

Case study

When should scientists share the data?

SOURCE: Barron, D. (2018). How freely should scientists share their data. *Scientific American Blog Network*. <https://blogs.scientificamerican.com/observations/how-freely-should-scientists-share-their-data>

Jack Gallant is a cognitive neuroscientist at the University of California, Berkeley who works on brain decoding technology. In 2016 he showed, what listening to the Moth podcast does to our brains. Before that he showed that basing only on brain activity it is possible to reconstruct images of movies people are watching. His analysis of Moth podcast was published in *Nature*, he has been interviewed by Freakonomics and NPR. Gallant's work has made him a prominent neuroscientist who runs a successful lab.

In 2018 on July 4 Gallant was promoting open science on his Twitter account. He argued that giving away free code is pointless if it only works within an expensive software system. The next day a theoretical physicist Manilo De Domenico tweeted in reply to Gallant: "Nice advice. But what about data? We keep trying to ask access to data in your Nature 2016, but we received not a single reply, yet". To which Gallant replied that "The original authors are still writing further primary research papers on these data so they haven't been released yet but we expect to be able to do that very soon." Another twitter user - Andre Brown pointed out that "We still want exclusivity to publish more papers' isn't a great excuse. Did you note data restrictions in the manuscript?" and referred to Nature's policy that, on publication, authors should make their data, code and protocols "promptly" and publicly available. Therefore, it appeared that Gallant has violated *Nature's* policy and fundamental principles of open science. De Domenico further complained that Gallant's paper has given him several ideas that he would like to test but not having access to Gallant's data he is not able to do that. To this Gallant answered: "And why do you assume that your project is better than the ones that we are continuing with these data? My students and postdocs are an awesome group of people, the stuff they have in the pipeline is great! But I can't afford for them to be scooped." Gallant then affirmed his commitment to open science, that he had shared many datasets in the past and gave further explanation of why he has not yet shared this particular data set - that complex data takes time to understand and his team wanted to work on data more before releasing it and that since his lab has competed for and has won the money to collect the data and then worked to collect it, they should be able to work on first before others do it. Many academics on Twitter were not happy about Gallant's answer. They called it a "nonsense excuse", "scandalous", etc. Someone on *Nature's* website wrote that "Jack Gallant refuses to share the data (in violation with Nature's Journal Policy and with his NSF grants)." Some called to boycott Gallant and to retract his paper.

Questions for discussion:

1. Who is right in this debate? Are the objections to Gallant's position justified? What do you think about Gallant's reasons of not sharing the data set? Does he violate the principles of open science?
2. Why scientists might have reservations about sharing their data?