Preregistration of Research Design

SOURCE: http://neurochambers.blogspot.co.uk/2012/10/changing-culture-of-scientific.html

What is preregistration?

Spell out in advance the methods and analyses you plan to do

Formally submit these plans to a journal for peer review and approval prior to conducting experiment

What are the requirements of preregistration?

What kinds of things do they want you to write down beforehand?

Distilled Preregistration

Stage 1: Registration review
Authors submit their initial manuscript prior to commencing their experiment(s). The initial submission would include the following sections

Background and Hypotheses
- A **brief review** of the relevant literature that motivates the research question, and a description of the aims and experimental **hypotheses**.

Methods
- Full description of proposed **sample characteristics**, including criteria for subject inclusion and exclusion, and detailed description of procedures for defining outliers.
- A description of **experimental procedures** in **sufficient detail** to allow another researcher to repeat the methodology
- Proposed **analysis pipeline**, including all preprocessing steps, and a precise description of every analysis that will be undertaken.

Optional bonus:
- A statistical **power analysis**.
- **Timeline** for completion of the study (don't have to verbally present, just include details on slide)

Example of a preregistration file from a replication study that was published in *Psychological Science*

Hypotheses
Procedure
Analysis Plan
Secondary Analyses

DESCRIPTION: https://osf.io/dr42m/wiki/home/
MATERIALS: https://dw2.psyc.virginia.edu/implicit/showfiles.jsp?user=chawkins&study=reprod4

What are the reasons to require preregistration?

Why should we want to make this change? The life sciences, in general, suffer from a number of serious problems including publication bias [1, 2], low statistical power [3, 4], undisclosed post-hoc analytic flexibility [5, 6, 7], and a lack of data transparency [8]. By valuing findings that are novel and eye-catching over those that are likely to be true, we have incentivised a range of questionable practices at individual and group levels. What's more, a worryingly high percentage of psychologists admit to engaging in dubious practices such as selectively reporting experiments that produced desirable outcomes (67%) and *p* value fishing (71%) [9].

…by requiring prospective authors to adhere to a preapproved methodology and analysis pipeline, it will eliminate a host of suspect but common practices that increase false discoveries…

"Because the study is accepted in advance, the incentives for authors change from producing the most beautiful story to producing the most accurate one."

Analysis Problems in Life Sciences

that could be solved by preregistration

Reproducibility Crisis: some reported effects fail to replicate in subsequent research

Publication bias: journals only publish studies that find significant differences (no record of how many similar studies failed to find same difference) – reported effects within 5%?

P-hacking: accusation that researchers finesse results using post hoc rules to exclude subjects as outliers, or to select statistically significant outcomes

Eating chocolate makes you thin

Posthoc interpretations

"Think of the measurements as lottery tickets. Each one has a small chance of paying off in the form of a "significant" result that we can spin a story around and sell to the media. The more tickets you buy, the more likely you are to win. We didn't know exactly what would pan out—the headline could have been that chocolate improves sleep or lowers blood pressure—but we knew our chances of getting at least one "statistically significant" result were pretty good."

The more measures you have, and the more tests you run, the more likely you'll find a significant effect somewhere. Need to be sure that it's not a **false positive**.

http://io9.com/i-fooled-millions-into-thinking-chocolate-helps-weight-1707251800

Posthoc interpretations

"With an exploratory analysis, whatever you find, you can never be sure it wasn't just a chance result. Perhaps I was lucky in having this brought home to me early in my career, when I had an alphabetically ordered list of stroke patients I was planning to study, and I happened to notice that those with names in the first half of the alphabet had left hemisphere lesions and those with names in the second half had right hemisphere lesions. I even did a chi square test and found it was highly significant. Clearly this was nonsense, and just one of those spurious things that can turn up by chance." Dorothy Bishop Blog

Posthoc interpretation would be: People with names in first half of alphabet more likely to have a left lesion than a right lesion, and vice versa for people in second half of alphabet.

http://deevybee.blogspot.co.uk/2013/07/why-we-need-pre-registration.html

Claim from preregistration folks:

If there is a record in advance of what you plan to do in your experiment, and which analyses you will conduct, post hoc interpretation of false positive effects will be eliminated.

What are the reasons NOT to require preregistration?

SOURCE: Chris Chamber article
SEE ALSO: http://andrewgelman.com/2014/01/23/discussion-preregistration-research-studies/

2. It all sounds too strict. Why would authors submit to this scheme when they can't change even one small aspect of their experimental procedure without being 'summarily rejected'? Even grant applications are not so demanding.

What would happen if ALL journals required preregistration?

Research that has been conducted but where a change was made in method part way through will never be published

3. Authors could game the system by running a complete study as per usual and submitting the methodology for registration review after the fact.

READ ABOUT HOW RESEARCHERS WORKED WITH HM: http://www.dana.org/News/Details.aspx?id=43051

5. A lot of the most interesting discoveries in science are serendipitous. Your approach will stifle creativity and data exploration.

Some whole areas of research are based on **sudden** and **unexpected findings**
Example: **Neuropsychology**

Patient HM was initially determined to have a specific memory deficit after surgery and researchers began testing within days after the surgery.
Other examples: Creolization, Children raised without language exposure, Awake brain stimulation

In these **perfectly legitimate** types of research, researchers decide what questions to ask about subjects **dynamically**, based on the subjects' responses.

