in imaging analyses. One thing that they have in common is hoping the best for the future. "The point of the salmon study isn't to prove that fMRI shouldn't be used or is worthless. Brain scientists can do things with fMRI machines they otherwise couldn't", remarked Ed a few years ago. Has the work of the duo and of many others who lit up the fallacies in functional imaging persuaded neuroscientists to check their lab-books? Has the fish been good food for the brain? Or has it just been swallowed down without being tasted? Only time will tell.

Madhuvanthi Kannan