Statistical vs. Practical Significance

Wow, it's statistically significant!

But is it significant? Or even better, is it meaningful?

We often hear the question: "How many responses do I need to make my study 'statistically significant'?" Hearing the words "statistically significant" is always confusing, and the usual response is often "What exactly do you mean by that?"

So what's the confusion?

- A survey or study by itself can't be "statistically significant," nor can a certain number of responses/participants. Only a test statistic (e.g., a calculated statistical quantity) can be statistically significant.
- To a statistician "statistically significant" has a very particular meaning related to hypothesis testing requiring a specific set of assumptions which are not always true in an educational setting.
- Researchers aren't always asking about the significance of the results in the context of their objectives. This should be their main focus.

Let's take a closer look:

For example, suppose a college transition program is developing a survey to assess student preferences for certain types of postsecondary options (i.e. community colleges, public senior colleges, private institutions, military service, full-time work, etc.). Often the first question programs ask is how many responses are needed to get "statistically significant" results. That is where the confusion starts; that question only makes sense in the context of a statistical hypothesis test. A survey may involve many hypotheses that we want to test.

A statistical hypothesis test requires both a hypothesis: women enroll in community colleges more than men, and a test statistic: the percent of women who enroll in community colleges minus the percent of men who enroll in community colleges. Now we can ask if the test statistic

(the difference between the two percentages) is “statistically significant.” That’s a legitimate question. A more meaningful question might be whether the difference is “practically significant.”

Practical significance: A calculated difference is practically significant if the actual difference it is estimating will ***affect a decision to be made***. (Should the program focus more community college-based services toward women than men?)

Statistical significance depends on the sample size. A difference of 3% (58% for women minus 55% for men) can be statistically significant if the sample size is big enough, but it may not be practically significant. 3% hardly seems big enough to warrant focusing on one group of students over the other.

A difference of 30% (65% for women minus 35% for men) may be practically significant (i.e., warrant a decision to focus more resources in one direction) but if the difference isn’t statistically significant (that depends on sample size) then you can’t be sure the difference you see (30%) is real, so you either need to get more data or treat the two groups as the same.

In short, statistical significance is mathematical - it comes from the data (sample size) and from your confidence (how confident you want to be in your results). Practical significance is more subjective and is based on other factors like cost, requirements, program goals, etc.