7. What if the authors never submit a final manuscript because the results disagree with some desired outcome (such as supporting their preferred explanation)? How can you prevent publication bias on the part of the authors?

Recall: One reason to require preregistration is to combat **publication bias** (where researchers don't publish null findings)

How will we be able to distinguish different types of retractions? Studies that were conducted but failed vs. Studies that were never conducted vs. Studies that changed their methods

10. What if the authors obtain IPA but then realise (after data collection commenced) that part of their proposed methods or analyses were incorrect or suboptimal?

Preregistration is not how science *really* works and is based on an idealistic caricature of the scientific method

Science is a dynamic and creative process

Researchers have ideas as they're conducting experiments (not just beforehand)
interactions with subjects, talking with students and colleagues, reading new articles

Researchers develop new hypotheses as they're working with data (not just beforehand)

Need to develop techniques for protecting against false positives but preregistration is not ideal, because it increases wasteful bureaucracy while decreasing discovery

SOPHIE SCOTT ARTICLE: https://www.timeshighereducation.com/comment/opinion/pre-registration-would-put-science-in-chains/2005954.article

Policies began in 2012…so how many registered reports were published this year?

In 2015:
Cortex 0
Psychological Science 1%

Why so few?

In-class preregistration examples

Green & Bavelier (together)

Senghas & Coppola (in groups, discuss & submit at the end)

Background (couple of sentences)
Hypothesis (couple sentences)
Design Type (Natural/Controlled, between/within)
Sample (how many, criteria, how assigned to groups?)
Tasks (stimuli, response rules, response method, dependent & independent variables)
Analysis Plan (type of statistic, between or within subject comparisons, posthoc tests, outliers removed?)

Green & Bavelier

Mini pre-registration of Green & Bavelier, just Experiments 1-4 shown here

Background
Perceptual learning has previously been shown only in the context of highly specialized tasks, and only over short time periods on the order of hours (see Sagi & Tanne, 1994 for review). Here we test whether real world perceptual learning with action videogames has long-term effects on cognitive performance over a range of visual tasks.

Hypothesis
Action videogame playing enhances visual attention.

Design
Natural Experiment, Between-Subjects

Sample
Two Groups: Videogame players and Non-videogame players
Inclusion criteria: 18-23 years, males only, no history of impairment. Videogame players must have played action videogames for at least 1 hour per day, 4 days per week in the past 6 months. Non-videogamers must have little to no videogame experience in the past 6 months.
N: 16-26, 8-13 per group, per task
Recruitment: Flyers on college campus, Psychology subject pool

Data Collection
Subjects will be tested in the laboratory using computerized displays with responses collected via mouse and keyboard.

This example preregistration is available on Blackboard under Course Materials: Example Preregistration

Green & Bavelier

Tasks
Flanker Task: Participants report a square or diamond appeared in a target ring (of 6 rings) while ignoring distractor shape presented outside the rings. Trials could be compatible (distractor same shape as target), incompatible (distractor different shape), and range from easy (only one filled ring) to hard (all 6 rings contain shapes) at 4 levels of difficulty (0, 1, 3, 5 extra filled rings).
Independent Variables: Videogame Experience (No, Yes), Difficulty (4 levels)
Dependent Variable: Response Time Difference Incompatible – Compatible Trials
Number of Trials: ?

Enumeration Task: Some dots are flashed on the screen and subjects use a keyboard report how many. Larger numbers are harder to name than smaller numbers.
Independent Variables: Videogame Experience (No, Yes), Difficulty (Numbers 1-10)
Dependent Variables: %Error
Number of Trials: ?

Useful Field of View Task: Subjects fixate the middle square in a radial 8-arm array of squares (6 squares per arm) and have to report using a mouse where a target appeared in the array. The target can appear close to the fixation point or farther away and the distance determines how difficult it is.
Independent Variables: Videogame Experience (No, Yes), Difficulty (Eccentricities 10, 20, 30)
Dependent Variables: %Correct
Number of Trials: ?

Attentional Blink Task: Subjects are instructed to fixate centrally during presentation of a stream of black random letters (15/s) and they must detect the white target letters. For each target, subjects report with a key press when they see an X following the white (target) letter. There can be 1 to 8 intervening items between the target and the presentation of the X.
Independent Variables: Videogame Experience (No, Yes), Difficulty (Lag 1-8 items)
Dependent Variables: %X detected
Number of Trials: ?

Analysis Plan
Each subject's mean response at each level of difficulty will be calculated for each dependent measure.
Between-subjects effects: ANOVAs Videogame Experience (2 levels) x Difficulty (3-10 levels)
Posthoc t-tests contrasting Videogame Experience Groups at each level of Difficulty
Correlations between subjects' reports of videogame experience and performance
Outlier removal of subjects > 2.5 standard deviations from mean on overall performance per task
Alpha = .05

Senghas & Coppola

Background (couple of sentences)
Hypothesis (couple sentences)
Design Type (Natural/Controlled, between/within)
Sample (how many, criteria, how assigned to groups?)
Tasks (stimuli, response rules, response method, dependent & independent variables)
Analysis Plan (type of statistic, between or within subject comparisons, posthoc tests, outliers removed?)