

http://www.sciencenews.org/view/feature/id/50295

Home / December 19th, 2009; Vol.176 #13 / [Feature](#)

TRAWLING THE BRAIN

New findings raise questions about reliability of fMRI as gauge of neural activity

By Laura Sanders

December 19th, 2009; Vol.176 #13 (p. 16)



[Enlarge](#)

Trawling the brain

A region of a dead Atlantic salmon's brain appeared to glow with activity (red) in response to emotional scenes. Statistical checks corrected the spurious findings.

fMRI: C. Bennett/UCSB, fish: Andreyuu/istockphoto

The 18-inch-long Atlantic salmon lay perfectly still for its brain scan. Emotional pictures — a triumphant young girl just out of a somersault, a distressed waiter who had just dropped a plate — flashed in front of the fish as a scientist read the standard instruction script aloud. The hulking machine clunked and whirred, capturing minute changes in the salmon's brain as it assessed the images. Millions of data points capturing the fluctuations in brain activity streamed into a powerful computer, which performed herculean number crunching, sorting out which data to pay attention to and which to ignore.

By the end of the experiment, neuroscientist Craig Bennett and his colleagues at Dartmouth College could clearly discern in the scan of the salmon's brain a beautiful, red-hot area of activity that lit up during emotional scenes.

An Atlantic salmon that responded to human emotions would have been an astounding discovery, guaranteeing publication in a top-tier journal and a life of scientific glory for the researchers. Except for one thing. The fish was dead.

The scanning technique used on the salmon — called functional magnetic resonance imaging — allows scientists to view the innards of a working brain, presumably reading the ebbs and flows of activity that underlie almost everything the brain does. Over the last two decades, fMRI has transformed neuroscience, enabling experiments that researchers once could only dream of. With fMRI, scientists claim to have found the brain regions responsible for musical ability, *schadenfreude*, Coca-Cola or Pepsi preference, fairness and even tennis skill, among many other highly publicized conclusions.

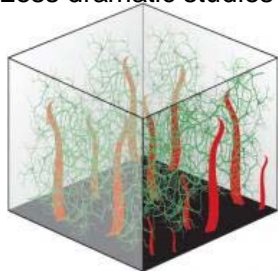
But many scientists say that serious issues have been neglected during fMRI's meteoric rise in popularity. Drawing conclusions from an fMRI experiment requires complex analyses relying on chains of assumptions. When subjected to

critical scrutiny, inferences from such analyses and many of the assumptions don't always hold true. Consequently, some experts allege, many results claimed from fMRI studies are simply dead wrong.

"It's a dirty little secret in our field that many of the published findings are unlikely to replicate," says neuroscientist Nancy Kanwisher of MIT.

A reanalysis of the salmon's postmortem brain, using a statistical check to prevent random results from accidentally seeming significant, showed no red-hot regions at all, Bennett, now at the University of California, Santa Barbara, and colleagues report in a paper submitted to *Human Brain Mapping*. In other words, the whole brain was as cold as a dead fish.

Less dramatic studies have also called attention to flawed statistical methods in



Enlarge

Voxel innards

A typical voxel contains a dense tangle of neurons and other kinds of brain cells (green) and blood vessels (orange). Blood vessels make up a relatively small amount of a voxel, but BOLD fMRI analyses assume that changes in blood flow represent the coordinated activity of millions of neurons.

T. Dube, adapted from N. Logothetis/Nature 2008

fMRI studies. Some such methods, in fact, practically guarantee that researchers will seem to find exactly what they're looking for in the tangle of fMRI data. Other new research raises questions about one of the most basic assumptions of fMRI — that blood flow is a sign of increased neural activity. At least in some situations, the link between blood flow and nerve action appears to be absent. Still other papers point out insufficient attention to insidious pitfalls in interpreting the complex enigmatic relationship between an active brain region and an emotion or task.

Make no mistake: fMRI is a powerful tool allowing neuroscientists to elucidate some of the brain's deepest secrets. It "provides you a different window into how mental processes work in the brain that we wouldn't have had without it," says Russell Poldrack of the University of Texas at Austin.

