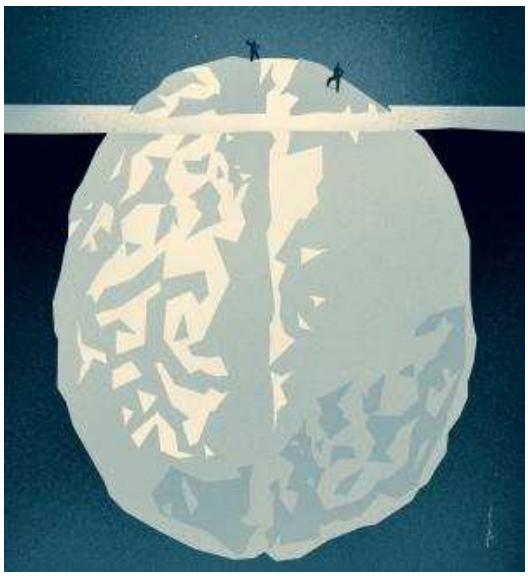
My New Scientist

Hidden depths: Brain science is drowning in uncertainty

17 October 2013 by **Ingfei Chen**
Magazine issue 2939. **Subscribe and save**
For similar stories, visit the **The Human Brain** Topic Guide

The edifice of research being built with brain scans is flawed. It's time to rethink the approach to build a more complete understanding of the mind

Editorial: "Neuroscience wrongs will make a right"



(Image: Paweł Jonca)

1 more image

IT'S FOUR in the afternoon when I meet John Ioannidis, but lines of fatigue are deepening under his eyes. He's exhausted with jet lag after a whirlwind tour of 20 European cities, where he's been lecturing and brainstorming with colleagues. In a corner of his office, I spot two oddly shaped bags, which hold gear for his sport of choice, épée fencing. It seems a fitting hobby for this soft-spoken professor, who is a crusader for good science.

Statistical logic and careful scrutiny of evidence are the weapons that Ioannidis nimbly wields. His previous targets have included spurious claims about drugs and other medical treatments from clinical trials backed by the pharmaceutical industry. Now his gaze has turned to the brain. Joining a growing army of critics, he has documented serious flaws in the ways that many – if not the vast majority of – neuroscience studies are designed, analysed and reported.

That should perhaps be a warning whenever we read headlines about studies capturing snapshots of the brain on "love", "fear", "religion" or "politics". It turns out that many of those colourful brain scans may offer little more than mirages, obscuring the true picture of the human mind in action.

Worst still, the problems are not just confined to a few misleading brain-scan reports. From experiments investigating the action of genes and individual molecules to studies linking brain structure to mental health, question marks are now hanging over the whole field of neuroscience. "Currently, I wouldn't put much trust in most of the literature," says Ioannidis, who is an epidemiologist at the Stanford University School of Medicine in California.

Amid these concerns, it might seem as if our understanding of the brain is set to disappear in a fog of uncertainty, and you will find many observers in the popular press who are now bashing "neuromania". But it's important not to forget the advances of the last century. And while the tough conclusions of Ioannidis and his colleagues are certainly reason to reassess our knowledge, their insights should only lead to more fruitful efforts in uncovering the mind's mysteries. "Neuroscience is moving forward," says Chris Baker of the US National Institute of Mental Health (NIMH). As the fog clears, more nuanced theories should, in time, emerge in sharp relief.

Although philosophers have long pondered the origins of thought, it was the invention of functional magnetic resonance imaging in 1991 that really sparked our love affair with neuroscience. fMRI is based on studying the flow of blood in the brain, with more blood rushing to the areas working hardest. The scans reveal bright splotches of neural activity inside people's heads as they engage in different tests of their capacity to see, feel, remember or think. We were instantly seduced by these technicolour insights.

But there were always some quiet grumblings about whether transient neural activity could reveal much about complex mental processes or behaviours. But the brain-scan backlash only really exploded into public view in late 2008, when psychology researchers Edward Vul and Harold Pashler at the University of California, San Diego, published a critique of what they cheekily dubbed "voodoo correlations". The pair had been baffled by a profusion of highly implausible fMRI results strongly linking behaviours or traits to one or just a few specific areas of the cortex. Examining 53 fMRI studies, Vul, Pashler and their colleagues concluded that half of them reported untrustworthy results that were simply too good to be possible, thanks to "seriously defective" methods (*Perspectives on Psychological Science*, vol 4, p 274).

