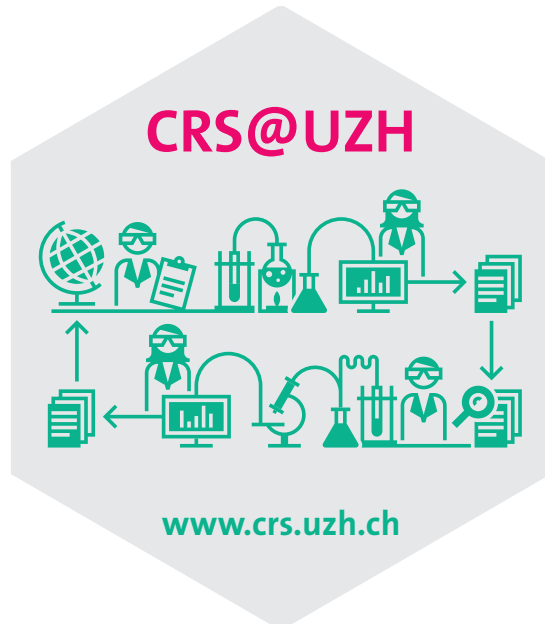
Tutorial on the R package ReplicationSuccess

Leonhard Held, Charlotte Micheloud, Samuel Pawel

Department of Biostatistics, Center for Reproducible Science



22.01.2020

Installation

- Linux / Windows

```
install.packages(pkgs = "ReplicationSuccess",
                 repos ="http://R-Forge.R-project.org")
```

- Mac

```
install.packages(pkgs = "ReplicationSuccess",
                 repos = "http://R-Forge.R-project.org",
                 type = "source")
```

Replication studies

Direct replication

- Repeating original study using the same methodology

$\rightarrow$ Tool to assess credibility of scientific discoveries

$\rightarrow$ Regulatory requirement

Replication studies

Direct replication

- Repeating original study using the same methodology

$\rightarrow$ Tool to assess credibility of scientific discoveries

$\rightarrow$ Regulatory requirement

Replication crisis

- Low replicability of many scientific discoveries

$\rightarrow$ Increased interest in meta-science

$\rightarrow$ Large-scale replication projects

Large-scale replication projects

– 2015: Reproducibility project psychology

Large-scale replication projects

- 2015: Reproducibility project psychology
- 2016: Experimental economics replication project

Large-scale replication projects

- 2015: Reproducibility project psychology
- 2016: Experimental economics replication project
- 2018: Experimental philosophy replicability project

