ADVANTAGES OF FACTORIAL DESIGNS

Continuing the theme of our numerical example, suppose that we were interested in the effectiveness of various treatments for hypertension. Specifically, we wonder whether biofeedback reduces blood pressure, and we are also interested in comparing Drugs *X, Y*, and *Z*. Would it be better to conduct a 2×3 factorial study or to perform two separate single-factor studies?

The most obvious advantage of the factorial design is that it enables us to test the existence of an interaction. With two single-factor studies, we could never learn that differences between drugs might differ depending on the presence or absence of biofeedback. In particular, stop and think about the single-factor study to compare the three drugs. This study most likely would be conducted in the absence of biofeedback, so in effect we are performing what would be a simple effects test in the factorial design. However, the simple effects test may convey only one-half of the relevant ways in which drugs differ. In summary, it may be of theoretical interest to discover an interaction, which implies the necessity of a factorial design.

What if an interaction is not expected? Is there still any advantage to the factorial design? Yes, because the factorial design enables greater generalizability. If our hypertension study is conducted as a factorial design and there is no interaction, we can conclude that drug differences are the same in the presence of biofeedback as in its absence. (However, an alternate explanation, especially with small sample sizes, is that the power to detect an interaction was inadequate. Even with sufficient power, one should really conclude that any differences that may exist are so small that we can regard them as non-existent for all practical purposes because we should not literally accept the null hypothesis.) In other words, we can generalize drug effects across two levels of the biofeedback factor. If we had instead conducted a single-factor study, we could not assess the extent of generalizability.

So far we have seen that a factorial design may be preferable to a series of single-factor studies because we can test interaction effects and we can assess generalizability (notice that these two advantages are really opposite perspectives on one advantage). However, don't factorial designs require larger sample sizes? Let's consider two hypothetical psychologists: Dr. Single and Dr. Multiple. Dr. Single decides to conduct two single-factor studies. The first study investigates the relative effectiveness of Drugs *X, Y*, and *Z*. Thirty subjects are assigned at random to each of the three drugs. In the second study, biofeedback is compared to a control. Forty-five individuals are assigned at random to each of the two groups. In the two studies combined, Dr. Single has used 180 participants. Dr. Multiple conducts a 2×3 factorial study investigating the effect of biofeedback and drug effects simultaneously. Fifteen individuals are assigned at random to each of the six groups. Of course, Dr. Multiple can test an interaction that Dr. Single cannot, but how else will their tests be different? Both will test whether biofeedback has an effect. Dr. Single's comparison involves 45 individuals in each group. But so does Dr. Multiple's, because there were 15 individuals at each level of drug, implying that 45 individuals received biofeedback, whereas 45 others did not. Both investigators will also test for drug differences. By the same logic, both Dr. Single and Dr. Multiple will have exposed 30 individuals to each type of drug. Thus, it should be the case that Dr. Multiple's statistical power for assessing biofeedback and drug effects should be equivalent to Dr. Single's. Does this mean that Dr. Single's and Dr. Multiple's approaches are equally good in how efficiently participants are used? Recall that Dr. Single used 180 individuals in all. However, Dr. Multiple used a total of $6 \times 15 = 90$ participants. Dr. Multiple's factorial design produced the same power with half as many subjects as Dr. Single's two separate studies! The implication is that the factorial design uses participants more efficiently than would a series of single-factor studies.[9]