Reproducibility, Open Science, and Preregistration

Data Analysis for Psychology in R 2

dapR2 Team

Department of Psychology
The University of Edinburgh

Learning Objectives

- What is open science, preregistration and why would we preregister studies.
- The connection between preregistration and reproducibility
- Concrete example of the benefits and difficulties of preregistration

Psychology has changed a lot in the last decade

- Open science
  - Open data (almost all journals)
  - All measures reported (many journals)
  - Push towards preregistration (more on that later)
  - Asking for direct replications
  - Increased sample size requirements and power analyses
  - And many, many other things

Why did the field change?

- Widely accepted practices in psychology were revealed to be fundamentally flawed
    - False positive rate in psychology MUCH higher than 1 in 20 findings
    - Extremely easy to find a statistically significant result → why is that?
- Researcher degrees of freedom
    - Many possible analytic decisions after looking at the data
    - Constant hunt for statistical significance (stargazing)

Researcher degrees of freedom

- How does this come about?
  - Ambiguity in what to do
  - Motivation

- Outcome is people end up searching for statistically significant results in the pile of tests they perform

- Decisions end up being unprincipled when post-hoc
  - What counts as a reason to exclude?
  - Missing all attention checks? 2?
  - Went too fast? Went to slow?

Academic Publishing

- Popular idea: Trust effects that are supported by many studies
- The problem is that scientists publish studies that "work" and file drawer those that do not "work"

Standard Scientific Practice Involves Analytic Freedom

- Stopping criterion: How many participants should I collect? Is 20 enough? 100?
- Exclusion criteria: Are there participants whom I should exclude? How would I figure that out? Is there a "bad" participant
- Analysis choices: What covariates should I include in my statistical model? How should I treat my dependent variable -- should I transform it in some way? How would that impact the conclusions I draw
- Family of comparisons: Do I have a key dependent variable or many I'm interested in? What are the implications of looking at one vs. many?

Nosek et al. 2018

"A vast number of choices in analyzing data could be made. If those choices are made during analysis, observing the data may make some paths more likely and others less likely. By the end, it may be impossible to estimate the paths that could have been selected had the data looked different"

Proposed Solution: Preregistrations

- Conduct exploratory or background research
- Form a prediction having explored some initial data / read the literature specifying the analysis (es) you aim to run completely.
- For example, DV ~ IV1 + IV2 vs. DV ~ IV1*IV2
- Test (either based on fit or test statistic) your model on this new data after having registered it
- Warning: This is more challenging than you might think for all but the simplest designs! And the less you know going into the study the harder these registrations will be.

Several ways to register your predictions

- https://aspredicted.org/
- https://osf.io/

What are included in registrations

- There are several templates for different branches of research, but the area you work in will likely affect what makes the most sense to use
- For social psychology, one template from vant Veer, A.E., and Giner-Sorolla, R. (2016)
- This template asks for information like the following:
    - Describe the hypotheses in terms of directional relationships between manipulated/measured variables
    - For interactions, describe the shape these interactions are likely to take
    - If you are manipulating a variable, make predictions for successful manipulation check variables or explain why no manipulation check is included
    - Describe the analyses which will test the main predictions and for each one include the relevant variables of interest, how they are calculated, the statistical technique / model and the rationale for including or omitting covariates

Other templates

More templates: https://osf.io/zab38/wiki/home/

- OSF Preregistration page: https://osf.io/k5wns/
- OSF How-To and Resources: https://www.cos.io/initiatives/prereg?ga=2.85949657.1114272946.1607948696-474133586.1547657474

Exploratory Analyses

- Exploration is good
  - Do not let concerns about preregistration interfere with asking questions about your data you didn't think of when you first registered your analyses!
  - In fact, exploration is very useful for discovering patterns in data that were not predicted and motivating future confirmatory research
- But, preregistrations make it clear (mostly to your future self but also to others) what analyses you considered primary when you began the project, and what was comparatively secondary.

Many Challenges

- Writing analysis plans is difficult and takes quite a lot of time
    - Foreseeing contingencies
    - The documents one creates can be pretty long
    - Upside: Once you get the data, the analysis is pretty fast. You've already thought a lot about the analyses you will run and even written the R script

Many Challenges

- With very new projects, it can be difficult to know exactly what to preregister
    - Initial preregistrations can be quite general and thus might not be completely convincing to the skeptical reader. But that doesn't mean they aren't worth doing.
    - Subsequent preregistrations of direct replications or follow-up experiments will often be more specific and constrained
    - Deviations and amendments are just fine, but the point is to be transparent about those deviations

What preregistration does not fix

- Deliberate dishonesty
- Preregistering predictions after you've already looked at the data
- Ignoring the preregistration (e.g., dropping a dependent variable that was a central part of your analytic plan)
- Results that don't generalize because of biased sampling, poorly designed experiments, or all the other stuff that can go wrong in science

Time for a break

Welcome Back!

Working through an example

- We're going to return to the example we discussed last week of looking at whether photos and sentiment predict retweeting and liking.

Prior set of findings

- We seem to find that both media type and the specific sentiment of the post predict retweeting!
- Let's turn to conducting a purely confirmatory analysis. How should we do that?

First, let's consider our prior findings and then write down our predictions.

- Tweets with photos were retweeted more not just on training data but also on test data
- Sentiment of a tweet, specifically negative sentiments, seemed to predict more retweeting as well (also on test data)

Issues:

- We used one metric of sentiment. Ideally, our findings should hold for other metrics of sentiment. *We should predict they will.*
- We've focused on retweeting entirely but we also see that retweets and favorites are extremely strongly correlated. *We should predict all of the same predictions will hold for favoriting just like retweeting. Or we need a good reason to distinguish them*
- We initially fit linear models but it's pretty clear those models are problematic. Need to fit a poisson model.
- Now let's write down our models that correspond to these hypotheses.

The things we listed above are going to form the backbone of our preregistration.

- Here are some of the models we are going to fit

```
# confirm.m2.poss.rt <- glm(retweet_count ~ media_type + sentiment , data = confirmsentiment1df, fa


# Photo + sentiment predict retweeting on new data set. Use the same sentiment dictionary.
# not run
```

```
# confirm.m2.poss.fav <- glm(favorite_count ~ media_type + sentiment , data = confirmsentiment1df,
# not run
```

```
# confirm.m2.poss.rt.s2 <- glm(retweet_count ~ media_type + value , data = confirmsentiment2df, fam

# Different metric of sentiment (dictionary 2) also predictive of retweeting.
# not run
```

```
# confirm.m2.poss.fav.s2 <- glm(favorite_count ~ media_type + value , data = confirmsentiment2df, f
# not run
```

```
# confirm.m2.poss.rt.s3 <- glm(retweet_count ~ media_type + sentiment , data = confirmsentiment3df,


# anger is reference condition of sentiment so all coefficients of sentiment are relative to that.
# not run
```

```
# confirm.m2.poss.fav.s3 <- glm(favorite_count ~ media_type + sentiment , data = confirmsentiment3d

# anger is reference condition of sentiment so all coefficients of sentiment are relative to that.
# not run
```

Where are we?

- After conducting exploratory analyses, we came up with a set of predictions we sought to confirm. Rather than just writing that information down for ourselves, we preregistered our predictions on a public repository.
- We went through the registration, but on the analytic front there were MANY things left unspecified.
  - What do we do about outliers? Why 20k observations? Why those activist organizations instead of others?
- But by preregistering something it makes it clear how open ended these analyses were, even though we were quite explicit in some ways.

Summary

- What covered what preregistration is, where to do it, and how to do it
- Concrete example of the benefits and difficulties of preregistration