Double dipping

To understand why, first consider that a typical fMRI scan of the whole brain contains as many as 100,000 three-dimensional pixels, called voxels – a vast amount of data to analyse. Researchers use specialised software to find clusters of voxels that light up when participants view images that trigger, say, empathy or emotional responses. However, the challenge is that true signals can be obscured by underlying random fluctuations in those voxels – a bit like the static noise on an untuned TV. fMRI software tries to filter that out but it cannot work miracles, so many areas will inevitably show some increased activity simply by fluke.

Ideally, neuroimagers should use two sets of scans. One set is for identifying which voxel clusters are highly activated during the experiment. Having found these regions, you then look at them specifically in the second set of scans to confirm that the response wasn't due to random fluctuations, and then measure its size. But Pashler and Vul found that many researchers instead made the mistake of using just one data set for both the initial and final analysis, which allows the random noise to inflate an apparent link to a behavioural response or trait. Such "double-dipping" led researchers to some exciting but premature conclusions, including overly simplistic ideas about the origin of personality traits. Neuroticism, for instance, was chalked up to stronger activity in a pair of almond-shaped regions called the amygdalae, which are known to be involved in fear and other negative emotions.

Confirming that the problem was spread far and wide, Baker and colleagues at NIMH looked at all the fMRI studies published in five top journals in 2008. Of the 134 papers, 42 per cent had made double-dipping errors (Nature Neuroscience, vol 12, p 535). The flawed method is also common in studies of single-neuron responses in animals, as well as in genetic analyses, Baker's team noted.

Neither critique went as far as overturning the broader conclusions of the studies in question. "It doesn't invalidate everything," Baker says of his work, "but it raises question marks."

The voodoo correlations study, in particular, set off an angry back-and-forth of rebuttals. One cause of criticism was that Vul and Pashler named offending studies, which some said was overly aggressive. "It came across as a little bit nasty," says fMRI specialist Russell Poldrack of the University of Texas in Austin, though he admits that it got people's attention. "I don't know that a paper that was written more nicely would have necessarily had as much impact."

"We spoke frankly and just kind of had a little fun," says an unrepentant Pashler, while acknowledging that since he and Vul do not themselves do brain-scanning research, they had not needed to worry about how their next studies or grant applications would be received.

After the furore died down, many fMRI researchers realised that the critiques were essentially right. Voodoo correlations and double-dipping appear to be less common now, and the idea that you can map complex personality traits to a few specific regions like the amygdalae is increasingly considered to be "a pipe dream", says cognitive neuroscientist Tal Yarkoni, also at the University of Texas at Austin. Personality traits are now thought to be associated "with lots of different brain regions interacting in complex ways", he says.

But as researchers patched up those holes in their methods, other equally serious concerns began to emerge. Last year, for instance, a jaw-dropping study from the University of Michigan demonstrated that an fMRI experiment could be analysed in nearly 7000 ways – and the results could vary hugely. With so much flexibility, neuroimagers can unintentionally (or indeed deliberately) analyse their experiments in a way that yields the most favourable results. One tongue-in-cheek report showed that even a dead salmon's brain could appear to be "thinking" inside a scanner if the wrong techniques were used.

The most alarming wake-up call came from Ioannidis this May, with a paper showing that the problems run much deeper than flawed fMRI studies. Working with experimental psychologists Katherine Button

and Marcus Munafo of the University of Bristol, UK, and others, he analysed 48 review papers that collectively had scrutinised 730 studies examining the risk factors and treatments for neurological disorders such as Alzheimer's disease and chronic pain. The experiments used many different methods, including measures of cognitive functioning, gene testing, and clinical trials. From this, the team estimated the odds that each study was able to detect something that was truly there to be discovered – otherwise known as its "statistical power".

The results were grim. The average overall power was about 20 per cent, largely because the number of subjects used in the experiments was simply too small for reliable results to come out of them, even if they passed the standard statistical tests (Nature Reviews Neuroscience, vol 14, p 365). In other words, four out of five studies might have been missing the actual biological effect or mechanism sought, and therefore reported false negatives.

But that's not all. The low power delivers a double whammy of uncertainty: not only are you likely to be missing the evidence even if it's under your nose, but "if you do detect something that seems to be significant, it has a higher chance of being a false positive", Ioannidis says.

The picture was even more troubling when looking specifically at structural MRI studies that studied the physical anatomy of the brain (as opposed to the changing neural activity that shows up in functional MRI). The average statistical power of studies linking structural abnormalities to mental health conditions such as depression or autism was a feeble 8 per cent – meaning that 92 per cent of the investigations would have failed to make true discoveries and in many cases detected something that was not really there.

