CHAPTER 8

Strategies for Using Statistics

There is a life behind repeated measurement ANOVA!
—Dr. Daniel Stahl, Senior Lecturer, King's College London

By the time you get to this section of the guide, you should not be asking which statistics to apply to your data. You should know the answer, which should be defined in your research proposal, driven by your key research question. It is a good idea to enter the data for your study and keep current with the coding as you test the infants or children. By keeping current, not only will you be able to assess whether you are getting positive results but also you may detect problems that you did not anticipate. When computing statistics, first get an overall sense of your data.

You do not want to send an email to your advisor that reads "Dear Professor, there is (or is not) a significant difference—here is the ANOVA." Before you compute the relevant statistics, you should examine such descriptive statistics as the means. Take a careful look at both your highest and lowest scores. Do these scores fall within reasonable parameters? For example, did one child gaze at an object for 20 seconds while another took only 2 seconds with the same object?

Always double-check and then triple-check your data before you even begin thinking about running statistics. Look at your descriptive statistics with your hypothesis in mind. Does everything make sense? Were any variables completely inverted in terms of their expected influence? Always play devil's advocate with your research peers and even with in your deliberations. If you are computing statistics, do your *F* values correspond with the number of participants in your study manuscript? Does every statistical result make sense? Just because a statistic is generated by a computer, do not assume that it is necessarily correct.

Step back!

What variables are you attempting to assess in your analysis? Does the output of your SPSS/data file make sense? If you are comparing how long infants look at a novel object versus a familiar object and whether there is a significant difference between these two time periods, how many *p* values would you expect? If you expect three *p* values, what analysis are they for, and why? If you are assessing whether there is a significant difference between two groups' numbers, you would expect one *p* value. Step back, question every number you see, and try to comprehend the logic implied by your statistical outputs. Do not send your colleagues random batches of SPSS outputs and Excel files that are not labeled. Rather, date all of your files in the top-right corner, and describe what each analysis means. **A file labeled simply "Output246" can be very dangerous for a variety of reasons.**

TIP: If you happen to have computed statistics for practice or because you were "just curious," be sure to delete these from your computer.

"Do I understand which statistics I should compute—and why?"

"Did I date and clearly label my files/output?"

"Did I compute all relevant descriptive statistics?"

"Did I search out any outliers or inconsistencies?"

"Did I plot a sample graph just to get a feel for the data?"

"Did the inferential statistics I computed fit the patterns suggested by the descriptive statistics and graphs?"

Look for the basic insights implied by your data. If you do compute a significant ANOVA and find, for example, that your children looked significantly longer at the red square than the blue one, specifically how many children followed this tendency? You may end up publishing a research paper based on these data, and any perceived shortcomings could conceivably damage your reputation.

Always be critical. Was anything else going on that might have affected the overall validity of your study results? For instance, you would want to assure in advance that Observer 1 did not test all 5-year-olds and Observer 2 all 10-year-olds. Otherwise, if your data end up indicating an age effect, it might arguably be attributable to differences in the tester. You also should keep the testers and the parents (and participants) uninformed as to your experimental hypotheses so as to minimize bias in their administration of the tests. You should try to assure that parents not influence their children's behavior unduly. One way to do this is by having them wear opaque glasses or simply observe their children through a one-way mirror. You should always strive for generalizability of your results whenever possible. For instance, if your sample is solely children at an elite prep school that costs $65K a year, your results would not likely be generalizable to all children. Similarly, external validity often becomes an issue when testing young infants. At times, some 50% of infants fail to make it through a study for one reason or another (fussiness, excessive movement, experimental error, etc.). In this case, you might want to assess the research literature if possible and explain in your paper that the experienced dropout rate was typical for this type of study (if indeed it was).