But like any powerful tool, fMRI must be used with caution. "All methods have shortcomings — conclusions they support and conclusions they don't support," Kanwisher says. "Neuroimaging is no exception."

BOLD assumptions

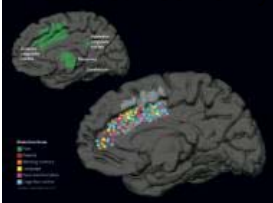
fMRI machines use powerful magnets, radio transmitters and detectors to peer into the brain. First, strong magnets align protons in the body with a magnetic

field. Next, a radio pulse knocks protons out of that alignment. A detector then measures how long it takes for the protons to recover and emit telltale amounts of energy. Such energy signatures act as beacons, revealing the locations of protons ensconced in specific molecules.

fMRI is designed to tell researchers which brain regions are active — the areas where nerve cells are abuzz with electrical signals. Scientists have known for a long time how to record these electrical communiqués with electrodes, which can sit on the scalp or be implanted in brain tissue. Yet electrodes outside the skull can't precisely pinpoint active regions deep within the brain, and implanting electrodes in the brain comes with risks. fMRI, on the other hand, offers a noninvasive way to measure neuron activity, requiring nothing more of the subject than an ability to lie in a big tube for a while.

But fMRI doesn't actually measure electrical signals. Instead, the most common fMRI method, BOLD (for blood oxygen level–dependent), relies on tiny changes in oxygenated blood as a proxy for brain activity. The assumption is that when neurons are working hard, they need more energy, brought to them by fresh, oxygen-rich blood. Protons in oxygen-laden hemoglobin molecules, whisked along in blood, respond to magnetic fields differently than protons in oxygen-depleted blood. Detecting these different signatures allows researchers to follow the oxygenated blood to track brain activity — presumably.

“There's still some mystery,” Bennett says. “There are still some things we don't



Enlarge

Not so simple

View larger version | Most brain regions and mental tasks don't match up one-to-one, confounding the interpretation of fMRI results. Pain activates many regions throughout the brain (upper left, green). One such region, the anterior cingulate cortex, is also activated by many other functions (lower right brain).

Angela Laird/University of Texas Health Science Center at San Antonio,
Jennifer Robinson/Scott & White Memorial Hospital

understand about the coupling between neural activity and the BOLD signal that we're measuring in fMRI.”

Researchers use BOLD because it's the best approximation to neural activity that fMRI offers. And for the most part, it works. But a study published in January in *Nature* reported that the link between blood flow and neural activity is not always so clear. In their experiments, Aniruddha Das and Yevgeniy Sirotnin, both of Columbia University, found that in monkeys some blood changes in the brain had nothing to do with localized neuron firing.

Das and Sirotnin used electrodes to measure neuronal activity at the same time and place as blood flow in monkeys who were looking at an appearing and disappearing dot. As expected, when vision neurons detected the dot and fired,

blood rushed into the scrutinized brain region. But surprisingly, at times when the dot never appeared and the neurons remained silent, the researchers also saw a dramatic change in blood flow. This unprompted change in blood flow occurred when the monkeys were anticipating the dot, the researchers found. The imperfect correlations between blood flow and neural firing can confound BOLD signals and muddle the resulting conclusions about brain activity.

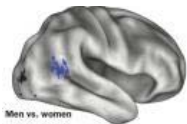
Mass action

Another fMRI difficulty arises from its view-from-the-top scale. Predicting a single neuron's activity from fMRI is like trying to tell which way an ant on the ground is crawling from the top of the Washington Monument, without binoculars. The smallest single unit measured by BOLD fMRI, called a voxel, is often a few millimeters on each side, dwarfing the size of individual neurons. Each voxel — a mashup of volume and pixel — holds around 5.5 million neurons, calculates Nikos Logothetis of the Max Planck Institute for Biological Cybernetics in Tübingen, Germany. Assuming that the millions of neurons in a voxel perform identically is like assuming every single ant on the National Mall crawls north at noon.

