

WARNING SIGNS

SOME INDICATIONS THAT THE RESEARCH MAY NOT BE OF THE HIGHEST QUALITY

BY BARBARA A. DRESCHER

USE OF THE WORD "PROVE"

SCIENCE IS NEVER CERTAIN

Science accepts a few "first principles" on which all arguments rest. If any one of these first principles is shown to be false, then everything we know is open to question. Therefore, science accepts that proof is not possible. Science is open-minded.

That said, some respected scientists have used this word in the popular press. I suspect that they mean "strong evidence", however, I believe this practice misleads the public and fuels common misunderstandings of science.

A TITLE THAT DESCRIBES THE IMPOSSIBLE

"THE EFFECTS OF..."

"THE INFLUENCE OF..."

"THE ROLE OF..."

Followed by ANYTHING that cannot be randomly assigned to subjects, such as:

- Gender
- Birth Order
- Race/Ethnicity/Religious Affiliation
- Intelligence
- Height
- Favorite Color... you get the idea

All of these statements imply causal relationships and nothing that cannot be randomly assigned can be said to *cause* anything (in a single experiment or study). The acceptance of causal relationships among variables that cannot be studied experimentally (meaning a "true" experiment) requires extraordinary converging evidence.

CORRELATION MISTAKEN FOR CAUSE

Like the titles that imply cause, the language used in conclusions can be misleading. This is particularly problematic in secondary sources such as the popular media.

REDEFINING TERMS

"IT DEPENDS ON WHAT YOUR DEFINITION OF 'IS' IS..."

Studies that claim relationships among variables, but define those variables in vague or unique terms should be scrutinized. You may claim that your diet results in weight loss by adding up all of the pounds people in sample have lost in a given time period, but your definition is a deliberate misrepresentation of weight loss if you do not include the weight people gain.

MISSING COMPARISONS

"EATING BRAND X REDUCES CHOLESTEROL..."

Compared to what? EVERYTHING in the world is relative. *Effects* are differences in measures among levels of a variable. A variable, by definition, is *something that varies*. A study may find that eating Brand X may result in lower cholesterol than eating Brand Y (given otherwise identical conditions), but no single value of a variable can *do* anything or *affect* anything.

When comparisons are missing in the communication of findings, the claim cannot be evaluated. For example, it is entirely possible that eating Brand Y *raises* cholesterol.

PRETEST-POSTTEST FINDINGS

One-group, pretest-posttest comparisons tell us absolutely nothing. For example, findings such as:

- People with headaches are given aspirin and experience pain relief within 1 hour.
- People enrolled in my weight-loss clinic lost an average of 10 pounds in 4 weeks.
- Walls are cleaner after being washed with "SuperClean".

Can easily be explained a number of ways, including:

- Headaches usually last about 45 minutes.
- People who are publicly weighed tend to lose weight to avoid embarrassment.
- Walls washed with anything are cleaner than they were before being washed.

USE OF THE WORD *SIGNIFICANCE* IN FINDINGS

OR MISUSE OF THE WORD *SIGNIFICANT*

This term has lost its meaning through repeated misuse.

- We do not "find significance" or seek it.

- Effects are not "significant" or "non-significant". By definition, if there is an effect, a statistically-significant result was found. If no significant result was found, there is no "effect".
- Statistical significance does not mean "significant" findings. When a researcher says "Group A had significantly higher scores than Group B," they mean that the difference in scores was probably not due to chance. We cannot conclude that the treatment results in "significant" differences in the dependent measure.

THE PURPOSE IS TO CONFIRM

A REPORT THAT STARTS WITH, "THE PURPOSE OF THIS STUDY IS TO FIND SUPPORT FOR THE HYPOTHESIS THAT..."

Scientists do not seek support for what they believe. Scientists seek the truth.

CONCLUSIONS THAT DEFY FINDINGS

"THE DIFFERENCE BETWEEN GROUP A AND GROUP B WAS NOT SIGNIFICANT, HOWEVER, THERE WAS A SLIGHTLY HIGHER MEAN..."

In hypothesis testing, there are no "trends toward significance". Lines are drawn at acceptable levels of error and moving those lines to suit one's wishes defies the purpose of the scientific method – to remove human biases.

THE BOTTOM LINE: If the differences are not statistically significant, then you must conclude that there are no differences in the population, or you must discuss the methodological flaws or omissions that resulted in the null findings. You can NEVER, however, reject the null hypothesis based on what you did NOT find.

SUSPICIOUS HYPOTHESES

HYPOTHESES WITHOUT A THEORETICAL FOUNDATION

Sometimes hypotheses seem to spring from nowhere, either somewhat unrelated to the statements made in the introduction, or even contradicting them. At times, the logic in the introduction is a bit twisted, or speculative, or requires a number of assumptions to go from the known literature to the hypothesis to be tested. This is a sign that the introduction was actually written *after* the study was completed.

Although it is a common practice to *write* after completion of the study, changing one's hypothesis to match findings is not. In fact, it is a serious violation of ethical principles.

Researchers using proper scientific method have a priori hypotheses that are based on what is known (findings of other studies, mostly). Whether this information is written before or after the study is conducted is not the issue.

HYPOTHESES THAT DO NOT MATCH THE METHOD

Sometimes the method of a study does not seem to directly test the hypotheses. This is another indication that the hypotheses themselves have been changed to suit the findings, making it appear as though the conclusions are more meaningful than they actually are. Studies often produce interesting findings that are not directly related to the original hypotheses. These should be discussed, but should never be over-analyzed using the same data, nor should they be allowed to replace the original hypothesis.

We may have good explanations for what we have found, but that does not mean our explanations are correct (or even close). Instead of reconstructing a study to fit the findings, new hypotheses should be tested in future studies.

STUDIES THAT ARE OVER-COMPLICATED

TOO MUCH ANALYSIS, TOO MANY FACTORS, MEDIATION & MODERATION

The canon of *parsimony* dictates that explanations that require the fewest assumptions are the most likely to be correct. The more assumptions that are required, the greater the probability that an error has been made. The more statistical tests that are conducted, the greater the probability of error. Studies can be overcomplicated in their procedure, analysis, or even hypotheses. Hypotheses regarding mediating or moderating variables must be tested with care.

CITATIONS THAT REFER TO WHAT SOMEONE SAID RATHER THAN WHAT SOMEONE *FOUND*

HEARSAY OR OPINION...

- Citing statements from the introductions of a research report usually involves hearsay. Joe said that Bart said... Only the findings of studies are evidence that can be evaluated.
- Citations of review articles, textbooks, or popular press are always filtered information, either in the form of opinion or, again, hearsay.

INCORRECT USE OF STATISTICS

DRAWING CONCLUSIONS ACROSS TESTS

A common mistake is drawing conclusions without direct comparisons. For example, a group of children were given a pretest on verbal ability, then received either "phonics" or "traditional"

teaching in language arts. They were tested again after 4 weeks of classes. Separate tests showed that the scores of students in the "phonics" condition improved, but the scores of those in the "traditional" condition did not. The researchers concluded that "phonics" was more effective than "traditional" teaching. This conclusion does not logically follow, since the scores of those in "phonics" were never compared to those in "traditional" teaching methods.

CAUSAL MODELING DOES NOT PERMIT CAUSAL CONCLUSIONS

This kind of analysis has become a popular means of analyzing data to show complicated relationships among variables. It is often mistaken, however, for a means to overcome the problem of random assignment. Causal models are no different from other models. Finding a model that fits data does not mean a causal relationship exists. Correlation, which is required for cause, is sufficient for a good fit. Only experimental methods can produce valid causal conclusions.