Planning: When to Stop Data Collection

A clear plan for ending data collection needs to be in place before collecting data. The pros and cons of that plan should be weighed carefully. A common approach is to a) set your final sample size in advance of collecting data and b) ensure the sample/cell size is sufficiently high as discussed below. In the lab, setting the sample size in advance is relatively easy. When collecting observational field data, it can sometimes be difficult to obtain the desired sample size—consequently it is also important to assess the likely sample size and ensure the intended analysis approach is consistent with the likely final sample size. Indeed, the sample size around which results stabilize are often much higher (http://www.sciencedirect.com/science/article/pii/S0092656613000858) than researchers expect or wish (see Schönbrodt & Perugini, 2013). Ensuring that your research has adequate power is so important that *Psychological Science* has indicated that low power articles will likely receive desk rejections and not be reviewed (see Lindsay, 2015). Moreover, other researchers have moved beyond Power, Type-I, and Type-II errors into thinking in terms of Type-M and Type-S errors, representing magnitude and sign, respectively (see Gelman & Carlin, 2014 (http://pps.sagepub.com/content/9/6/641.full)).

Setting sample/cell size in advance

p-values cannot meaningfully be interpreted unless sample sizes are set in advance, rather than adjusted dynamically in light of whether a finding is significant (i.e., an optional-stopping approach based on p-values should not be used). Likewise, an optional-stopping strategy will upwardly bias effect-size estimates. Work by Simmons, Nelson, and Simonsohn (2011) (http://pss.sagepub.com/content/22/11/1359) illustrates how the practice of determining sample size dynamically has enormous implications for false-positive findings. Setting the sample size in advance avoids Questionable Research Practice (http://pss.sagepub.com/content/23/5/524) #2 ("Deciding whether to collect more data after looking to see whether the results were significant") and #4 ("Stopping collecting data earlier than planned because one found the result one had been looking for") (John et al., 2012 (http://pss.sagepub.com/content/23/5/524)). Indeed, in an excellent editorial (http://pss.sagepub.com/content/26/12/1827.full), the Editor-in-Chief of *Psychological Science*, Stephen Lindsay, classified the optional-stopping approach as a *p*-hacking strategy and indicated the journal will be on the "lookout" (Lindsay, 2015, p. 2) for this strategy. He even went so far as to say "*Whether these sorts of things are done innocently or nefariously, readers need to know about them to assess the replicability of the research. Senior authors must ensure that their supervisees understand the risks of p-hacking and understand that hiding p-hacking is unethical.*" (p. 2) Avoid this questionable research practice and set your sample/cell size in advance.

Determining needed sample/cell size

A large sample size decreases the width of confidence intervals associated with findings. As well, a large sample size increases the chance of obtaining a significant finding if there is an underlying effect. We note that a properly conducted sample size analysis (e.g., safeguard power analysis or confidence interval with sample size analysis) will often result in sample sizes larger than are typical in most fields. This result occurs because the sample sizes used in psychology are typically too small. Data from Open Science Collaboration (http://science.sciencemag.org/content/349/6251/aac4716) (2015) was recently re-analyzed using a Bayesian approach. This analysis revealed that 64% of studies had sample sizes sufficiently small that neither the null or alternate hypothesis could be strongly supported (Etz & Vandekerckhove, 2016 (http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0149794)). The authors concluded that "traditional sample sizes are insufficient." Thus, sample sizes are typically too small in psychology. Buttons et al. (2013) (http://www.nature.com/nrn/journal/v14/n5/full/nrn3475.html) went further and stated that the "*consequences of this include overestimates of effect size and low reproducibility of results. There are also ethical dimensions to this problem, as unreliable research is inefficient and wasteful.*" Consequently, we encourage students to ensure their sample/cell size is sufficiently large.

There are several approaches to determining the sample size needed for a study. There are two steps to determining sample size. First, generate an estimate of the effect size (point or interval estimate). Second, use the effect size estimate to calculate the required sample size using via power analysis or analysis of desired confidence interval widths. When conducting these analyses, think about all of your analyses. Indeed, it's relatively easy to obtain the required power of a one-way analysis of variance and not have sufficient power for the subsequent *a priori* paired comparisons. Ensure all planned analyses have sufficient power. We encourage a third step which involves comparing the required sample size to the one likely to be obtained—and then seeing if a different analysis/design is needed in light of possibly limited likely sample size. Because our Psychology Participant Pool imposes practical limits on obtaining a desired sample size – this can be an important question.

