

Journalist's Resource

Research on today's news topics

Eight questions to ask when interpreting academic studies: A primer for media

Reading [scholarly studies](#) can help journalists integrate rigorous, unbiased sources of information into their reporting. These studies are typically carried out by professors and professional researchers -- at universities, think tanks and government institutions -- and are published through a peer-review process in which those familiar with the study area ensure that there are no major flaws.

Even for people who carry out research, however, interpreting scientific (and social science) studies and making judgments about their quality can be difficult tasks. In a now-famous article, Stanford professor John Ioannidis argues that “[most published research findings are false](#)” due to inherent limitations in how researchers design studies. (Health and medical studies can be particularly attractive to media, but be aware that there is a [long history of faulty findings](#).) Occasionally, too, studies can be the product of outright fraud: A [1998 study](#) falsely linking vaccines and autism is now perhaps the canonical example, as it spurred widespread and [long-lasting societal damage](#). Journalists should also always examine the funding sources behind the study, which are frequently declared at the study's conclusion.

Before journalists write about research and speak with authors, they should be able to both interpret a study's results generally and understand the appropriate degree of skepticism that a given study's findings warrant. This requires [data literacy](#), some familiarity with [statistical terms](#) and a basic knowledge of [hypothesis testing and construction of theories](#).

Journalists should also be well aware that most academic research contains careful qualifications about findings. The common complaint from scientists and social scientists is that news media tend to pump up findings and hype studies through catchy headlines, distorting public understanding. But landmark studies sometimes do no more than [tighten the margin of error](#) around a given measurement -- not inherently flashy, but intriguing to an audience if explained with rich context and clear presentation.

On the following pages are some important questions to ask when reading a scientific study.

1. What are the researchers' hypotheses?

A hypothesis is a research question that a study seeks to answer. Sometimes researchers state their hypotheses explicitly, but more often their research questions are implicit. Hypotheses are testable assertions usually involving the relationship between two variables. In a study of smoking and lung cancer, the hypothesis might be that smokers develop lung cancer at a higher rate than non-smokers over a five-year period.

It is also important to note that there are formal definitions of [null and alternative hypotheses](#) for use with statistical analysis.

2. What are the independent and dependent variables?

Independent variables are factors that influence particular outcomes. Dependent variables are measures of the outcomes themselves. In the study assessing the relationship between smoking and lung cancer, smoking is the independent variable because the researcher assumes it predicts lung cancer, the dependent variable. (Some fields use related terms such as “exposure” and “outcome.”)

Pay particular attention to how the researchers define all of the variables -- there can be quite a bit of nuance in the definitions. Also look at the methods by which the researchers measure the variables. Generally speaking, a variable measured using a subject's response to a survey question is less trustworthy than one measured through more objective means -- reviewing laboratory findings in their medical records, for example.

3. What is the unit of analysis?

For most studies involving human subjects, the individual person is the unit of analysis. However, studies are sometimes interested in a different level of analysis that makes comparisons between classrooms, hospitals, schools or states, for example, rather than between individuals.

4. How well does the study design address causation?

Most studies identify correlations or associations between variables, but typically the ultimate goal is to [determine causation](#). Certain study designs are more useful than others for the purpose of determining causation.

At the most basic level, studies can be placed into one of two categories: [experimental and observational](#). In experimental studies, the researchers decide who is exposed to the independent variable and who is not. In observational studies, the researchers do not have any control over who is exposed to the independent variable -- instead they make comparisons between groups that are already different from one another. In nearly all cases, experimental studies provide stronger evidence than observational studies.

Here are descriptions of some of the most common study designs, presented along with their respective values for inferring causation:

- **Randomized controlled trials** (RCTs), also known as clinical trials, are experimental studies that are considered the “gold standard” in research. Out of all study designs, they have the most value for determining causation although they do have limitations. In an RCT, researchers randomly divide subjects into at least two groups: One that receives a treatment, and the other -- the control group -- that receives either no treatment or a simulated version of the treatment called a **placebo**. The independent variable in these experiments is whether or not the subject receives the real treatment. Ideally an RCT should be **double-blind** -- the participants should not know to which treatment group they have been assigned, nor should the study staff know. This arrangement helps to avoid bias. Researchers commonly use RCTs to meet regulatory requirements, such as evaluating pharmaceuticals for the Food and Drug Administration. Due to issues of cost, logistics and ethics, RCTs are fairly uncommon for other purposes. **Example:** “[Short-Term Soy Isoflavone Intervention in Patients with Localized Prostate Cancer](#)”
- **Longitudinal studies**, like RCTs, follow the same subjects over a given time period. Unlike in RCTs, they are observational. Researchers do not assign the independent variable in longitudinal studies -- they instead observe what happens in the real world. A longitudinal study might compare the risk for heart disease among one group of people who are exposed to high levels of air pollution to the risk of heart disease among another group exposed to low levels of air pollution. The problem is that, because there is no random assignment, the groups may differ from one another in other important ways and, as a result, we cannot completely isolate the effects of air pollution. These differences result in [confounding](#) and other forms of bias. For that reason, longitudinal studies have less validity for inferring causation than RCTs and other experimental study designs. Longitudinal studies have more validity than other kinds of observational studies, however. **Example:** “[Mood after Moderate and Severe Traumatic Brain Injury: A Prospective Cohort Study](#)”
- **Case-control studies** are technically a type of longitudinal study, but they are unique enough to discuss separately. Common in public health and medical research, case-control studies begin with a group of people who have already developed a particular disease and compare them to a similar but disease-free group recruited by the researchers. These studies are more likely to suffer from bias than other longitudinal studies for two reasons. First, they are always **retrospective**, meaning they collect data about independent variables years after the exposures of interest occurred -- sometimes even after the subject has died. Second, the group of disease-free people is very likely to differ from the group that developed the disease, creating a substantial risk for confounding. **Example:** “[Risk Factors for Preeclampsia in Women from Colombia](#)”.

