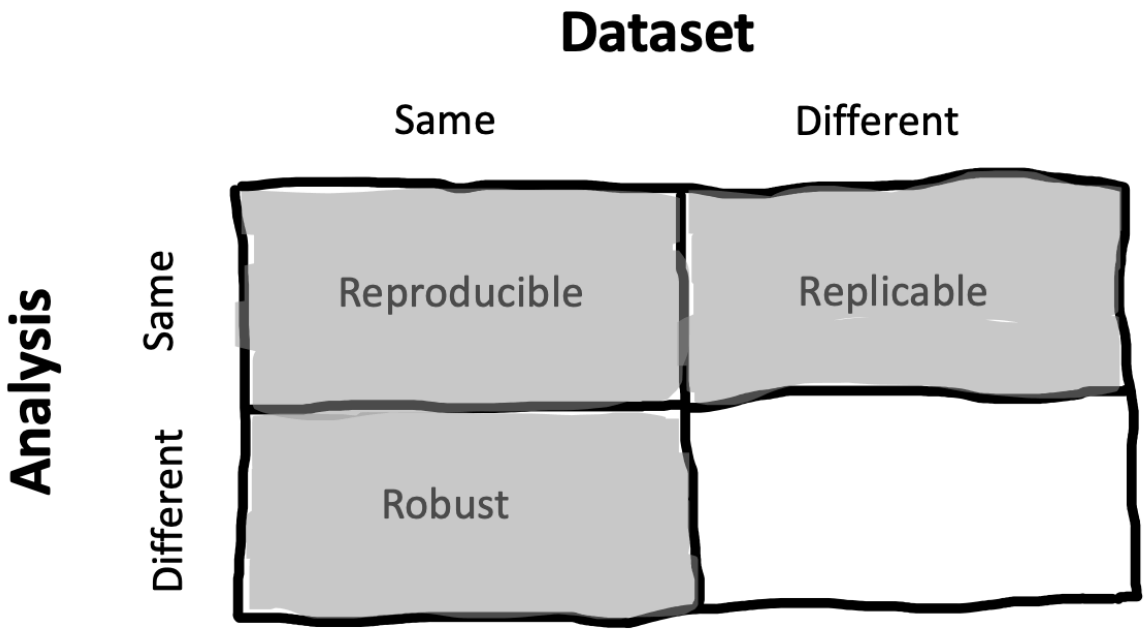
Preregistration & Open Science

(many slides created by Sebastian Schuster)

Open Science Collaboration, 2015; but see C. Anderson et al., 2016; Gilbert et al., 2016

Terminology



direct vs. conceptual replication

analytic flexibility (*p*-hacking) and publication bias (file drawer problem)

