Results Section

General Advice

Be nice to your audience. Don't throw numbers at your audience. Walk them through your results, holding their hand and explaining everything clearly.

Don't hunt for a significant result. Keep in mind that the statistical power of your study is likely to be quite low. This means that there is a very good chance that your hypothesis will not be supported with a statistically significant *p*-value even if you do everything perfectly. When your central hypothesis is not supported by the data, there is a strong temptation to begin analyzing everything in the hopes of finding at least one significant finding. Please resist this temptation. In the first place, your goal is not to support your hypothesis but to test your hypothesis, and the answer to the test may be "no." There is no shame in not supporting your hypothesis. In fact, that is the most likely outcome when you have low statistical power.

Active voice. APA recommends writing in the active voice rather than the passive voice. The passive voice occurs when the subjects of the sentence are having things done to them rather than the subjects themselves doing something. In the sentence, "The questionnaires were completed by participants," the subject of the sentence is the questionnaire and it is being acted upon by participants. To rewrite this in the active voice, you might write, "Participants completed the questionnaires."

First person. It is preferable to write "We conducted a 2 x 2 ANOVA" rather than "The researchers conducted a 2 x 2 ANOVA." However, this can become awkward when discussing results, in which case you could make the results themselves the subject of the sentence: "The results from a 2 x 2 ANOVA indicated..."

Setting the Stage

Before you get to your central results, there is often some statistical housekeeping you have to take care of. This typically involves clearing up any doubts the reader may have about your measures or manipulations and is often called "setting the stage."

Manipulation checks. A manipulation check is a procedure sometimes used to test whether the levels of your independent variable differ on the variable you intend. For example, if you will be having participants interacting with two confederates, one of whom will be acting high in competence and the other of whom will be acting low in competence, you might have participants complete a "first impressions" questionnaire in which they rate their partners on, among other things, competence. If you can report that participants rate the "high competence" confederate significantly higher in competence than they rate the "low competence" confederate according to a *t*-test, your readers will have more confidence that you have successfully manipulated competence.

Reliability. You need to report the reliability of any dependent variables as well as any changes you made to improve reliability. If you have several dependent variables, it may be more effective to present their reliabilities immediately before you present the statistical tests they were used for, but otherwise you can present reliabilities in this section. If you used reliability analysis to improve reliability by deleting items, you should report the number of items in your original scale, the number of items in your final scale, and the Cronbach alpha of your final scale.

Conversions of raw observations into data. If the responses you collected consist of Likert scale scores, you don't need to worry about this because your raw observations are already in the form they need to be in for analysis. But let's say that you had participants

perform a word-completion task in which you gave them several letters in a word and had them provide a missing letter to create a word: "coff_ _" could be completed as either "coffee" or "coffin." If you are measuring the availability of death-related ideas, you might score "coffin" as a 1 and "coffee" as a 0. That scoring system, or any system you used to convert raw observations into numbers that you will analyze, needs to be carefully described to the reader. It is usually sufficient to provide a couple of examples, give the general rule used to code responses, and say that a complete list of possible responses and the decision rules for coding are available from the authors.

Creation of mean scores. It is very common to combine several responses into a single mean response for purposes of analysis. For example, participants might complete a 10-item measure of anxiety. You need to describe how those ten numbers get turned into a single number. You might write, "A mean anxiety score was computed by averaging participants' responses on the 10-item anxiety measure after reverse-scoring the appropriate items." It is a good idea to *interpret high and low scores*: "High scores on the anxiety measure indicate a high level of anxiety or worry, while low scores indicate little to no anxiety or worry." Finally, choose labels for your variables that are easy to understand. If you are studying the relation between stress and health and all your health items are phrased in terms of frequency of illness, it would be a mistake to use the label "health" for the mean of the illness items, with high "health" scores corresponding to high levels of illness. This would be very confusing for readers. You could reverse-score the illness items, so that high scores correspond to *low* levels of illness, or you could relabel the variable "illness" instead of "health."

Excluded data. There are good and bad reasons for excluding data from analysis. A bad reason is because the data don't support your hypothesis. A good reason is because a data value is implausibly high or low. For example, if you ask participants how much they think a painting would receive at auction and one participant responds "20 billion dollars," which is 20,000 times larger than the next highest estimate, you might want to create a rule that would replace any value above a cutoff level with either a missing value or the cutoff level. If you are collecting reaction time data and typical reaction times are approximately 800 milliseconds, you might consider a rule that excludes reaction times that are more than 3,000 milliseconds because they probably indicate that the participant has become distracted. Whatever the reason for excluding data, you must provide all the rules you used to exclude them. If you excluded *all* the data from one or more participants, you would explain that in the *Participants* subsection of the *Method* section. If you did not exclude any data, you don't need to mention it at all.