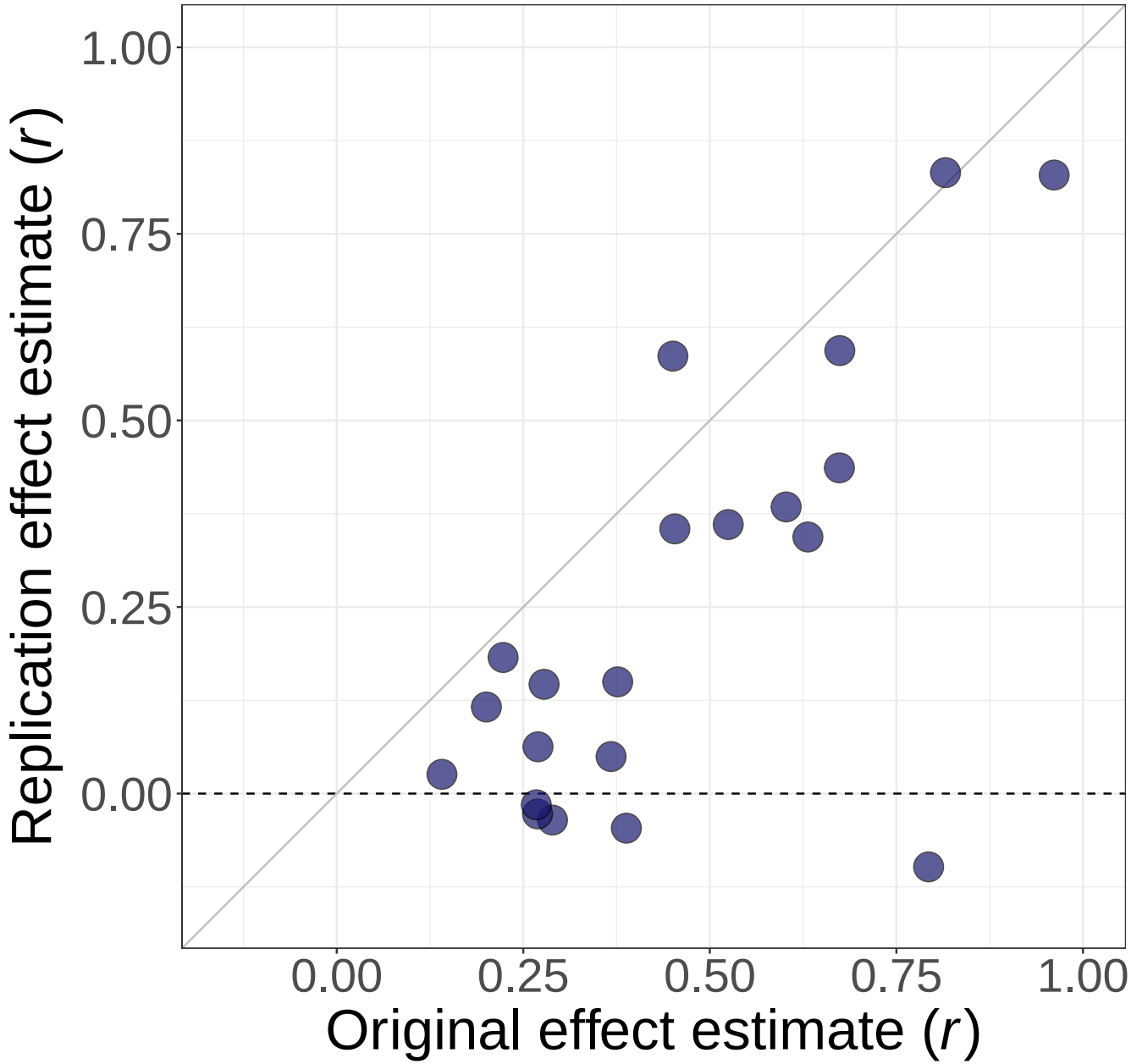
Large-scale replication projects

- 2015: Reproducibility project psychology
- 2016: Experimental economics replication project
- 2018: Experimental philosophy replicability project
- 2018: Social sciences replication project