A typical psycholinguistics study

Hypothesis

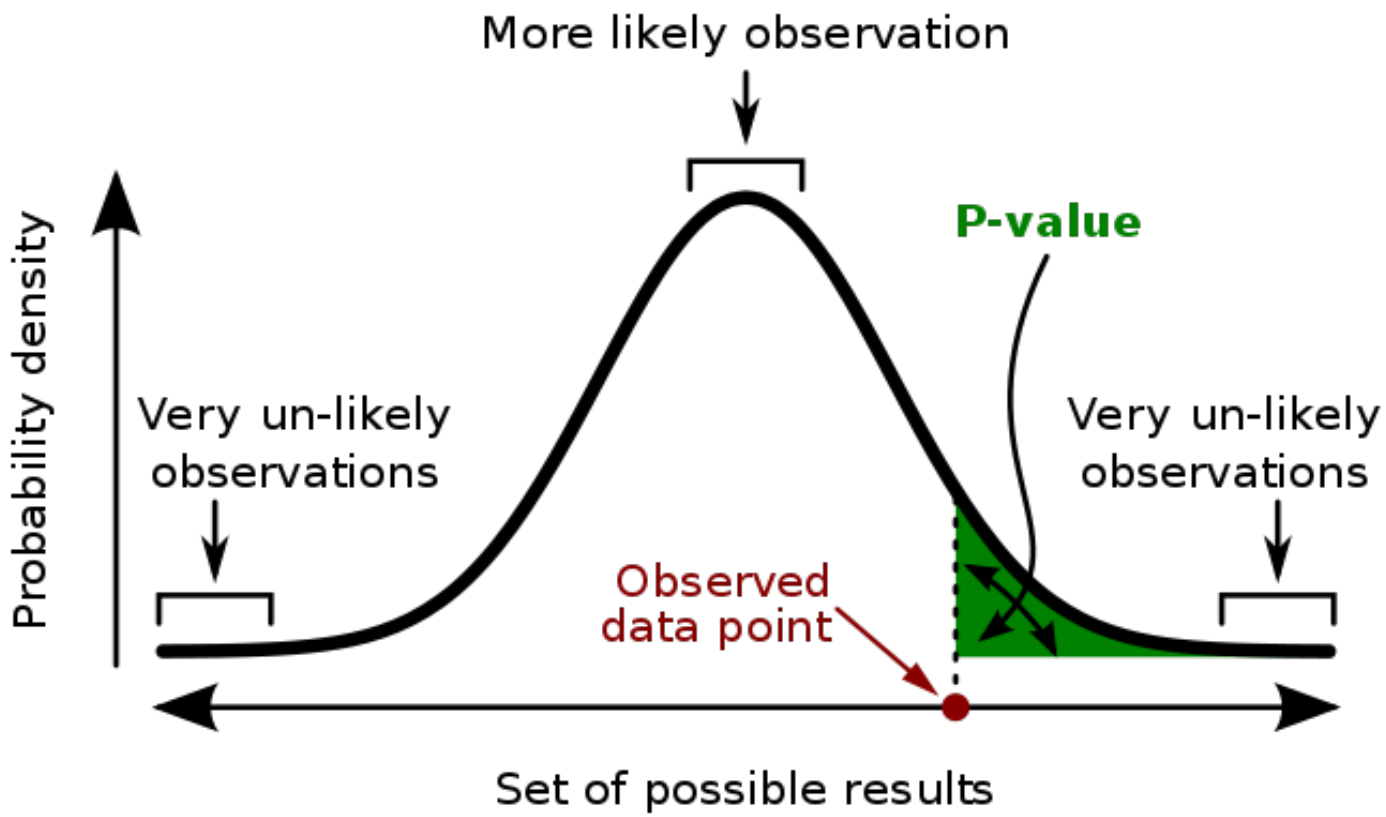
Reading of sentences with reduced relative clauses is slower than reading of sentences with overt complementizer

H_0: Average reading times of both sentence types are equal

The horse raced past the barn fell into a ditch

The horse that was raced past the barn fell into a ditch

p-values



A **p-value** (shaded green area) is the probability of an observed (or more extreme) result assuming that the null hypothesis is true.

Sketchy things you can do to get a significant *p*-val

1. Number of subjects per condition

    a. Run 10 subjects per condition

    b. Perform a t-test

    c. If $p < .05$: Publish paper!

       Otherwise: Go to step a.

Simmons et al. (2011)

Sketchy things you can do to get a significant *p*-val

2. Have multiple dependent variables

    a. Run tests to predict each of the variables

    b. Pick the dependent variable that gives you a significant *p*-value

The horse raced past the barn fell nto a ditch
The horse that was raced past the barn fell nto a ditch

Simmons et al. (2011)

Sketchy things you can do to get a significant *p*-val

3. Run models with many different independent variables

    a. Have a set of many independent variables

    b. Run models with various combinations and interactions until your manipulation is significant

Simmons et al. (2011)

Sketchy things you can do to get a significant *p*-val

4. Have conditions that you don't report on

    a. Run n > 2 conditions

    b. Pick 2 conditions which differ significantly and don't tell anybody about the other conditions

DON'T DO ANY OF THESE THINGS!!!

Simmons et al. (2011)

Simulated false-positive rates

	Significance level		
Researcher degrees of freedom	$p < .1$	$p < .05$	$p < .01$
Situation A: two dependent variables ($r = .50$)	17.8%	**9.5%**	2.2%

Simmons et al. (2011)

Simulated false-positive rates

	Significance level		
Researcher degrees of freedom	$p < .1$	$p < .05$	$p < .01$
Situation A: two dependent variables ($r = .50$)	17.8%	**9.5%**	2.2%
Situation B: addition of 10 more observations per cell	14.5%	**7.7%**	1.6%

Simmons et al. (2011)

Simulated false-positive rates

	Significance level		
Researcher degrees of freedom	$p < .1$	$p < .05$	$p < .01$
Situation A: two dependent variables ($r = .50$)	17.8%	**9.5%**	2.2%
Situation B: addition of 10 more observations per cell	14.5%	**7.7%**	1.6%
Situation C: controlling for gender or interaction of gender with treatment	21.6%	**11.7%**	2.7%

Simmons et al. (2011)