The Main Event

After setting the stage, you should present your central results first, then report any other results. Start by restating your prediction: "We expected that greater emotional expressiveness would lead to fewer health problems." Then report the general pattern you observed, referring the reader to any figures or tables: "As Figure 1 shows, this prediction was supported. Participants who showed greater emotional expression were less likely to visit the infirmary." Follow this with any statistical tests, briefly describing what pattern you intended to detect with the test: "The correlation between emotional expressiveness and health problems was $r(52) = -.49$, $p < .01$." Finally, provide an interpretation: "Thus, the prediction was supported: higher levels of emotional expressiveness tended to be found with lower levels of illness."

Figures and Tables. Figures and tables require special formatting that is described in detail in the *APA Publication Manual* and that almost always deviates considerably from the default formatting of programs like SPSS or Microsoft Excel. Every table or figure must be referred to in the text: "As Figure 4 shows..." or "...(see Figure 4)." Typically, it is a good idea to

present your *central* results in a figure, if possible, to provide your audience with a clear, visual way of considering and remembering your findings. If you use a scatterplot, it is usually a good idea to also include a best-fitting line that you describe in the results section. If you present means in a figure, you should also visually represent confidence intervals for those means. The figure caption or table title should capture what the table shows: "Mean smoking rates by extroversion score." If your manuscript is likely to be the final version of your report (as contrasted with a manuscript that is submitted for publication), then you are permitted to place your tables and figures directly into the text of your manuscript at the location where they would be most useful. Otherwise, figures and tables each go on their own page, after References, Appendixes, and Footnotes.

Discussion Section

The Discussion section is where you interpret your findings and place them in context. This section often makes the difference between a good paper and a great paper, so don't skimp on it. Begin with a clear statement about your research question or hypothesis, then present the pattern of your findings and, if you had a hypothesis, tell whether it was supported.

If your hypothesis was supported, discuss the implications of your findings. What have you learned about the concepts you were studying? Are your findings consistent or inconsistent with previous research and theory? Consider any practical applications of your findings: could they be used to change the way...doctors treat patients? ...witnesses are questioned by police? ...advertisements are presented during television programs?

If your hypothesis was not supported, make that clear. Be honest and unapologetic about your findings. If your results are inconsistent with previous research, state that and consider some of the reasons why you may not have found support: low statistical power, different population, different methods, etc. Be specific. You should acknowledge that it is possible that your hypothesis was incorrect, that you did not find a significant effect because the effect does not exist.

Reconnect to ideas in your introduction, especially questions that you may have posed. Have they been answered? If you spent a lot of time discussing a particular issue in the introduction, make sure it has been adequately covered in the discussion. You can improve the stylistic quality of your paper by looking for ways to connect the intro and discussion.

Limitations

An important section of your discussion is a consideration of the limitations of your study. Focus especially on potential limitations of your materials and procedures. If you could do your study over again, what would you change? If you showed participants photographs or film clips, were there any changes that could have made these more impactful? Were there any questionnaire items that you think could have been confusing or misinterpreted by participants? Should you have waited longer before giving participants feedback on their performance? A common limitation of experiments is the sample: it is typically unrepresentative of even a college population, let alone the population of all human adults. When you acknowledge the limitations of your sample, consider how your results might differ with different cultures or other demographic categories (age, gender, ethnicity).

Future Directions

Another important section of the discussion is speculation about what research questions remain and what research might help to answer those questions. You can sometimes integrate the limitations and future directions sections by following a limitation with a suggestion for how the research could be done differently in future studies. Although it is tempting to say

that a particular question would be “interesting” to explore in future studies, it is better not to use the word interesting but instead to describe the question in such a way that readers can see why it would be interesting.

Conclusion

Think about ways to pull together your discussion in the final paragraph. It’s common to restate your major findings (“In summary, we found...”) along with any themes that developed during your paper. Although it can be difficult, it is possible to add a touch of humor or irony to your conclusion. Here’s the last paragraph from Gilbert et al.’s (1998) article on the *durability bias* – the tendency for people to overestimate how long an event will make them feel bad:

> Although we see little evidence of underestimation in either our lives or our laboratories, it is certainly possible that such evidence is simply waiting to be found. For now, we will place a public bet on the predominance of the durability bias, fully confident that should our faith prove misplaced, we will not be embarrassed for long. (p. 636)