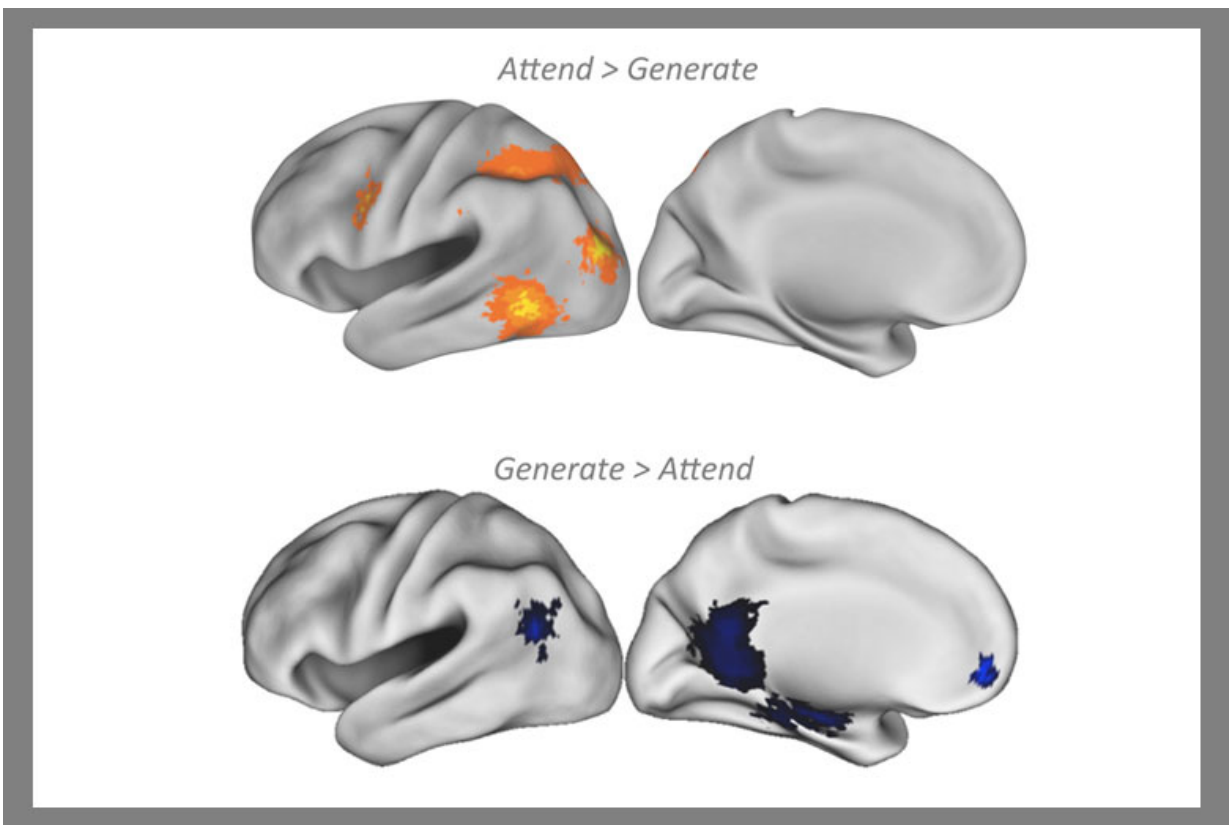


Scientists show fMRI memory detectors can be easily fooled

June 5 2015, by Bjorn Carey



Top images show brain activity as participant pretends to see a familiar face for the first time; bottom images show brain activity when the participant is pretending to recognize a novel face. Both instances fooled the memory decoding algorithm into thinking the subject was telling the truth. Credit: Stanford Memory Laboratory

For the past several years, Anthony Wagner has been developing a computer program that can read a person's brain scan data and surmise, with a high degree of certainty, whether that person is experiencing a memory. The technology has great promise to influence a number of fields, including marketing, medicine and evaluation of eyewitness testimony.

Now, Wagner, a professor of psychology and neuroscience at Stanford, and his colleagues have shown that with just a little bit of coaching and concentration, subjects are easily able to obscure real memories, or even create fibs that look like real memories, on brain scans. The work both confirms the power of the technology when applied to cooperative subjects and underscores the need for more research before applying the science to high-stakes situations.

The results are published in the current issue of the *Journal of Neuroscience*.

The act of creating, storing, recalling and replaying a [memory](#) is a marvelously complex multi-process biological event. First, the stimulus (a person, an event, etc.) must be sensed, interpreted and stored, each step occurring within different brain structures. Recalling the memory relies on yet another set of neural mechanisms to find the distributed information that constitutes an event memory, reassemble it into a full "story," and then replay it in the mind.

