The New York Times | http://nyti.ms/2bYUMQm

SundayReview | NEWS ANALYSIS

Do You Believe in God, or Is That a Software Glitch?

By KATE MURPHY AUG. 27, 2016

We've all seen them, those colorful images that show how our brains "light up" when we're in love, playing a video game, craving chocolate, etc. Created using functional magnetic resonance imaging, or fM.R.I., these pictures are the basis of tens of thousands of scientific papers, the backdrop to TED talks and supporting evidence in best-selling books that tell us how to maintain healthy relationships, make decisions, market products and lose weight.

But a study published last month in the Proceedings of the National Academy of Sciences uncovered flaws in the software researchers rely on to analyze fM.R.I. data. The glitch can cause false positives — suggesting brain activity where there is none — up to 70 percent of the time.

This cued a chorus of "I told you so!" from critics who have long said fM.R.I. is nothing more than high-tech phrenology. Brain-imaging researchers protested that the software problems were not as bad nor as widespread as the study suggested. The dust-up has caused considerable angst in the fM.R.I. community, about not only the reliability of their pretty pictures but also how limited funding and the pressure to publish splashy results might have allowed such a mistake to go unnoticed for so long. The remedial measures some in the field are now proposing could be a model for the wider scientific community, which, despite breathtaking technological advances, often produces findings that don't hold up over time.

“We have entered an era where the kinds of data and the analyses that people run have gotten incredibly complicated,” said Martin Sereno, the chairman of the cognitive neuroimaging department at the University of California, San Diego. “So you have researchers using sophisticated software programs that they probably don’t understand but are generally accepted and everyone uses.”

Developed in the 1990s, fM.R.I. creates images based on the differential effects a strong magnetic field has on brain tissue. The scans occur at a rate of about one per second, and software divides each scan into around 200,000 voxels — cube-shaped pixels — each containing about a million brain cells. The software then infers neural activity within voxels or clusters of voxels, based on detected blood flow (the areas that “light up”). Comparisons are made between voxels of a resting brain and voxels of a brain that is doing something like, say, looking at a picture of Hillary Clinton, to try to deduce what the subject might be thinking or feeling depending on which area of the brain is activated.

But when you divide the brain into bitty bits and make millions of calculations according to a bunch of inferences, there are abundant opportunities for error, particularly when you are relying on software to do much of the work. This was made glaringly apparent back in 2009, when a graduate student conducted an fM.R.I. scan of a dead salmon and found neural activity in its brain when it was shown photographs of humans in social situations. Again, it was a salmon. And it was dead.

This is not to say all fM.R.I. research is hooey. But it does indicate that methods matter even when using whiz-bang technology. In the case of the dead salmon, what was needed was to statistically correct for false positives that arise when you make so many comparisons between voxels.

Think of the brain as a jam-packed sports arena and the voxels as all the fans. If you ask everyone in the stadium a bunch of questions, you might, by chance, see a pattern emerge, such as a cluster of people standing in line for the bathroom who love pistachio ice cream and skipped a grade in school. You need to statistically account for the possibility of coincidence before drawing any conclusions about ice cream, intellect and bladder control, just as you would for areas in the brain that light up or don't light up in response to stimuli.

The authors of the paper on the software glitch found that a vast majority of published papers in the field do not make this "multiple comparison" correction. But when they do, they said, the most widely used fM.R.I. data analysis software often doesn't do it adequately.

Other statistical problems in analyzing fM.R.I. data have been pointed out. But these kinds of finger-wagging methodological critiques aren't easily published, much less funded. And on the rare occasions they do make it into journals, they don't grab headlines as much as studies that show you what your brain looks like when you believe in God.

"There is an immense amount of flexibility in how anybody is going to analyze data," said Russell Poldrack, who leads a cognitive neuroscience lab at Stanford University and is a co-author of the "Handbook of functional MRI Data Analysis." And, he continued, "some choices make a bigger difference than others in the integrity of your results."

To try to create some consistency and enhance credibility, he and other leaders in the field recently published a lengthy report titled "Best Practices in Data Analysis and Sharing in Neuroimaging Using MRI." They said their intent was to increase transparency through comprehensive sharing of data, research methods and final results so that other investigators could "reproduce findings with the same data, better interrogate the methodology used and, ultimately, make best use of research funding by allowing reuse of data."

The shocker is that transparency and reproducibility aren't already required, given that we're talking about often publicly funded, peer-reviewed, published research. And it's much the same in other scientific disciplines.

Indeed, a study published last year in the journal Science found that researchers could replicate only 39 percent of 100 studies appearing in three high-ranking psychology journals. Research similarly hasn't held up in genetics, nutrition, physics and oncology. The fM.R.I. errors added fuel to what many are calling a reproducibility crisis.

“People feel they are giving up a competitive advantage” if they share data and detail their analyses, said Jean-Baptiste Poline, senior research scientist at the University of California, Berkeley’s Brain Imaging Center. “Even if their work is funded by the government, they see it as their data. This is the wrong attitude because it should be for the benefit of society and the research community.”

There is also resistance because, of course, nobody likes to be proved wrong. Witness the blowback against those who ventured to point out irregularities in psychology research, dismissed by some as the “replication police” and “shameless little bullies.”

Nevertheless, the fM.R.I. community seems determined to be an exemplar. The next issue of the journal NeuroImage: Clinical will lead with an editorial announcing that it will no longer publish research that has not been corrected for multiple comparisons, and there is a push for other journals to do the same, as well as to require authors to make publicly available their data sets and analyses. Data-sharing platforms such as OpenfMRI and Neurovault have already been established to make fM.R.I. data and statistical methods more widely accessible. In fact, it was data sharing that revealed the fM.R.I. software glitch.

Data repositories have been established in other branches of science, too. Many are supported by the National Institutes of Health, which now requires researchers who receive $500,000 or more in federal funding (as well as those doing large-scale genomic research regardless of funding level) to have a data-sharing plan, although there remain some loopholes as well as limits on who can access the data.

The Wellcome Trust and the Bill & Melinda Gates Foundation have begun to make receiving grants contingent on unfettered data sharing. And the International Committee of Medical Journal Editors has floated a proposal to require authors to share data underlying clinical trials no later than six months after publication of their studies. This, despite stiff opposition from some researchers.

Perhaps the most encouraging sign of increasing transparency is the colorful and meticulously detailed new brain map recently released by the N.I.H.-supported Human Connectome Project. It was compiled using shared data from a variety of technologies, including fM.R.I., so that it could be reviewed, added to and improved upon.

”If we don’t have access to the data, we cannot say if studies are wrong,” said Anders Eklund, an associate professor at Linkoping University in Sweden and a co-author of the study that found the fM.R.I. software bug. “Finding errors is how scientific fields evolve. This is how science gets done.”

Kate Murphy is a journalist in Houston who writes frequently for The New York Times.

Follow The New York Times Opinion section on Facebook and Twitter (@NYTopinion), and sign up for the Opinion Today newsletter.

A version of this news analysis appears in print on August 28, 2016, on page SR4 of the New York edition with the headline: Do You Believe in God, or Is That a Software Glitch?.