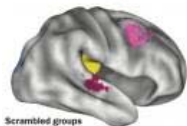
"fMRI is a measure of mass action," Logothetis says. "You almost have to be a professional moron to think you're saying something profound about the neural mechanisms. You're nowhere close to explaining what's happening, but you have a nice framework, an excellent starting point." BOLD signals could reflect many different events, he says. For instance, some neurons send signals that stop other neurons from firing, so increased activity of these dampening neurons could actually lead to an overall decrease in neuron activity.

Kanwisher points out that words such as "activity" and "response," mainstays of fMRI paper titles, are intentionally vague. Pinning down the details from such a zoomed-out view, she says, is impossible. "What exactly are the neurons doing in there? Is one inhibiting the other? Are there action potentials? Is there synaptic activity? Well, we have no idea," she says. "It would be nice to know what the neurons are doing, but we don't with this method. And that's life."

Inadvertent mischief



Men vs. women



Scrambled groups

Enlarge

Uninterpretable brains

In language tasks, men and women showed brain activation in 13 different regions (two shown, top). But when the groups were scrambled, 14 brain regions (three shown, bottom) differed in activity between the random groups, suggesting that fMRI findings need to be replicated.

S.K.Z. Ihnen et al./NeuroImage 2008

After BOLD signals have been measured and the patient has been released from the machine, researchers must sort the red-hot voxels from the dead fish. Statistics for dealing with these gigantic data sets are so complex that some researchers outsource the analyses to professional number crunchers. Choosing criteria to catch real and informative brain changes, and guarding against spurious results, is one of the most important parts of an fMRI experiment, and also one of the most opaque.

“It’s hellishly complicated, this data analysis,” says Hal Pashler, a psychologist at the University of California, San Diego. “And that creates great opportunity for inadvertent mischief.”

Making millions, often billions, of comparisons can skew the numbers enough to make random fluctuations seem interesting, as with the dead salmon. The point of the salmon study, Bennett says, was to point out how easy it is to get bogus results without the appropriate checks.

Bennett and colleagues have written an editorial to appear in *Social Cognitive and Affective Neuroscience* that argues for strong measures to protect against false alarms. Another group takes the counterpoint position, arguing that these protections shouldn’t be so strong that the real results are tossed too, like a significant baby with the statistical bathwater.

One of the messiest aspects of fMRI analysis is choosing which part of the brain to scrutinize. Some studies have dealt with this problem by selecting defined anatomical regions in advance. Often, though, researchers don’t know where to focus, instead relying on statistics to tell them which voxels in the entire brain are worth a closer look.

In a paper originally titled “Voodoo correlations in social neuroscience” in the May issue of *Perspectives on Psychological Science*, Edward Vul of MIT, Pashler and colleagues called out 28 fMRI papers (of 53 analyzed) for committing the statistical sin of “nonindependence.” In nonindependent analyses, the hypothesis in question is not an innocent bystander, but in fact distorts the experiment’s outcome. In other words, the answer is influenced by how the question is asked.

One version of this error occurs when researchers define interesting voxels with one set of criteria — say, those that show a large change when a person is scared — and then use those same voxels to test the strength of the link between voxel and fear. Not surprisingly, the correlation will be big. “If you have many voxels to choose from, and you choose the largest ones, they’ll be large,” Vul says.

In a paper in the May *Nature Neuroscience*, Nikolaus Kriegeskorte of the Medical Research Council in Cambridge, England, and colleagues call the non-independence issue the error that “beautifies” results. “It tends to clean things up at the expense of a veritable representation of the data,” Kriegeskorte says.

Digging through the methods sections of fMRI papers published in 2008 in *Nature*, *Science*, *Nature Neuroscience*, *Neuron* and the *Journal of Neuroscience* turned up some sort of nonindependence error in 42 percent,

Kriegeskorte and colleagues report in their paper. Authors “do very complicated analyses, and they don’t realize that they’re actually walking in a very big circle, logically,” Kriegeskorte says.