Data dredging

As in many fields, published neuroscience studies tend to show more positive results than would be expected, something Ioannidis and his colleagues confirmed through further work examining bias in fMRI investigations and animal studies of neurological illnesses. Some of this bias arises simply because negative studies are not published very often. But another possibility, Ioannidis says, is "data dredging" – researchers fishing through and analysing subsets of their results until they find something favourable to publish.

To know exactly which or how many of the reports are right or wrong would mean attempting to replicate all the findings, which usually is not done. But based on his experience with other research fields, Ioannidis thinks that the vast majority of neuroscience studies published these days are likely to be incorrect. "Neuroscience is in serious trouble," he says.

What is to be made of this damning assessment? For a start, it does not mean ditching everything. Conclusions that have stood the test of time are more believable, and Ioannidis is not questioning textbook knowledge of brain anatomy and function. Injury from a stroke in Broca's area, for instance, obviously impairs the ability to speak, so we can be sure of its role in language production. Such big effects can be discerned even by studying just a small number of people, and have been corroborated by many strands of evidence.

It is probably the newer findings that we should take with a pinch of salt, particularly as neuroscientists tease apart the finer processes that are likely to underlie many complex mental tasks, behaviours or differences in personality traits. Such phenomena are much harder to measure, and since the patterns of brain activity are so faint, a lot more data must be collected before the true signal can be detected above the background noise.

Unsurprisingly, Ioannidis has ruffled many feathers, though many neuroscientists agree with the gist of his findings. The big concern is that he is being too alarmist. Poldrack, for instance, is concerned that the ideas may be "spun into this kind of global nihilism that all of neuroscience is bullshit". Certainly, no one should be saying that. Many fMRI findings have held up over time, including observations that the frontal cortex is always activated during short-term recall and that the hippocampi are active during sleep, perhaps as they work to consolidate memories.

"I wouldn't keep doing science if every time I found something, I later found that it was unreliable," Poldrack says. While there are certainly problems, he adds, "many of us are doing what we can to try to address them". But some researchers worry that if governments get the wrong message, they may starve labs of funding, killing revolutionary research.

For his part, Ioannidis is adamant that transparency is the best way to keep the public's confidence in science. "I don't like hiding things under the carpet. I prefer to identify issues and solve them."

And he does have a prescription to cure many of those ills. For example, bigger sample sizes – such as in rigorous multi-centre studies – are often the most obvious way to increase statistical power when looking at small, hard-to-detect effects. Alternatively, for some research questions, studying a few subjects can still produce reliable results, if you gather enough data from each person. Increasing the size of fMRI studies can be challenging, however, since it costs around $500 per hour to use a machine, though arguably the funds are better spent on larger but fewer studies.

Another approach is to encourage brain scientists to disclose their data and replicate others' findings and so weed out some of the false positives. For instance, in 2010, Poldrack and several colleagues launched the web-based Open fMRI Project, which lets investigators upload their raw data sets so that others can reanalyse and validate their results. Replication is a thankless task, though, since researchers don't get promoted for being right or confirming ideas – they get promoted for publishing intriguing new results.

If all this seems like a struggle, neuroscientists may take heart from genetics research, which faced a similar upheaval a decade ago after a flood of small studies overemphasised the role of particular genes in disease and personality traits. Now, with much bigger studies and consistent rules for reporting and sharing data, that field has gone from a replication rate of 1 per cent to more than 90 per cent reliability, Ioannidis says.

Indeed, a couple of large initiatives are already tackling these challenges. In the $40 million Human Connectome Project, neuroscientists across a dozen institutions are building a detailed wiring diagram of the brain's circuitry. They will be scanning a large sample – 1200 people – using fMRI and a technique called diffusion imaging, and the data will be openly shared. It promises to give us our best view yet of the way the brain's anatomy shapes thought and behaviour.

Building bridges

The forthcoming BRAIN Initiative, meanwhile, will receive $100 million of US government funding to develop techniques that will pick out the finer circuits in the brain, bridging the gaps between studies examining single neurons and the large-scale fMRI maps. That will include rethinking or refining existing techniques, such as "optogenetic" methods that allow you to control neuronal activity with pulses of light, as well as inventing entirely new technologies.