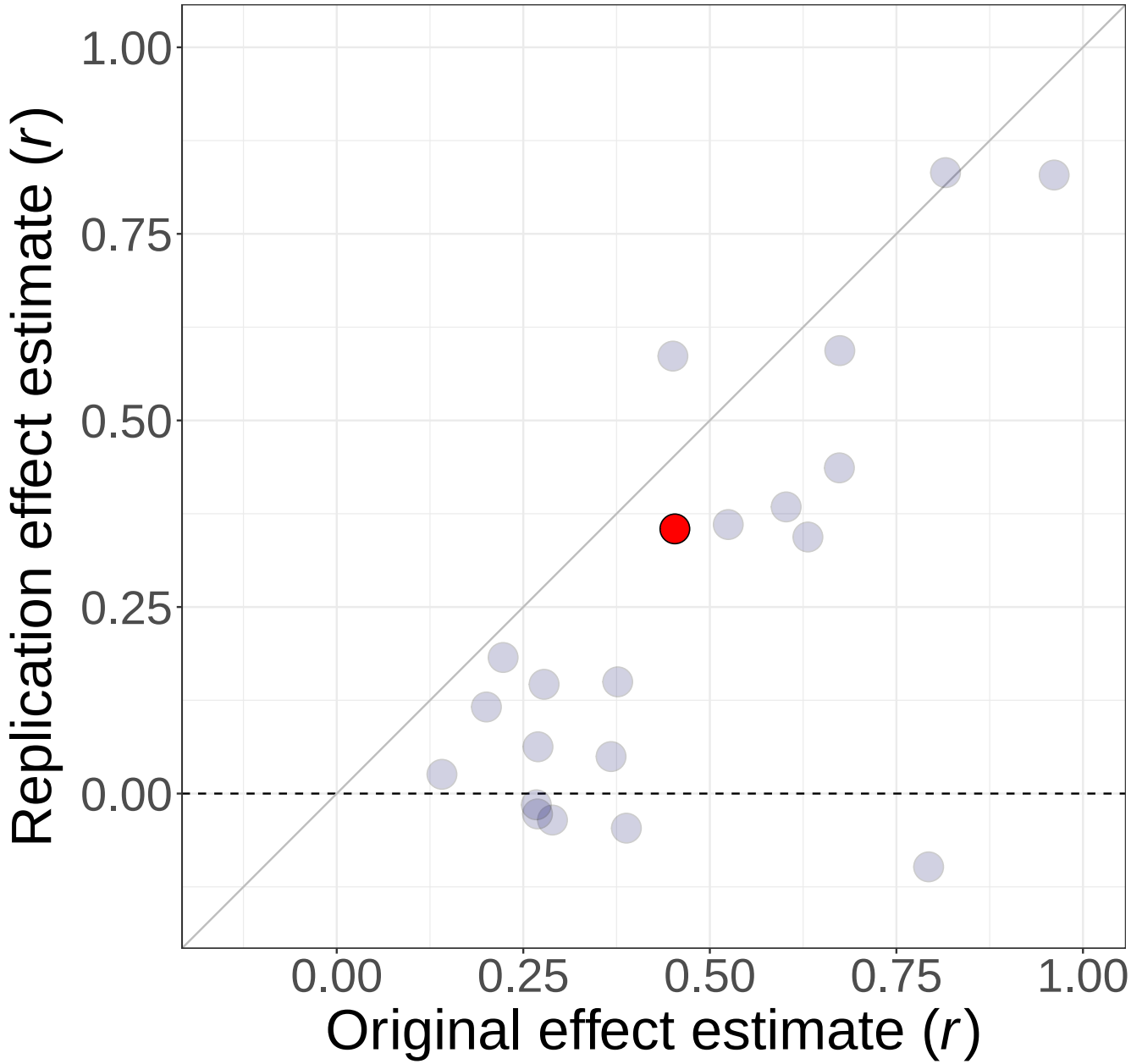
Large-scale replication projects

- 2015: Reproducibility project psychology
- 2016: Experimental economics replication project
- 2018: Experimental philosophy replicability project
- 2018: Social sciences replication project