O versus 0

Be sure to label and code variables carefully in your files so that there is no ambiguity in understanding them. Be consistent with the terms that you select. Remember that someone may seek to examine the data file that you developed some 20 years from now. I once had a student distinguish infants in two conditions by placing a triangle or a circle next to each infant's name. Not only did the failure to formally define the symbols mean that no one would know what they meant 10 years later (or for *ever*, for that matter), but moreover triangles and circles look all too similar (as opposed to *X*s and an *O*). There was just too much room for potential error. The lesson is to always write and mark clearly and consistently on your coding sheets and logs, as it is important not to leave any room for ambiguity. Also, do not assume that someone will know that 0 means "no" and 1 means "yes" in your data files. Display clear descriptions for any notations used so that anyone will know immediately what a given code means.

What statistic should I compute?

To answer this question, return to your proposal and look up "What is my question?" There are great sources available for those seeking to learn about the bases as well as current trends in statistics. To be on the safe side, you should also consider working closely with and publishing with a professionally trained statistician. Statisticians often bring expertise to bear that developmental scientists lack. They keep up with the latest methods and trends in the field and can help a great deal in deciding what to do about the small sample sizes or missing data points often encountered when researching children.

TIP: **Statisticians often make for great collaborators. Involve your statistician in your research project from the start and in your publications if possible. Your statistician is not just a number cruncher or someone to run to when SPSS gives you an output that you do not like. Your statistician likely cares deeply about your research question. Plan a strategy to get him or her involved in your research from the start.**

What is the key to working with researchers? What should I remember when working with a statistician?

To address these questions, I asked my collaborator and statistician Dr. Daniel Stahl, Senior Lecturer in Biostatistics at King's College London.

1. Make a plan for how the data will be stored and how data will be analyzed, especially if you have missing data, which is very common for longitudinal studies!
2. Take randomization seriously, and try to do as much blinded as possible.
3. Talk with a statistician *before* you start the study.
4. Don't think that statistics did not advance in the last 50 years!
5. Don't measure everything possible without having a reason for it. You will get too many false positive results.

Any tips on how to plan how many participants will be needed in my study?

Again, from Dr. Stahl:

> *Think of what kind of mean score you expect at the beginning (use perhaps pilot data or published data from similar studies), and define the minimum change you would consider as important. Do a simple power analysis for this change (post–pre) or, if you compare two groups, the difference: (post–pre of group 1) – (post–pre of group 2). You can do a simple power analysis based on a t-test (paired for situation 1 and independent for situation 2) using any sample size calculation software such as G*power or Nquery. Estimate how many subjects you expect to drop out, and add this number to your sample size. In general, you should have at least 25–30 subjects per group so that you do not have to worry too much about distributional assumptions.*

How do I manage missing data points?

Once more, from Dr. Stahl:

> *One way is by using multilevel modeling/mixed modeling. This is now possible to perform in most all software packages including SPSS (mixed model module). Remember to check if the participants who are missing differ in baseline variables and demographics as compared to completers. It will help you to identify if the missing participants have something in common, such as tend to be male or if they have a common social background. It will tell you if you still can generalize your results to the whole population. If a variable predicts missingness, include it in your model as a sensitivity analysis.*

Summary (check off your achievements)

- Develop a suitable plan for storing your data. ______
- Reduce the possibility for error by making all your records clear and straightforward. ______
- Always double-check your data entry with fresh eyes. ______
- Know which statistics you will compute before you begin your study. ______
- Measure only what the question posed by your study requires. ____

- Keep up to date on developments in statistics. ______
- Collaborate with a statistician whenever possible. ______
- Label all your files clearly and consistently. ______

EXERCISES

1. Find eight graphs in the news. Present these graphs to your class or your lab. Would you change these graphs—and, if so, in what ways?

2. Find three research papers. Examine all the figures and graphs used. Are the graphs labeled properly? Would you change these—and, if so, in what ways?

3. Take the average of males = 5, 28, and 52 seconds and females = 22, 17, and 42 seconds. Plot a bar graph of mean seconds, with seconds on the *x* axis.