Kanwisher, who just cowrote a book chapter with Vul about the nonindependence error, says that researchers can lean too heavily on “fancy” math. “Statistics should support common sense,” she says. “If the math is so complicated that you don’t understand it, do something else.”

The problem with blobology

An issue that particularly irks some researchers has little to do with statistical confounders in fMRI, but rather with what the red-hot blobs in the brain images actually mean. Just because a brain region important for a particular feeling is active does not mean a person must be feeling that feeling. It’s like concluding that a crying baby must be hungry. True, a hungry baby does cry, but a crying baby might be tired, feverish, frightened or wet while still well-fed.

Likewise, studies have found that a brain structure called the insula is active when a person is judging fairness. But if a scan shows the insula to be active, the person is not necessarily contemplating fairness; studies have found that the insula also responds to pain, tastes, interoceptive awareness, speech and memory.

In most cases, the brain does not rely on straightforward relationships, with a specific part of the brain responsible for one and only one task, making these reverse inferences risky, Poldrack points out.

“Researchers often assume that there are one-to-one relations between brain areas and mental functions,” he says. “But we don’t actually know if that is true, and there are many reasons to think that it’s not.” Inferring complex human emotions from the activity of a single brain region is not something that should be done casually, as it is often is, he says.

Sometimes, reverse inference is warranted, though, as long as it is done with care. “There’s nothing wrong with saying there’s a brain region for x,” Kanwisher says. “It just takes many years to establish that. And like all other results, you establish it, and it can still crash if somebody presents a new piece of data that argues against it.”

Marco Iacoboni of the University of California, Los Angeles and colleagues drew heat from fellow neuroscientists for a *New York Times* op-ed in November 2007 in which the team claimed to have ascertained the emotional states of undecided voters as they were presented with pictures of candidates. For instance, the researchers concluded that activity in the anterior cingulate cortex meant that subjects were “battling unacknowledged impulses to like Mrs. Clinton.” Poldrack and 16 other neuroscientists quickly wrote their own editorial, saying that the original article’s claims had gone too far.

Iacoboni counters that reverse inference has a valuable place in research, as long as readers realize that it is a probabilistic measure. “A little bit of reverse inference, to me, is almost necessary,” he says.

Careful language and restrained conclusions may solve some of the issues swirling around fMRI interpretations, but a more serious challenge comes from fMRI's noise. Random fluctuations masquerading as bona fide results are insidious, but the best way to flush them out is simple: Do the experiment again and see if the results hold up. This built-in reality check is time-consuming and expensive, Kanwisher says, but it's the best line of defense against spurious results.

A paper published April 15 in *NeuroImage* clearly illustrates the perils of one-off experiments. In an fMRI experiment, Bradley Schlaggar of Washington University in St. Louis and colleagues found differences in 13 brain regions between men and women during a language task. To see how robust these results were, the researchers scrambled the groups to create random mixes of men and women. Any differences found between these mixed-up groups could be chalked up to noise or unknown factors, the researchers reasoned. The team found 14 "significant" different regions between the scrambled groups, undermining the original finding and rendering the experiment uninterpretable.

"The upshot of the paper is really a cautionary one," Schlaggar says. "It's easy and common to find some group differences at some statistical threshold. So go ahead and do the study again."

In many ways, fMRI has earned its reputation as a powerful neuroscience tool. In the laboratories of capable, thoughtful researchers, the challenges, exceptions and assumptions that plague fMRI can be overcome. Its promise to decode the human brain is real. fMRI "is a great success story of modern science, and I think historically it will definitely be viewed as that," Kriegeskorte says. "Overwhelmingly it is a very, very positive thing."

But the singing of fMRI's praises ought to be accompanied by a chorus of caveats. fMRI cannot read minds nor is it bogus neophrenology, as Logothetis pointed out in *Nature* in 2008. Rather, fMRI's true capabilities fall somewhere between those extremes. Ultimately, understanding the limitations of neuroimaging, instead of ignoring them, may propel scientists toward a deeper understanding of the brain.