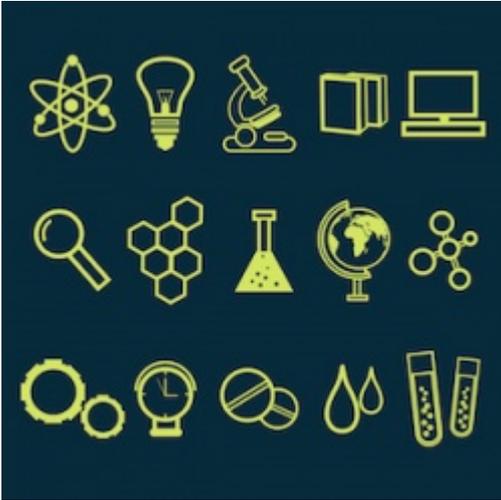# Why Good Science Is Hard to Do

*By Alex Berezow, PhD — September 6, 2018*



Credit: Storyblocks [1]

Neil Sedaka once noted that breaking up is hard to do. Even harder, it seems, is doing good science. Why?

Dr. Ray Dingledine, a pharmacologist at Emory University in Atlanta, attempts to answer this question in a new paper published in the journal *eNeuro.*

His thesis begins with the discomfiting statement that "psychological aspects of decision-making introduce preconceived preferences into scientific judgment that cannot be eliminated by any statistical method." In other words, the fact that a research paper is backed by really solid statistics is meaningless if preconceived assumptions or scientific judgments are incorrect. No amount of fancy math can fix that.

Dr. Dingledine believes this problem arises from several factors: (1) "our drive to create a coherent narrative from new data regardless of its quality or relevance"; (2) "our inclination to seek patterns in data whether they exist or not"; and (3) our negligence to "always consider how likely a result is regardless of its P-value."

**Solutions Require Identifying the Correct Problem**

Dr. Dingledine notes that the crisis in reproducibility has caused a greater emphasis on statistical know-how and adherence to proper scientific experimental design. However, he believes this will be insufficient to address the problem. Instead, he believes that the only real solution is to understand how humans make judgments in the first place, and he turned to the financial and economics literature for guidance.

The author was aware of research conducted by Daniel Kahneman and others that examined

biases in decision-making. Their primary finding was that people tended to rely on unconscious emotions when confronted with unfamiliar data and over-relied on intuition. But since Dr. Kahneman examined undergraduates, Dr. Dingledine thought that he would get different results with a more mature, scientifically literate audience. He did not.

Using the same questions that Dr. Kahneman had created, Dr. Dingledine sent a survey to science faculty, post-docs, grad students, and senior research technicians. Of the 70 he contacted, 44 responded. His results were eye-opening.

He showed that a majority of the respondents could not correctly answer questions involving probability, sampling methods, or statistics. He also found that the respondents were likelier to identify patterns in one set of data but not another, even if both were equally likely. In other words, Dr. Dingledine found that scientists are just like everybody else: They find statistical concepts inherently difficult to grasp and have a natural tendency to seek patterns, even if they don't exist.

**The Trouble with Statistics**

Finally, Dr. Dingledine brilliantly shows why we must consider the likelihood that a result is correct, regardless of its p-value. Just because a p-value is "significant" does not mean that the hypothesis proposed is likely to be true. Using basic math, Dr. Dingledine shows why.

Consider 1,000 hypotheses, 100 of which we know to be true. If our statistical power (i.e., the likelihood of correctly identifying the true hypothesis as true) is 80%, then we will have 80 true positives and 20 false negatives. Of the remaining 900 hypotheses, which we know to be false, an alpha level of 0.05 means we will generate 45 false positives and 855 true negatives. This is shown in Box A below.

**A.**
1000 hypotheses tested
10% are true
α = 0.05
Power = 80%

|  | True + | True − |
|---|---|---|
| Expected outcomes | 80 | 855 |
| | 45 | 20 |
|  | False + | False − |

Chance of a single positive result being true: **64%**
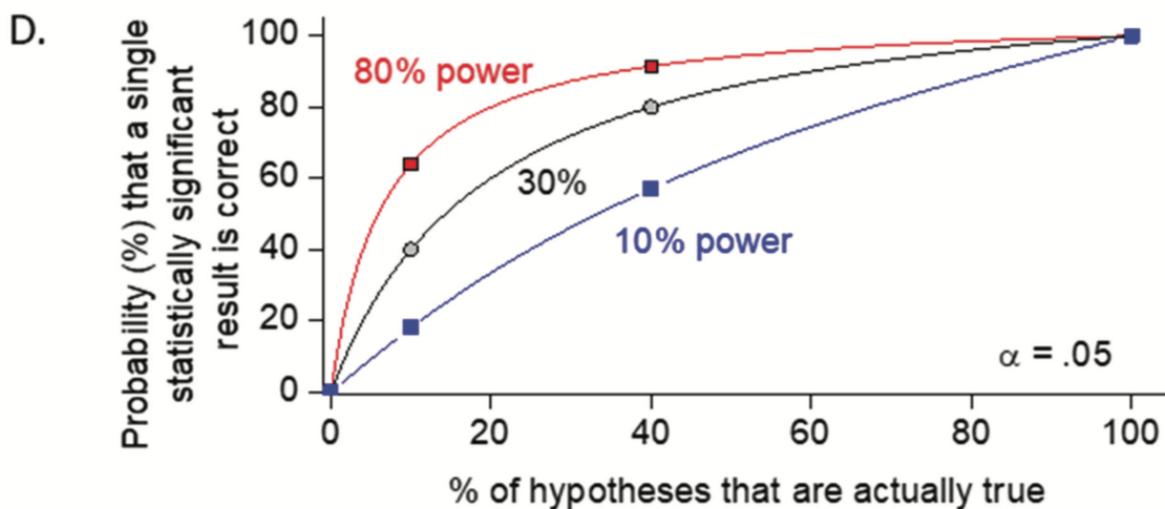
**B.**
1000 hypotheses
10% are true
α = 0.05
Power = **30%**

|  | True + | True − |
|---|---|---|
| | 30 | 855 |
| | 45 | 70 |
|  | False + | False − |

**40%**

**C.**
1000 hypotheses
**40%** are true
α = 0.05
Power = 30%

|  | True + | True − |
|---|---|---|
| | 120 | 570 |
| | 30 | 280 |
|  | False + | False − |

**80%**

**D.**

Y-axis: Probability (%) that a single statistically significant result is correct
X-axis: % of hypotheses that are actually true

80% power
30%
10% power
α = .05

Now, let's pretend that you just performed an experiment and got a positive result. What is the likelihood that it is actually true? It's not 95%, as many believe. No, it is 80/125 = 64%. That's no good at all. If statistical power is even lower than 80% (which is often the case in many studies), this problem gets even worse (see Box B).

The good news, according to Dr. Dingledine? With a little discipline, all of the human flaws he identified can be fixed. That is, of course, assuming that all of the economic and psychology data he cited can itself be reproduced.

Source [2]: Ray Dingledine. "Why Is It So Hard To Do Good Science?" *eNeuro.* Published: 4 September 2018. DOI: 10.1523/ENEURO.0188-18.2018

**Source URL:** https://www.acsh.org/news/2018/09/06/why-good-science-hard-do-13385

**Links**

[1] https://www.storyblocks.com/stock-image/science-concept-hmzlo8-px_wj6gsa3lx

[2] http://www.eneuro.org/content/early/2018/09/04/ENEURO.0188-18.2018