~~Putt~~ing hypotheses to the test: We must hold ourselves accountable to decisions made before we see the data.

In the daily practice of doing research, it is easy to confuse what is being done. There is often confusion over whether a study is exploratory (hypothesis-generating) research or confirmatory (hypothesis-testing) research. By defining how a hypothesis or research question will be tested at the outset of research, preregistration eliminates this ambiguity. **David Mellor** *outlines the value of preregistration for determining statistical significance and introduces a new prize for scholars aimed at encouraging this practice.*

We are giving $1,000 prizes to 1,000 scholars simply for making clear when data were used to generate or test a hypothesis. Science is the best tool we have for understanding the way the natural world works. Unfortunately, it is in our imperfect hands. Though scientists are curious and can be quite clever, we also fall victim to biases that can cloud our vision. We seek rewards from our community, we ignore information that contradicts what we believe, and we are capable of elaborate rationalizations for our decisions. We are masters of self-deception.

Yet we don't want to be. Many scientists choose their career because they are curious and want to find real answers to meaningful questions. In its idealized form, science is a process of proposing explanations and then using data to expose their weaknesses and improve them. This process is both elegant and brutal. It is elegant when we find a new way to explain the world, a way that no one has thought of before. It is brutal in a way that is familiar to any graduate student who has proposed an experiment to a committee or to any researcher who has submitted a paper for peer-review. Logical errors, alternative explanations, and falsification are not just common – they are baked into the process.

Image credit: Winnowing Grain Eastman Johnson Museum of Fine Arts, Boston

Using data to generate potential discoveries and using data to subject those discoveries to tests are distinct processes. This distinction is known as exploratory (or hypothesis-generating) research and confirmatory (or hypothesis-testing) research. In the daily practice of doing research, it is easy to confuse which one is being done. But there is a way – preregistration. Preregistration defines how a hypothesis or research question will be tested – the methodology and analysis plan. It is written down in advance of looking at the data, and it maximizes the diagnosticity of the statistical inferences used to test the hypothesis. After the confirmatory test, the data can then be subjected to any exploratory analyses to identify new hypotheses that can be the focus of a new study. In this way, preregistration provides an unambiguous distinction between exploratory and confirmatory research.The two actions, building and tearing down, are both crucial to advancing our knowledge. Building pushes our potential knowledge a bit further than it was before. Tearing down separates the wheat from the chaff. It exposes that new potential explanation to every conceivable test to see if it survives.

To illustrate how confirmatory and exploratory approaches can be easily confused, picture a path through a garden, forking at regular intervals, as it spreads out into a wide tree. Each split in this garden of forking paths is a decision that can be made when analysing a data set. Do you exclude these samples because they are too extreme? Do you control for income/age/height/wealth? Do you use the mean or median of the measurements? Each decision can be perfectly justifiable and seem insignificant in the moment. After a few of these decisions there exists a surprisingly large number of reasonable analyses. One quickly reaches the point where there are so many of these reasonable analyses, that the traditional threshold of statistical significance, $p < .05$, or 1 in 20, can be obtained by chance alone.

If we don't have strong reasons to make these decisions ahead of time, we are simply exploring the dataset for the path that tells the most interesting story. Once we find that interesting story, bolstered by the weight of statistical significance, every decision on that path becomes even more justified, and all of the reasonable, alternative paths are forgotten. Without us realizing what we have done, the diagnosticity of our statistical inferences is gone. We have no idea if our significant result is a product of accumulated luck with random error in the data, or if it is revealing a truly unusual result worthy of interpretation.

This is why we must hold ourselves accountable to decisions made *before* seeing the data. Without putting those reasons into a time-stamped, uneditable plan, it becomes nearly impossible to avoid making decisions that lead to the most interesting story. This is what preregistration does. Without preregistration, we effectively change our hypothesis as we make those decisions along the forking path. The work that we thought was confirmatory

becomes exploratory without us even realizing it.

I am advocating for a way to make sure the data we use to create our explanations is separated from the data that we use to test those explanations. Preregistration does not put science in chains. Scientists should be free to explore the garden and to advance knowledge. Novelty, happenstance, and unexpected findings are core elements of discovery. However, when it comes time to put our new explanations to the test, we will make progress more efficiently and effectively by being as rigorous and as free from bias as possible.

Preregistration is effective. After the United States required that all clinical trials of new treatments on human subjects be preregistered, the rate of finding a significant effect on the primary outcome variable fell from 57% to just 8% within a group of 55 cardiovascular studies. This suggests that flexibility in analytical decisions had an enormous effect on the analysis and publication of these large studies. Preregistration is supported by journals and research funders. Taking this step will show that you are taking every reasonable precaution to reach the most robust conclusions possible, and will improve the weight of your assertions.

Most scientists, when testing a hypothesis, do not specify key analytical decisions prior to looking through a dataset. It's not what we're trained to do. We at the Center for Open Science want to change that. We will be giving 1,000 researchers $1,000 prizes for publishing the results of preregistered work. You can be one of them. Begin your preregistration by going to https://cos.io/prereg.

About the Author:

David Mellor *is a Project Manager at the Center for Open Science and works to encourage preregistration. He received his PhD from Rutgers University in Ecology and Evolution has been an active researcher in the behavioral ecology and citizen science communities.*