Obtaining an effect size estimate for sample size determination

i. **Approach 1 Expert Judgment:** Use expert judgment to estimate the likely effect size. Use this approach only if the approaches below are not possible. We suggest assuming a small to medium effect size in the absence of other knowledge.

ii. **Approach 2 Past Research Point Estimate:** Use an effect size from a previous study (or better a meta-analysis) as your estimate of the effect size. Not preferred – may result in problematically wide confidence intervals when interpreting results.

iii. **Approach 3 Past Research Interval Estimate**: Use an effect size from previous research and construct a confidence interval around it (Perugini et al. 2014 (http://pps.sagepub.com/content/9/3/319.abstract)). This approach is preferable because it acknowledges that effect sizes in past studies were influenced by sampling error. The confidence interval indicates a plausible range of population effects sizes that could have created the original effect size. Using the lower-limit of the confidence interval in the power analysis increases the probability of power analysis providing a meaningful estimate of the sample size needed to obtain significance if an effect exists. This approach is most likely to result in a narrow confidence interval when interpreting results.

Sample size determinations

i. **Traditional power analysis approach** using one of the effect size estimates above. Using the lower-limit of the confidence interval from a previous study's effect size is likely the most successful approach – see the details of safeguard power analysis (http://pps.sagepub.com/content/9/3/319.abstract) (Perugini et al. 2014).

ii. **Confidence interval width approach**. Determine the sample size needed to ensure the confidence interval width does not exceed the effect size (or some other criteria). For example, if a *d*-value of .50 is expected, then the upper-limit of the confidence interval minus the lower limit of the confidence interval should be less than .50—see Brand and Bradley (2016) (http://prx.sagepub.com/content/118/1/154.abstract). Note that to ensure the width of the confidence interval does not exceed the effect size, you should increase the sample size. Do not decrease the level of confidence (e.g., from 95% to 80%) to meet this objective. In light of the ASA statement and commentary below—a confidence interval approach appears justified.

Authors must decide the rule for terminating data collection before data collection begins and report this rule in the thesis.

CHECK: ____

Open Science Considerations (/psychology/book-page/open-science-considerations)

Up (/psychology/graduate/thesis-statistics)

Planning: Assessing Viability of Planned Sample Sizes (/psychology/book-page/planning-assessing-viability-planned-sample-sizes)

Printer-friendly version (/psychology/book/export/html/503)

Undergraduate Studies (/psychology/undergraduate)

Graduate Programs (/psychology/graduate)

College of Social & Applied Human Sciences (CSAHS)

- CSAHS website (http://uoguelph.ca/csahs)
- Department of Family Relations & Applied Nutrition (http://www.uoguelph.ca/family/)
- Department of Geography (http://www.uoguelph.ca/geography)
- Department of Political Science (http://www.uoguelph.ca/polisci/)
- Department of Psychology (/psychology/)
- Department of Sociology & Anthropology (http://www.uoguelph.ca/socioanthro/)

- Faculty/Staff Login (/psychology/user)

Mail:

Department of Psychology
University of Guelph
50 Stone Road East (http://maps.google.com/maps/ms?ie=UTF8&hl=en&msa=0&msid=106115594900620713079.000436308185e9184c678&ll=43.531407,-80.228584&spn=0.007607,0.014591&z=16&om=1)
Guelph, Ontario, N1G 2W1
Canada

Courier and Shipping:

Department of Psychology
University of Guelph
50 Stone Road East
4010 MacKinnon Ext. (Bldg. 154)
(https://www.google.com/maps/place/MacKinnon+Bldg,+University+of+Guelph,+87+Trent+Ln,+Guelph,+ON+N1G+1Y4,+Canada/@43.532459,-80.227324,16z/data=!4m2!3m1!1s0x
hl=en-US)

Guelph, Ontario, N1G 2W1
Canada
Telephone: 519-824-4120