

PSY 102 – What do the Research Results Mean?

Researchers generally consider their findings noteworthy only if they are *statistically significant*. The term “significant” usually means “important,” but in this case the term does not have to do with how important or large a finding is. It’s unfortunate that researchers a long time ago decided to use that term, because it’s misleading. All it means is that it is statistically unlikely that the researcher would have found the correlation or the difference between two groups that she did if in fact there the correlation was .00 in the population, or if the two groups did not differ at all in the population.

To be worth reporting in research journals, results have to be statistically significant. However, that does not mean that the results are large or important. The term ***practical significance*** refers to the actual size of the difference between groups, or the correlation between variables.

What we are concerned with here is the degree to which particular research findings help us better understand individuals. For example, two or more groups may show statistically significant group differences on some variable, yet those differences might be very small in absolute terms such that they have little practical meaning. In other words, the research results may be so small that they do not help us better understand what happens in our own lives.

Unfortunately, when research results are presented in a way that entails describing one group as having more of a certain experience than a second group, it is easy to assume that the members of the first group *typically* or always differ from members of the second group. This may indeed have been the case. However, a statistically significant difference between the groups may have resulted from a small subset of people in the first group who reported extreme amounts of sexual experience. In this case, the *typical* member of each group may have given very similar responses to the questions researchers posed, yet the groups differ in the *average* response because of those relatively few respondents who gave very atypical reports (and hence inflated their group’s average). Describing the members of the groups as being different in general would be misleading.

Misleading Statistics and Characterizations

It is possible that, regardless of the size of the difference or relationship found, the way researchers describe or visually present their findings implies greater relationships or differences than were actually found. I am not saying that researchers intentionally attempt to deceive readers, but rather there are times when the ways research results are reported can lead to misrepresentation of the findings. Suppose you read that in a particular sample of college students men were more likely than women to agree with the statement, “If a woman gets raped, she gets what she deserves.” This description portrays men as relatively callous at best and perhaps dangerous at worst.

Now, suppose that the respondents rated their degree of agreement with the statement using a seven-point scale where 1 corresponded to *Strongly Disagree*, 4 corresponded to *Neither Agree nor Disagree*, and 7 corresponded to *Strongly Agree*. This is a common method for measuring attitudes in sexuality research (and social science research in general). Further suppose that the researcher found that the average rating for women was 1.78 and the average rating for men

was 2.56. This may have been a statistically significant difference, yet to say that men were in greater *agreement* with the statement is misleading because it is clear that *both men and women* generally indicated disagreement with the statement. Although it may be more accurate to say that men were less likely to disagree with the statement, the findings may not be presented this way. Perhaps there is a somewhat natural tendency for people to make sense out of findings in terms of a "positive" direction (which group displayed more of something, in this case agreement) rather than a "negative" direction.

Percentages too can be misleading. As an example, consider the results from analyses performed by Wortley and Fleming (1997). These researchers examined the prevalence of AIDS in men and women in the United States, particularly from the period 1991 to 1995. Based on their findings, Wortley and Fleming (1997) concluded that "The incidence of AIDS has increased more in women than men in recent years" (p. 914). If all one read was the conclusion, one might be led to believe that AIDS is more prevalent in women than in men, or quickly becoming so. Examining their data more closely, however, reveals a potential problem with percentages.

It is true that, from 1991 to 1995, new cases of AIDS among men increased about 11% compared to an increase of about 60% among women (see Wortley & Flemming, 1997, Figure 1). Even these percentages imply more of an epidemic among women than among men. However, in terms of *absolute number* of new cases of AIDS in 1995, there were about 50,000 cases among men compared to about 13,000 cases among women. Clearly, most cases of AIDS were still occurring among men (in 1991 it was about 45,000 new cases among men compared to about 8,000 cases among women). It was accurate to say that the percentage increase from 8,000 cases to 11,000 cases among women is larger than the percentage increase from 45,000 cases to 50,000 cases among men, but without considering the absolute number of cases, the overall conclusion regarding the percentages may mislead the reader.

Even these numbers do not tell the whole story for women. While it was true that new cases of AIDS were occurring at a faster *rate* among women than among men from 1991 to 1995, the incidence of AIDS was not equally distributed among women in general. A new case of AIDS in the United States in 1995 occurred among every 33,300 women of European descent, every 62,500 Asian-American women, every 4,200 Latina women, and every 2,000 African-American women (see Wortley & Flemming, 1997, Table 1). Obviously, to make generalizations about "American women," as though they were a homogeneous group, would also lead to misrepresentation of the data.

A similar way percentages may be interpreted in a way that is confusing has to do with describing the difference between two groups in relative rather than absolute terms. In other words, the conclusion that "Members Group A was four times more likely than members of Group B to report having a miserable marriage" implies a substantial difference between the two groups. However, the absolute percentages may have been 8% in Group A compared to 2% in Group B. Similarly, we may read that the incidence of some behavior or disease or experience has "doubled" (100% increase) in the past year or decade or whatever. This conclusion may be based on the finding that the incidence went from one out of every 10,000 people to two out of every 10,000 people. The increase sounds more dramatic when presented as a percentage compare to the absolute numbers. If a writer does not present the actual numbers, but rather provided only percentages, we should be skeptical as to the magnitude of

the problem.

Do the Conclusions Match the Research Design?

Last, are the conclusions the authors draw from their findings, or the ways in which they characterize their findings, legitimate given the nature of the research? In asking this question, we have come full circle from considering the research design and the inherent limitations involved with each type of research. Now we are considering whether, given those limitations, the conclusions seem warranted.

As an example, suppose researchers claim that the results of their research show that living together before marriage results in increased likelihood of divorce within five years after marriage. The researchers are basing this conclusion on the finding that a greater proportion of people who lived together prior to marriage divorced within five years when compared to respondents who had not lived together. Although the researchers found such a relationship, the conclusion implies that there is something about premarital cohabitation that *causes* an increased risk for divorce.

Recall, however, that correlational research does *not* allow us to draw conclusions about causality. In this case, there may be some other variable or set of variables that is related to *both* cohabitation and divorce that better explains their apparent relationship. For example, people who lived together before marriage may be less religious and hold more liberal attitudes regarding relationships, sexuality, and family compared to those couples who did not live together. People who are less religious and hold liberal attitudes about relationships may also be more open to the possibility of divorce when things are not going well in the marriage. These group differences in religiosity and attitudes (which existed before cohabitation) may best explain why those who chose premarital cohabitation were also more likely to divorce.

Rosenthal (1994) referred to researchers' tendencies to imply causal relationships between their variables as "causism." He noted that writers may not come out and say that their correlational results indicate a causal relationship, yet they may describe their findings using such words as *effect, impact, consequence, or the result of*. In using such words when describing relationships between two variables, the implication is that one variable caused the other or at least influenced it. Such words distract the reader from the important point that the results are simply correlational, and that all we can say for sure is that the two variables demonstrate a statistical relationship (and perhaps a weak, yet statistically significant, one).