Simulated false-positive rates

	Significance level		
Researcher degrees of freedom	$p < .1$	$p < .05$	$p < .01$
Situation A: two dependent variables ($r = .50$)	17.8%	**9.5%**	2.2%
Situation B: addition of 10 more observations per cell	14.5%	**7.7%**	1.6%
Situation C: controlling for gender or interaction of gender with treatment	21.6%	**11.7%**	2.7%
Situation D: dropping (or not dropping) one of three conditions	23.2%	**12.6%**	2.8%

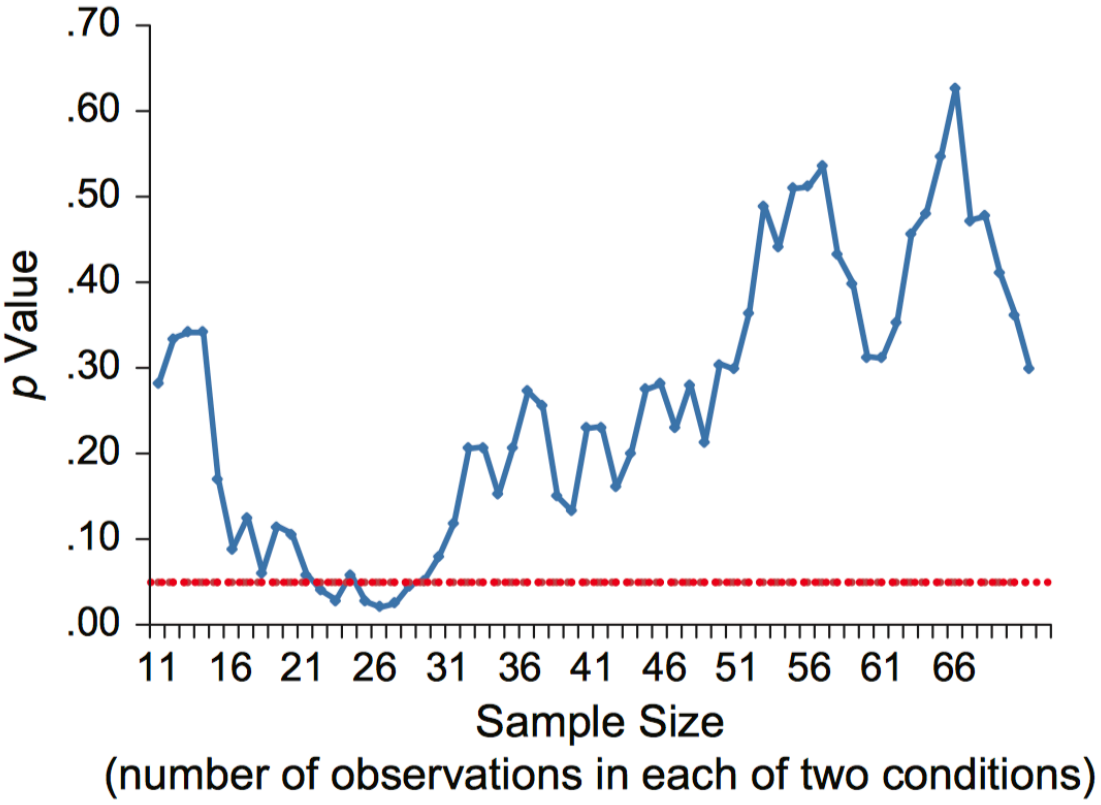
Simmons et al. (2011)

Simulated false-positive rates

Researcher degrees of freedom	Significance level		
	$p < .1$	$p < .05$	$p < .01$
Situation A: two dependent variables ($r = .50$)	17.8%	9.5%	2.2%
Situation B: addition of 10 more observations per cell	14.5%	7.7%	1.6%
Situation C: controlling for gender or interaction of gender with treatment	21.6%	11.7%	2.7%
Situation D: dropping (or not dropping) one of three conditions	23.2%	12.6%	2.8%
Combine Situations A and B	26.0%	14.4%	3.3%
Combine Situations A, B, and C	50.9%	30.9%	8.4%
Combine Situations A, B, C, and D	81.5%	60.7%	21.5%

