

RESEARCH TRANSPARENCY

Why null results rarely see the light of day

“File drawer” study proposes registry for unpublished social science data

By Jeffrey Mervis

Researchers have put numbers on the “file drawer” phenomenon, in which scientists abandon results that they believe journals are unlikely to publish.

In a study published online this week in *Science* (<http://scim.ag/AFranco>), a team at Stanford University in Palo Alto, California, traced the publication outcomes of 221 survey-based experiments funded by the National Science Foundation. Nearly two-thirds of the social science experiments that produced null results, those that did not support a hypothesis, were simply filed away. In contrast, researchers wrote up 96% of the studies with statistically strong results (see graph).

Such practices can skew the literature and lead to wasteful duplication, the authors argue. Their remedy: Deposit all data and study designs into public registries. But while most scientists agree that a registry would be valuable, some worry that it would become burdensome and could even introduce new biases. “I wouldn’t want to take all the unpublished findings and give them the same prominence as those containing strong results,” says political scientist Gary King of Harvard University.

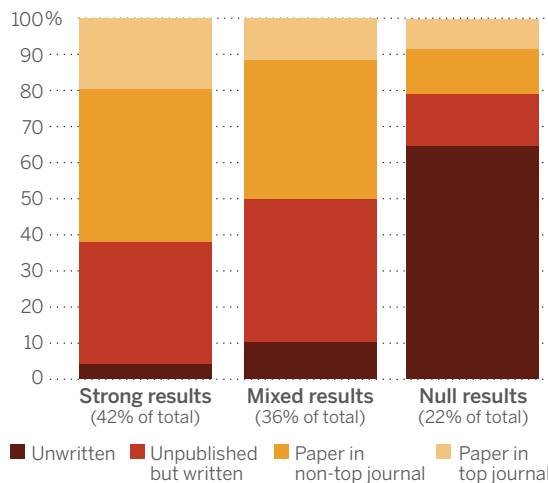
The question of what to do with null results—when researchers fail to see an effect that should be detectable—has long been hotly debated among those conducting medical trials, where the results can have a big impact on lives and corporate bottom lines. More recently, the debate has spread to the social and behavioral sciences, which also have the potential to sway public and social policy. There were little hard data, however, on how often or why null results were squelched. “Yes, it’s true that null results are not as exciting,” King says. “But I suspect another reason they are rarely published is that there are many, many ways to

produce null results by messing up. So they are much harder to interpret.”

In the new study, Stanford political economist Neil Malhotra and two of his graduate students examined every study since 2002 that was funded by a competitive grants program called TESS (Time-sharing Experiments for the Social Sciences). TESS allows scientists to order up Internet-based surveys of a representative sample of U.S. adults to test a particular hypothesis (for example, whether voters tend to favor legis-

Most null results are never written up

The fate of 221 social science experiments



Source: A. Franco et al., *Science* (28 August)

lators who boast of bringing federal dollars to their districts over those who tout a focus on policy matters).

Malhotra’s team tracked down working papers from most of the experiments that weren’t published, and for the rest asked grantees what had happened to their results. In their e-mailed responses, some scientists cited deeper problems with a study or more pressing matters—but many also believed that journals just wouldn’t be interested. “The unfortunate reality of the publishing world are that null effects do not tell a clear story,” said one scientist. Said another, “Never published, definitely

disappointed to not see any major effects.”

Their answers suggest to Malhotra that rescuing findings from the file drawer will require a shift in expectations. “What needs to change is the culture—the author’s belief about what will happen if the research is written up,” he says.

Not unexpectedly, the statistical strength of the findings made a huge difference in whether they were ever published. Overall, 42% of the experiments produced statistically significant results. Of those, 62% were ultimately published, compared with 21% of the null results. However, the Stanford team was surprised that researchers didn’t even write up 65% of the experiments that yielded a null finding.

Scientists not involved in the study praise its “clever” design. “It’s a very important paper” that “starts to put numbers on things we want to understand,” says economist Edward Miguel of the University of California, Berkeley.

He and others note that the bias against null studies can waste time and money when researchers devise new studies replicating strategies already found to be ineffective. Worse, if researchers publish significant results from similar experiments in the future, they could look stronger than they should because the earlier null studies are ignored. Even more troubling to Malhotra was the fact that two scientists whose initial studies “didn’t work out” went on to publish results based on a smaller sample. “The non-TESS version of the same study, in which we used a student sample, did yield fruit,” noted one investigator.

A registry for data generated by all experiments would address these problems, the authors argue. They say it should also include a “preanalysis” plan, that is, a detailed description of what the scientist hopes to achieve and how the data will be analyzed. Such plans would help deter researchers from tweaking their analyses after the data are collected in search of more publishable results.

Some researchers are wary. Requiring a preanalysis plan, for instance, could breed resentment, says James Coan, a social psychologist at the University of Virginia in Charlottesville. “It’s part of a scientist’s job to be canny enough” to do the most appropriate statistical analyses, he says. “The implicit message is that scientists are not to be trusted with those decisions.”

Miguel, a member of the recently formed Berkeley Initiative for Transparency in the Social Sciences, is one of many who say a registry should be encouraged, but not imposed. “We want the community to adopt good norms, and then promote them,” he says. “That’s certainly the most attractive outcome.” ■