Social sciences replication project



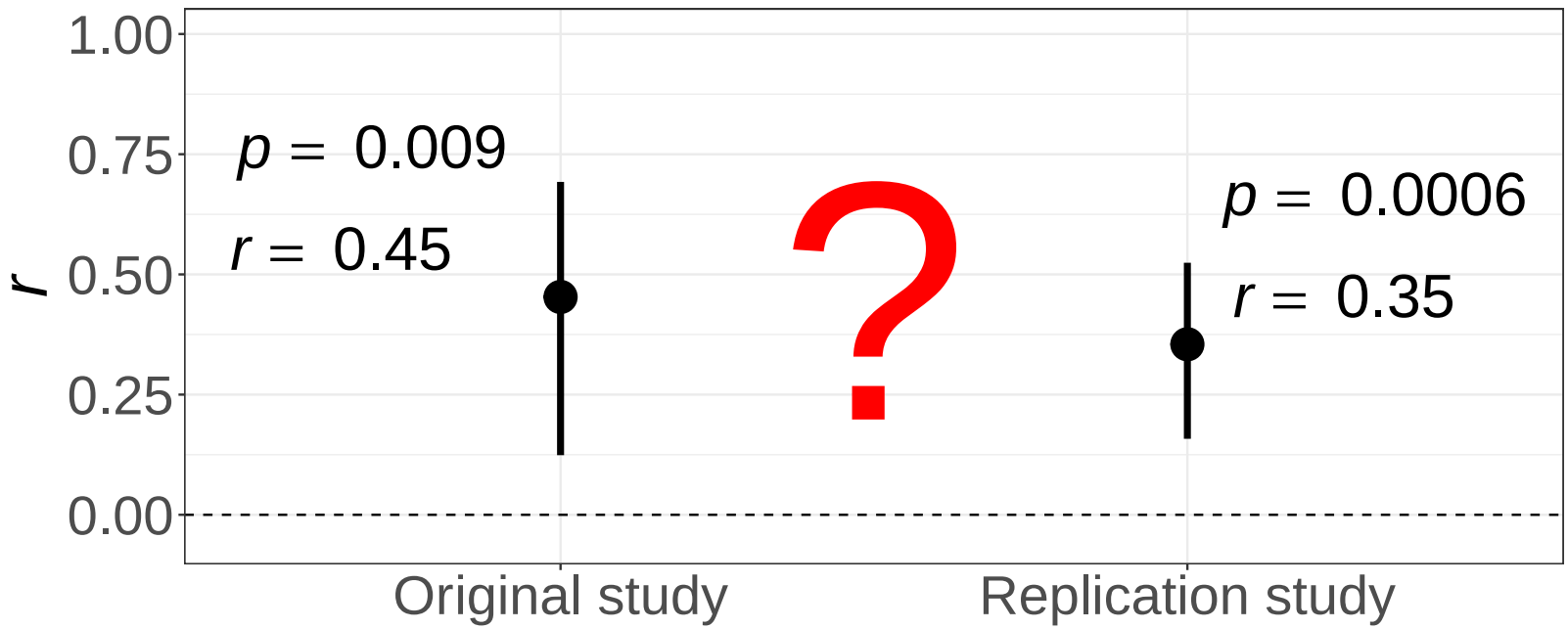
Social sciences replication project



Morewedge et al. (2010). Science

Original discovery