All this may mean we will finally be able to appreciate the complexity of the brain. Jack Gallant at the University of California, Berkeley, for instance, points out that there is so much more for us to see if only we pay attention to the bigger picture. At the moment, it's as if we've been peering at the brain through a lousy microscope, he says – partly due to the fact that most MRI data is thrown away to focus on a few selective results. "We're missing huge things," he says.

Consider our understanding of face recognition – a knotty task for the brain, given just how much our expressions can vary. Typical fMRI experiments would compare just two conditions, such as showing volunteers pictures of faces versus places. Based on such investigations, neuroscientists used to think that one region of the brain – the so-called fusiform face area (FFA) – uniquely responds whenever a person sees a face. But the story has grown more complicated as further research turned up a network of other regions that cooperate to recognise faces.

And as neuroimaging grows more sophisticated, so too does our view of the brainscape. Gallant's experiments, for instance, collect hours of brain-scan data from a few subjects as they watch movie trailers, which allows the team to track the brain's changing reaction to an immensely wider range of stimuli. The researchers then skip some of the usual processing steps that lead to data loss, so that they can draw as much meaningful data as possible from the experiment.

Their soon-to-be-published results show the FFA to be even more intricate than previously imagined, suggesting that it can be subdivided into three separate areas. While these all respond to faces, each is also involved in processing different categories of other objects, such as flags, crucifixes and snakes – making the FFA something like a Swiss Army knife for visual recognition. That doesn't mean the initial view of the "face area" was wrong. It was just incomplete, Gallant says. (Although others point out more studies will be necessary to confirm that the same principles apply to the average brain, and not just the few volunteers in this sample.)

Such deepening complexity in the understanding of the brain is to be expected as one's microscope gets better and better. But Gallant says it just goes to show that we are still at the tip of the iceberg when it comes to the prevailing theories. Even with all the work on visual perception, for instance, no one can yet build a robot that sees like a human, let alone a machine that accurately recognises

people. And phenomena like emotions or moral judgement are even murkier.

Will the understanding come eventually? Gallant remains an optimist, pointing out that fMRI was invented only 20 years ago. Back then, nobody knew how best to design the experiments and analyse the vast data they generated, but he thinks the lessons of past mistakes, such as double-dipping, will be learned. The study of the brain just gets better all the time, he says.

Even the épée-fencing Ioannidis agrees that our understanding of the brain will eventually correct itself where wrong. But the question is, how quickly? "If it takes several years for something to be refuted, that could be a real waste of effort." As he points out, "the brain is more complex than almost any other system".

Few questions are as profound as the mysteries between our ears, and there is no doubt that solving them will need the finest tools wielded with the greatest skill.

This article appeared in print under the headline "Hidden depths"

A sceptic's guide to neuromania

While the neuroscientist's toolkit comes into question (see main article), there are also many common pitfalls in the way the results are interpreted to explain complex traits and behaviours.

For instance, you will often read about differences in brain activity or structure that appear to be linked to psychopathic tendencies, with studies showing that convicted murderers have reduced activity in areas associated with empathy when they see images of people suffering. Defence lawyers might use this as evidence that a defendant had diminished responsibility, and some pundits have even pondered whether it might be possible to identify people who are more likely to commit a crime. But there are probably plenty of people who show the similar quirks in the brain scanner, with no criminal intentions. (Indeed, doctors are thought to tone down their own empathic response to pain to help them manage a patient's distress.) And differences in a murderer's brain may be the result of their past brutality, not the cause.

Similar "neurocentric" arguments are sometimes used when talking about drug abuse as a "brain disease". There is no doubt that addictive substances do create long-lasting changes to our neural circuity, but as psychiatrist Sally Satel and clinical psychologist Scott Lilienfeld point out in their book *Brainwashed* (Basic Books, 2013), this view can devalue many other factors, including stress, the influence of friends, and access to drugs. In this way, it might distract addicts from psychological strategies such as avoiding cues that trigger a craving. Satel and Lilienfeld also point out that placing all the blame on the brain's circuits could diminish people's belief in their own self-control, whereas about 80 per cent of addicts do manage to kick the habit.

Brain science clearly has big potential for medicine and the law. But it is crucial to realise that our neurology need not rule our fate. **David Robson**

Ingfei Chen *is a writer based in Santa Cruz, California*

From issue 2939 of New Scientist magazine, page 32-37.

MORE FROM NEW SCIENTIST

Clucking hell: The nightmare world without chickens

Battle-scarred Earth: How war reshapes the planet

Risky business: Rare events happen surprisingly often

Feedback: Striving to name polyfailure

Back to article