Hypothesizing

Hypothesizing Questionable Research Practices (Wicherts et al, 2016, Table 1)

1. Conducting exploratory research without any hypothesis (and later characterizing it as confirmatory).
2. Studying a vague hypothesis that fails to specify the direction of the effect (e.g., a two-sided hypothesis when a one-sided hypothesis is appropriate).

Guidance:

A common problem in psychology is specifying hypotheses in a vague way that makes it easier to engage in questionable research practices (also known as *p*-hacking). Indeed, Lakens (2017) noted that, "statistics teaching should focus less on how to answer questions and more on how to ask questions." Consequently, we encourage you to ask your question in a manner that does not facilitate later *p*-hacking as described below.

When hypothesizing you should clearly specify both the direction and the magnitude of an effect. Avoid vague statements like, "there will be an interaction." Instead, specify the exact pattern of the interaction: "For men, there will be a large effect of aggression (approx. d= 0.80), such that men in the high crowding condition will be more aggressive than men in the low crowding condition. For women, the effect of crowding will be smaller and negligible." Note that sample size planning (e.g., power analysis) requires an effect size estimate, so why not incorporate that effect size into your hypothesis?

Effect Size Specification

There are three primary approaches to specifying the effect size you expect for a hypothesis:

1. Published / Pilot Study.

Researchers often use the effect size from a previous study(ies) to guide sample size analysis. Although common, this approach is problematic because the effect size likely will be a substantial overestimate, given the combined effects of sampling error and publication bias. In practice, published effect sizes tend to be double of those from replications (Reproducibility Project). Therefore, one approach is to use an expected effect size that is half the published effect size. Alternatively, you can use a safeguard approach to determining expected effect size. A safeguard approach involves using the lower bound of the effect size confidence interval, which is easy to calculate if not provided. You should also consult Anderson et al. (2017) for additional solutions to this issue, and associated software. Regardless, published effect sizes should not be used "as is" in power estimates given that they are likely overestimates that will result in underpowered studies. Even meta-analytic estimates of the literature are likely biased (e.g., ego-depletion research). Consequently, if you choose this approach, you are well advised to use the lower bound of a meta-analytic effect size confidence interval following the safeguard strategy.

2. Standard/Common effect sizes (small, medium, large).

Another approach is to use small, medium, and large effect sizes. A common practice is to refer to Cohen's standards for small, medium, and large (correlations .10, .30, .50, respectively; *d*-values 0.20, 0.50, 0.80, respectively; partial-eta squared .01, .06, .14, respectively). However, these classifications are controversial (Ellis, 2010, p. 40). In terms of historical context, these effect sizes are based on an evaluation of effects of a single volume (Volume 61) of the *Journal of Abnormal and Social Psychology*.

The problem with this approach is that what constitutes a small, medium or large effect size varies greatly with research context. More recent evaluations of effect sizes have been done that are more comprehensive in nature – and domain specific. Note that, due to publication bias, the effect size estimates below are likely overestimates. Moreover, it may be difficult to know how an effect size was calculated – so you may want to calculate it on your own from the summary statistics provided in the article.

Industrial Organizational Psychology. A review of approximately 150,000 focal and non-focal relations over a 30-year period revealed a median correlation of .16 with an interquartile range of .07 (25^{th}percentile) to .32 (75^{th}percentile; Bosco, Aguinis, Singh, Field, & Pierce, 2015). Thus, in I-O Psychology, small, medium, and large correlations correspond to .07, .16, and .32.

Social Psychology.A review of 100 years of social psychological research revealed a mean correlation of .21, median correlation of .18, and a standard deviation of correlations across literatures of .15. (Richard, Bond, Stokes-Zoota, 2003). Thus, in Social Psychology, small, medium, and large correlations could be approximated (using the SD) as .03, .18, and .33. Likewise, small, medium, and large d-values correspond roughly to .06, .36, and .69.

Cognitive Neuroscience.A review of approximately 10,000 cognitive neuroscience articles revealed an inter-quartile range of *d*= 0.34 (25^{th}percentile) to *d*= 1.22 (75^{th}percentile; Szucs & Ioannidis, 2017). Thus, in cognitive neuroscience a small effect size is *d*= 0.34 and a large effect size is *d*= 1.22.

Clinical Child and Adolescent Psychology. To our knowledge, this field has not undertaken a self-study and so we suggest that student researchers in this area use the effect sizes from the analysis of the I/O or Social Psychology fields, which are comparable.

3. Minimum effect size of interest.

This approach is favoured among some statisticians in psychology due to the fact that the researcher must make it clear the direction and size of the effect in which he/she is interested. For example, let's say that you set the minimum effect size of interest as a *d*-value of 0.80 and use this value in your sample size analysis. In doing so, you are saying that a *d*-value of 0.75 (or anything below 0.80) is not of theoretical importance or interest. Moreover, you are saying that if you obtain an effect of *d*= 0.75 you are "ok" with it being non-significant due to its lack of theoretical importance. This approach gives the researcher the greatest *a priori*flexibility in determining what effect sizes are of interest.

Preregistration of Hypotheses

One solution to the problem of hypothesizing after results are known is the preregistration of hypotheses. This practice prevents researchers from conducting exploratory analyses and later reporting them as confirmatory. As noted above, preregistration of hypotheses before data collection is becoming increasingly important – and a requirement for journals using Level 3 of the TOP guidelines. Note that preregistration of associated data analysis plans is also important. It is discussed in the analysis section of this document.

The committee meeting in which you obtain approval of your thesis is, effectively, a process in which you preregister your thesis hypotheses with the committee. Why not go the extra step and register your

hypotheses with the Open Science Foundation (https://osf.io) or at As Predicted (https://aspredicted.org)?

These four links may be of interest: OSF 101, Preregistration: A Plan, Not a Prison, Open Science Knowledge Base, and the Open Science Training Handbook.

Student Check List 1 of 5: Hypothesizing

_____ The student created directional hypotheses.

_____ The student indicated what effect size should be expected for each hypothesis.

_____ The student explicitly indicated which analyses will be exploratory.

Source URL:https://www.uoguelph.ca/psychology/book-page/hypothesizing