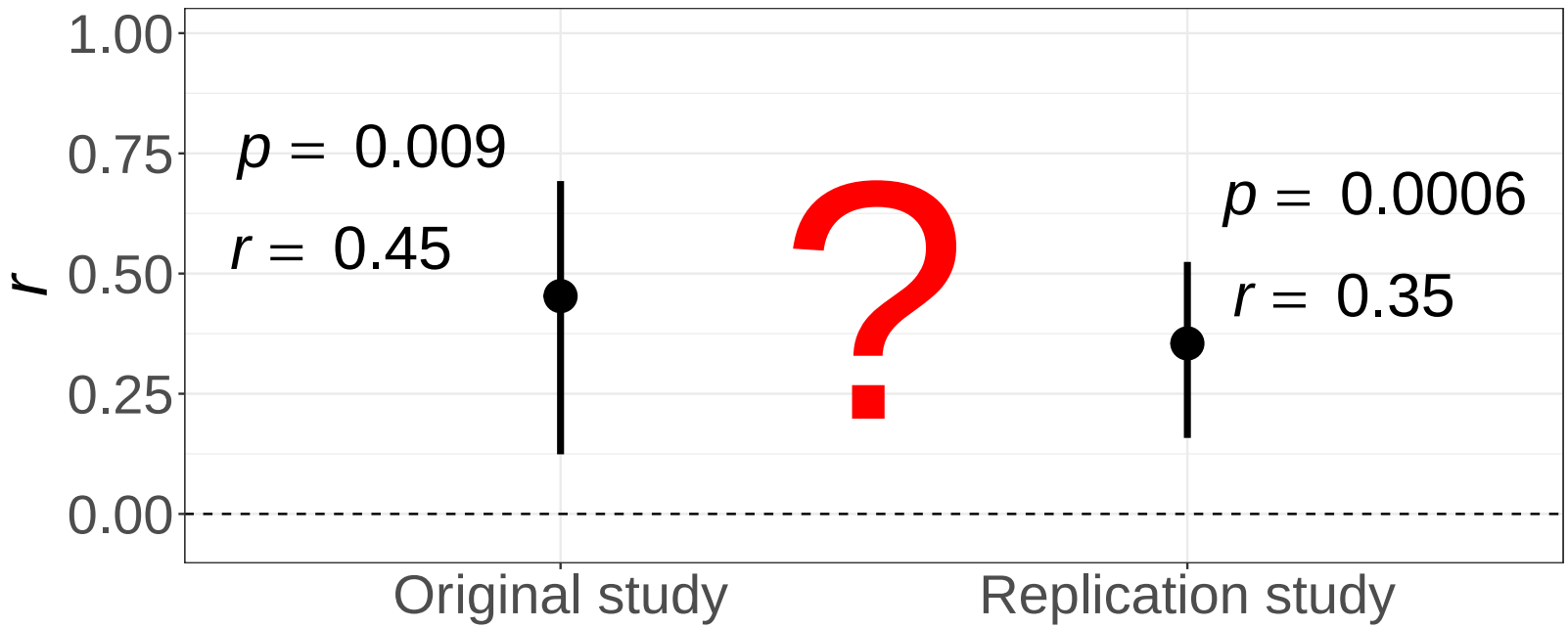
"Repeatedly imagining eating a food subsequently reduces the actual consumption of that food"



When is a replication successful?