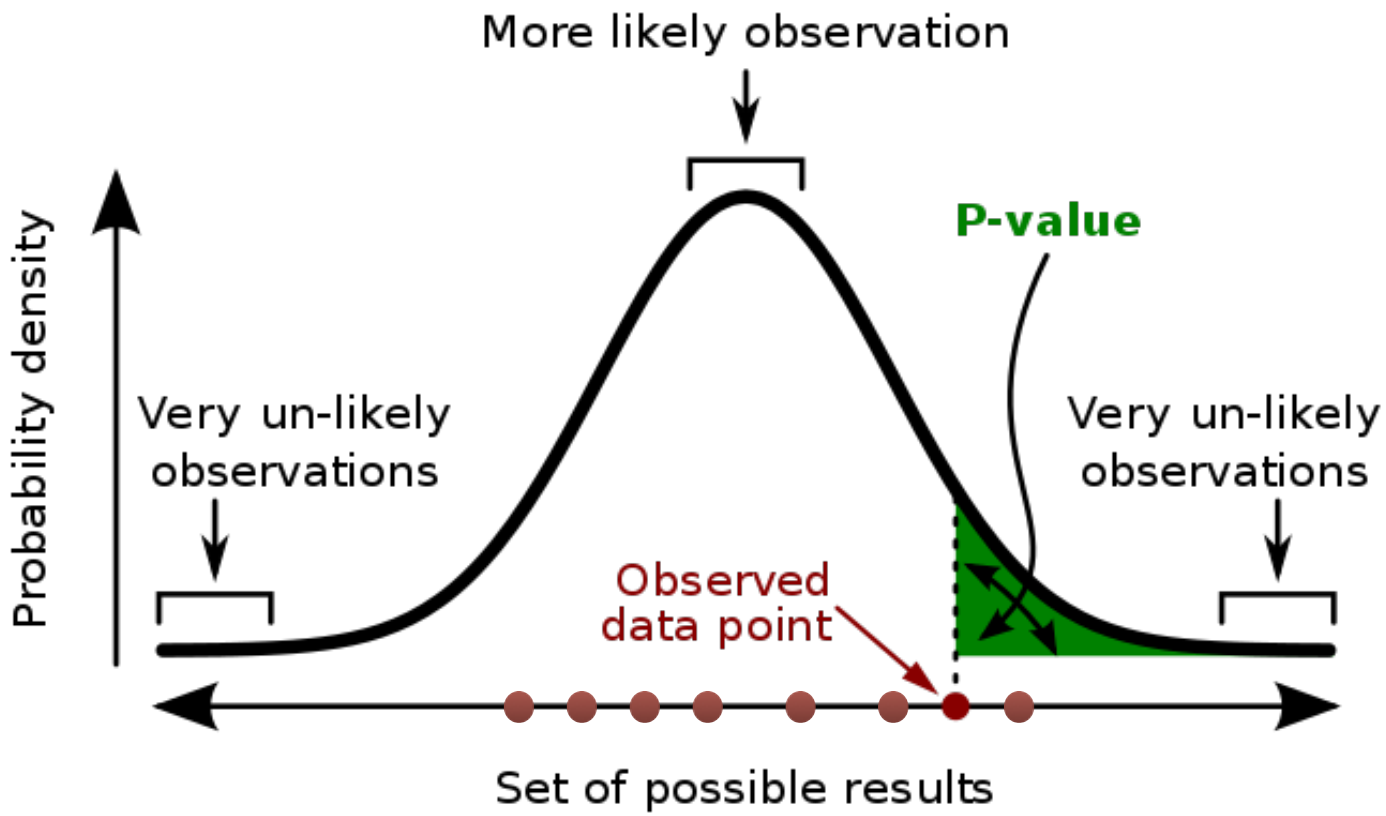
Simmons et al. (2011)

Simulated false-positive rates



Simmons et al. (2011)

p-values



A **p-value** (shaded green area) is the probability of an observed (or more extreme) result assuming that the null hypothesis is true.

Pre-registration

- To keep *p*-value meaningful, fix the following things **before collecting data**
  1. number of subjects you'll run
  2. exclusion criteria:
     Which data points are you going to exclude from your analysis
  3. dependent variable
  4. independent variables
  5. experimental conditions

Preregistration

- Preregistering provides you (and reviewers and readers of your paper) with proof that you actually fixed all these things

- Only requires filling out a short questionnaire which is permanently stored on a pre-registration platform

A good preregistration workflow

1. Come up with and implement experiment
2. Run pilot study with ~4 subjects
3. Write analysis scripts and test them with pilot data
4. Preregister study and upload analysis scripts to OSF
5. Run actual study
6. Analyze data with pre-registered analysis script
7. (optional) Do **exploratory** post-hoc analyses

www.osf.io

Guidelines

- Decide rule for terminating data collection before starting
- Collect at least 20 observations per cell
- List all variables collected in study
- Report all experimental conditions, including failed manipulations
- If observations are excluded also report results without exclusion
- Report results with and without covariates (where applicable)