Although the process is complex, the fact that different parts of the brain play distinct roles in memory allows neuroscientists like Wagner to pair [brain activity](#) signals with machine learning pattern analyses to detect complex patterns that signal whether a person is remembering a target stimulus or perceiving it as novel.

"We can train our algorithms on one person's data to recognize the

pattern of when they're first seeing something versus when they're remembering something," said study lead author Melina Uncapher, a research scientist in Wagner's lab. "But what's exciting is [that] we can then train the algorithm further so that we are able to use it on other people with fairly high certainty."

In previous studies, the algorithm performed well, detecting a memory in single trials with cooperative subjects with 75 to 95 percent accuracy. But Wagner and Uncapher wanted to know how it would perform against people who attempted to conceal their memory state – a scenario that will be important in, for example, legal contexts. They designed an experiment to establish the current boundary conditions of the technology.

Cheating on a memory test

Wagner and Uncapher recruited 24 subjects to participate in a two-day test. On the first day, the subjects had their brains scanned in an fMRI machine while seeing photos of 200 novel faces, with instructions to try to remember the faces. They were given two seconds to look at each face, and then eight more seconds to create a "story" about that person – a trick that would possibly help them remember the face. After a short break, the faces were shuffled and the subjects repeated the training. The algorithm was 67 percent accurate in detecting the presence or absence of memories when subjects were cooperative or truthful.

On the second day, the subjects returned to the fMRI machine, this time to see 400 faces – an additional 200 novel faces were mixed in with the 200 faces they had seen the previous day. For the first half of the new, combined batch of photos, the subjects were instructed to answer truthfully, "Do you remember seeing this face?" This allowed the algorithm to set a baseline of brain activity for when each subject answered truthfully.

For the remaining half of the study, the subjects were told to try to hide their memory states. If they thought they had seen the face before, they were to say it was new; if it was a novel face, they were to report they had seen it the previous day.

"We gave them two very specific strategies: If you remember seeing the face before, conceal your memory of that face by specifically focusing on features of the photo that you hadn't noticed before, such as the lighting or the contours of the face, anything that's novel to distract you from attending to the memory," Uncapher said. "Likewise, if you see a brand-new face, think of a memory or a person that this face reminds you of and try to generate as much rich detail about that face as you can, which will make your brain look like it's in a state of remembering."

With just two minutes of coaching and training, the subjects became proficient at fooling the algorithm: The accuracy of the decoder fell to 50 percent, or no better than a coin-flip decision.

Interestingly, the subjects reported that it was easier to make up new memories about novel faces than it was to bury real memories from the previous day. The brain scans reflected this shift. When subjects looked at a novel face and pretended they had seen it before, activity spiked in parts of the brain responsible for recalling memories. A similar flip in brain activity was recorded when subjects tried to suppress the memory of a face seen previously.

A significant hurdle

Wagner, who served as an expert witness on one of the first criminal cases involving fMRI lie detection and is a member of the MacArthur Foundation's Research Network on Law and Neuroscience (which funded this research), sees this result as potentially troubling for the goals of one day using fMRI to judge "ground truth" in law cases.

"With this class of technology, you can take a single snapshot of somebody's brain, just 10 seconds of data, and if they're cooperative, you can know above chance if they're recognizing the stimulus or not," said Wagner, who is also a cofounder of the Stanford Program in Neuroscience and Society. "Whether this kind of technology will be able to pull out the truth about memory in all context, I think it's unlikely. In fact, our new data demonstrate one instance in which the technology fails."

Getting to the bottom of that question will be difficult. There are several factors that influence all phases of memory, such as age, personal motivations, the passage of time between encoding and decoding, and whether that memory has morphed over time due to multiple recalls and edits.

In particular, scientists know little about the role of stress in memory, especially during the recall phase, due in large part because in all studies on memory decoding, there are no stakes for the subjects.

"In our studies, and in other published studies, there is no meaningful impact on the [subjects](#) if we are able to decode their memory state from their brain patterns," Wagner said. "In many real world situations, people are stressed when their memories are being probed. What we don't know is what will happen if the classifier algorithm is trained on data where people aren't stressed, and then applied to a stressed subject – who is being deceitful, or trying too hard to be truthful – we don't know what those [brain](#) patterns will look like, and whether we can decode memories from them."

More information: "Goal-Directed Modulation of Neural Memory Patterns: Implications for fMRI-Based Memory Detection." *Journal of Neuroscience*, 3 June 2015, 35(22): 8531-8545; [DOI: 10.1523/JNEUROSCI.5145-14.2015](#)

Provided by Stanford University

Citation: Scientists show fMRI memory detectors can be easily fooled (2015, June 5) retrieved 6 May 2024 from <https://medicalxpress.com/news/2015-06-scientists-fmri-memory-detectors-easily.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.