Some proposed criteria for replication success

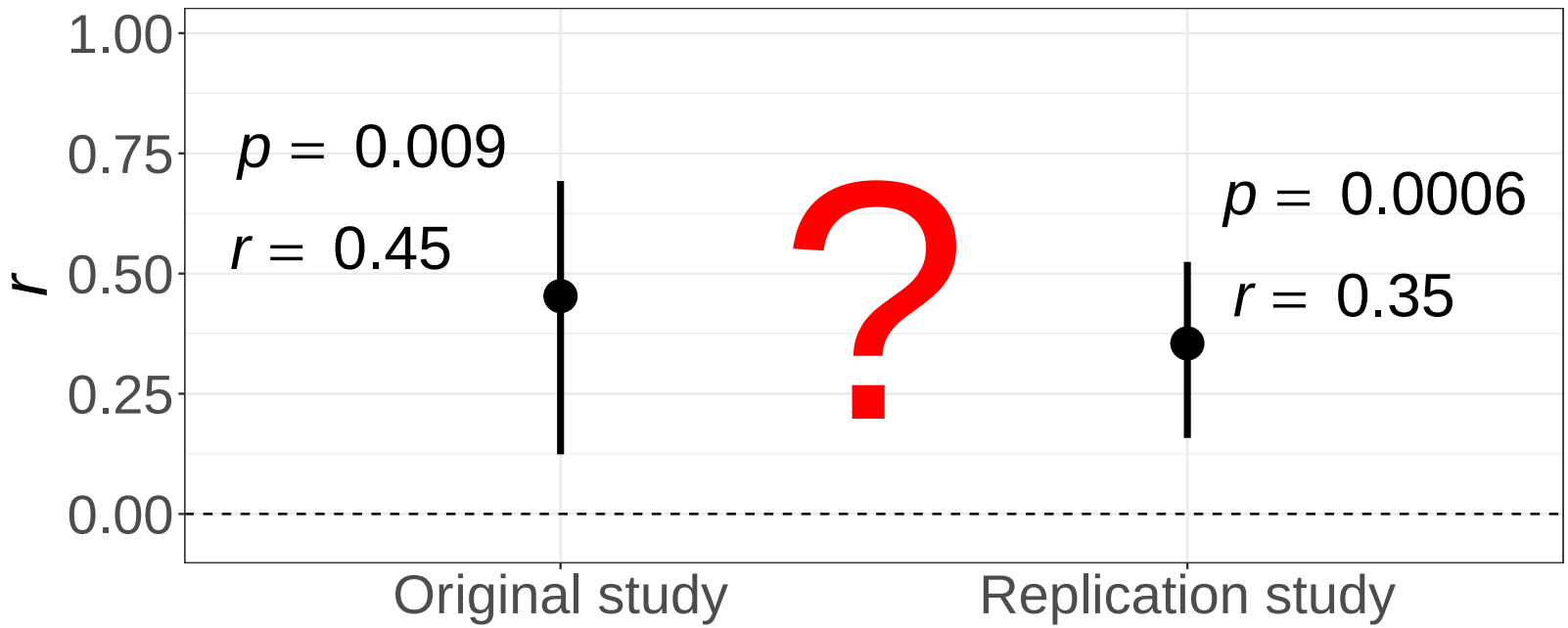
1. Statistical significance



When is a replication successful?

Some proposed criteria for replication success

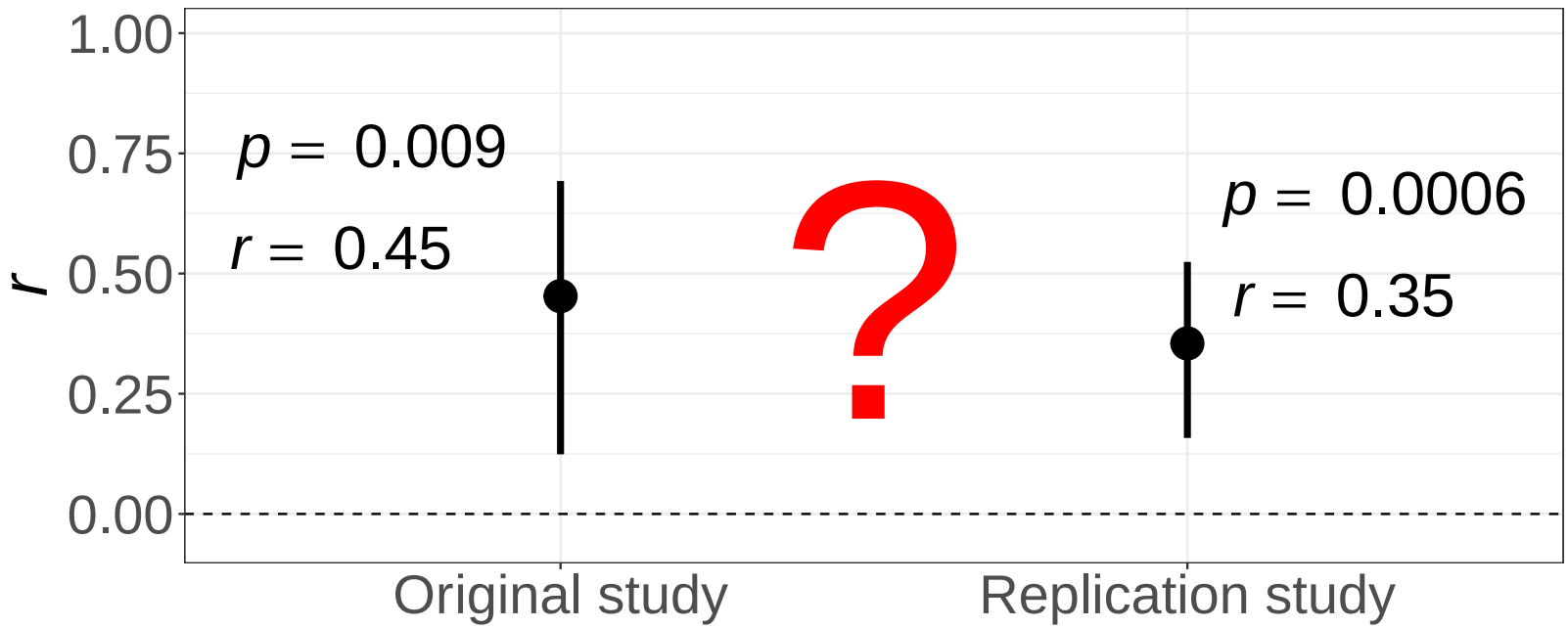
1. Statistical significance
2. Comparison of effect estimates