- **Cross-sectional studies** are a kind of observational study that measure both dependent and independent variables at a single point in time. Although researchers may administer the same cross-sectional survey every few years, they do not follow the same subjects over time. An important part of determining causation is establishing that the independent variable occurred for a given subject before the dependent variable occurred. But because they do not measure the variables over time, cross-sectional studies cannot determine that a hypothesized cause precedes its effect, so the design is limited to making inferences about correlations rather than causation. **Example:** “[Physical Predictors of Cognitive Performance in Healthy Older Adults](#)”
- **Ecological studies** are observational studies that are similar to cross-sectional studies except that they measure at least one variable on the group-level rather than the subject-level. For example, an ecological study may look at the relationship between individuals’ meat consumption and their incidence of colon cancer. But rather than using individual-level data, the study relies on national cancer rates and national averages for meat consumption. While it might seem that higher meat consumption is linked to a higher risk of cancer, there is no way to know if the individuals eating more meat within a country are the same people who are more likely to develop cancer. This means that ecological studies are not only inadequate for inferring causation, they are also inadequate for establishing a correlation. As a consequence, they should be regarded with strong skepticism. **Example:** “[A Multi-country Ecological Study of Cancer Incidence Rates in 2008 with Respect to Various Risk-Modifying Factors](#)”
- **Systematic reviews** are surveys of existing studies on a given topic. Investigators specify inclusion and exclusion criteria to weed out studies that are either irrelevant to their research question or poorly designed. Using keywords, they systematically search research databases, present the findings of the studies they include and draw conclusions based on their consideration of the findings. Assuming that the review includes only well-designed studies, systematic reviews are more useful for inferring causation than any single well-designed study. **Example:** “[Enablers and Barriers to Large-Scale Uptake of Improved Solid Fuel Stoves.](#)” For a sense of how systematic reviews are interpreted and used by researchers in the field, see “[How to Read a Systematic Review and Meta-analysis and Apply the Results to Patient Care.](#)” published in the *Journal of the American Medical Association* (JAMA.)
- **Meta-analyses** are similar to systematic reviews but use the original data from all included studies to create a new analysis. As a result, a meta-analysis is able to draw conclusions that are more meaningful than a systematic review. Again, a meta-analysis is more useful for inferring causation than any single study, assuming that all studies are well-designed. **Example:** “[Occupational Exposure to Asbestos and Ovarian Cancer](#)”

5. What are the study's results?

There are several aspects involved in understanding a study's results:

- Understand whether or not the study found [statistically significant](#) relationships between the dependent and independent variables. If the relationship is statistically significant, it means that any difference observed between groups is unlikely to be due to random chance. [P-values](#) help researchers to decide whether observed differences are simply due to chance or represent a true difference between groups.
- If the relationship is statistically significant, it is then important to determine the **effect size**, which is the size of the difference observed between the groups. Subjects enrolled in a weight loss program may have experienced a statistically significant reduction in weight compared to those in a control group, but is that difference one ounce, one pound or ten pounds? There are myriad ways in which studies present effect sizes -- such obscure terms as regression coefficients, odds ratios, and population attributable fractions may come into play. Unfortunately, research articles sometimes fail to interpret effect sizes in words. In these cases, it may be best to consult an expert to help develop a plain-English interpretation.
- Even if there is a statistically significant difference between comparison groups, this does not mean the effect size is meaningful. A weight loss program that leads to a total weight reduction of one ounce on average or a policy that saves one life out of a billion may not be meaningful. Again, consulting an expert in the field can help to determine how meaningful an effect size is, a determination that is ultimately a subjective judgment call.

6. How generalizable are the results?

Study results are useful because they help us make inferences about the relationship between independent and dependent variables among a larger population. The subjects enrolled in the study must be similar to those in the larger population, however, in order to generalize the findings. Even a perfectly designed study may be of limited value when its results cannot be generalized. It is important to pay attention to the composition of the study sample. If the unit of analysis is the individual, important factors to consider regarding the group's composition include age, race/ethnicity, gender, socioeconomic status, and geographic location. While some samples are deliberately constructed to be representative of a country or region, most are not.

7. What limitations do the authors note?

Within a research article, authors often state some of the study's limitations explicitly. This information can be very helpful in determining the strength of the evidence presented in the study.

8. What conclusions do similar studies draw?

With some notable exceptions, a single study is unlikely to fundamentally change what is already known about the research question it addresses. It is important to compare a new study's findings to existing studies that address similar research questions, particularly systematic reviews or meta-analyses if available.

Further: One hidden form of bias that is easily missed is what's called "selecting on the dependent variable," which is the research practice of focusing on only those areas where there are effects and ignoring ones where there are not. This can lead to exaggerated conclusions (and thereby false media narratives). For example, it is tempting to say that "science has become polarized," as survey data suggest significant differences in public opinion on issues such as climate change, vaccinations and nuclear power. However, on most scientific issues, there is [almost no public debate or controversy](#). Additionally, the reality of "[publication bias](#)" -- academic journals have traditionally been more interested in publishing studies that show effects, rather than no effects -- can create a biased incentive structure that distorts larger truths.

For an updated overview, see a 2014 paper by Stanford's John Ioannidis, "[How to Make More Published Research True.](#)"

By Justin Feldman and John Wihbey, May 26, 2015.