When is a replication successful?

Some proposed criteria for replication success

1. Statistical significance
2. Comparison of effect estimates
3. Reverse Bayes methods



Drawbacks of approaches

- Signficiance can always be achieved by increasing sample size
- Estimates can be compatible but uninformative

Design of replication studies

– Discuss sample size calculation important and depends on analysis strategy – Discuss what was done in projects – Ad hoc shrinkage

Statistical framework of package

– normality assumption – relative quantities – DesignPrior: Predictive and conditional

Significance

– powerSignficance + arguments – sampleSizeSignificance + arguments – exercises

Comparison of effect size

– predictionInterval – sampleSizePI – sampleSizePIwidth

Reverse Bayes

– pSceptical – powerReplicationSuccess – sampleSizeReplicationSuccess

Outlook

– Interim – Heterogeneity – EB shrinkage

References

Camerer, C. F., Dreber, A., Forsell, E., Ho, T., Huber, J., Johannesson, M., Kirchler, M., Almenberg, J., Altmejd, A., Chan, T., Heikensten, E., Holzmeister, F., Imai, T., Isaksson, S., Nave, G., Pfeiffer, T., Razen, M., and Wu, H. (2016). Evaluating replicability of laboratory experiments in economics. *Science*, 351:1433 – 1436.

Camerer, C. F., Dreber, A., Holzmeister, F., Ho, T., Huber, J., Johannesson, M., Kirchler, M., Nave, G., Nosek, B., Pfeiffer, T., Altmejd, A., Buttrick, N., Chan, T., Chen, Y., Forsell, E., Gampa, A., Heikenstein, E., Hummer, L., Imai, T., Isaksson, S., Manfredi, D., Rose, J., Wagenmakers, E., and Wu, H. (2018). Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015. *Nature Human Behavior*, 2:637 – 644.

Cova, F., Strickland, B., Abatista, A., Allard, A., Andow, J., Attie, M., Beebe, J., Berniūnas, R., Boudesseul, J., Colombo, M., Cushman, F., Diaz, R., N'Djaye Nikolai van Dongen, N., Dranseika, V., Earp, B. D., Torres, A. G., Hannikainen, I., Hernández-Conde, J. V., Hu, W., Jaquet, F., Khalifa, K., Kim, H., Kneer, M., Knobe, J., Kurthy, M., Lantian, A., Liao, S.-y., Machery, E., Moerenhout, T., Mott, C., Phelan, M., Phillips, J., Rambharose, N., Reuter, K., Romero, F., Sousa, P., Sprenger, J., Thalabard, E., Tobia, K., Viciana, H., Wilkenfeld, D., and Zhou, X. (2018). Estimating the reproducibility of experimental philosophy. *Review of Philosophy and Psychology*.

Held, L. (2019). A new standard for the analysis and design of replication studies (with discussion). *Journal of the Royal Statistical Society: Series A (Statistics in Society)*.

Open Science Collaboration (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251):aac4716.

Pawel, S. and Held, L. (2019). Probabilistic forecasting of replication studies. Preprint.