

# PROPOSAL

## Bayesian Evidence Synthesis

Thom Volker, 5868777

Supervisor: Irene Klugkist

*Methodology and Statistics for the Behavioural, Biomedical and Social Sciences*

*Utrecht University*

September 06, 2020

Word count: 662

# 1 Introduction

Word count is inclusief comments, zonder comments is de word count 479.

In recent years, a meta-analytic way of thinking has been advocated in the scientific community, an approach that is grounded in the belief that a single study is merely contributing to a larger body of evidence (Cumming, 2014). Additionally, the importance of replication has been legitimately supported (e.g., Open Science Collaboration, 2015; Baker, 2016; Brandt et al., 2014). However, most of the attention has been focused on studies that are highly similar, using an identical methodology and research design. These studies, commonly referred to as exact, direct or close replications, are merely concerned with the statistical reliability of the results. Unfortunately, if the results of these studies depend on methodological flaws, inferences from all studies will lead to suboptimal or invalid conclusions (Munafò & Smith, 2018). A safeguard against this deficiency is available in the form of conceptual replications, which primarily assess the validity of a study. That is, conceptual replications are a way of investigating whether the initial conclusions hold under different conditions, using varying measurement instruments or choosing different operationalizations.

As a consequence, multiple studies regarding the same hypotheses arise and as per the cumulative nature of science, synthesizing the results is required to build a robust and solid body of evidence. When the studies are highly similar, established methods as (Bayesian) meta-analysis and Bayesian updating can be used to pool the results (Glasauer, 2019; Lipsey & Wilson, 2001; Sutton & Abrams, 2001). Heel vreemd, maar ik kan echt zo goed als niks over Bayesian Sequential Updating vinden. Ondanks dat het in deze bron (Glasauer) wel oke uitgelegd wordt, is het niet echt ons interessegebied geloof ik. However, when researchers purposefully alter varying aspects of an earlier study with similar hypotheses, as is the case with conceptual replications, or unintentionally make different data-analytic choices as compared to earlier research, a situation that is referred to as the garden of forking paths (Gelman & Loken, 2014), these methods are not applicable. Namely, these methods are restricted to combine parameter estimates that (i) share a common scale, and (ii) arise from analyses with identical functional forms. [Twee opmerkingen mbt voorgaande zin: (1) Is het noodzakelijk om "with similar hypotheses" toe te voegen? Voor mijn gevoel is dat namelijk logisch gegeven de nadruk op replicatie, maar ik zit ondertussen al wat verder in het onderwerp. (2) Ik twijfel of ik de *garden of forking paths* erin moet betrekking, want enerzijds geeft het extra gewicht aan de noodzaak om dit te onderzoeken,

anderszijds komt het zo misschien wat uit de lucht vallen, en ik weet niet zo goed of ik het waard vind om het veel uitgebreider uit te leggen binnen de slechts 750 woorden die ik heb]. Consequently, Kuiper, Buskens, Raub, & Hoijtink (2013) proposed to use Bayesian Evidence Synthesis (BES), which allows researchers to pool the evidence for a specific hypothesis over multiple studies, even if the studies have seemingly incompatible designs.

Voor mijn gevoel is de komende alinea vrij rommelig, met erg veel nieuwe termen, die hiervoor nog niet gebruikt worden, waardoor het voor mijn gevoel nogal verwarrend is.

BES combines seemingly incompatible parameter estimates by expressing the information that these estimates contain in terms of a Bayes Factor. Bayes Factors render the relative amount of evidence for the hypothesis of interest relative to another hypothesis(???). This procedure is undertaken in every single study, after which the obtained Bayes Factors can be combined into one overall measure of evidence by means of multiplication (Kuiper et al., 2013). The focus thus is shifted from the parameter estimates, to the hypotheses underlying the parameters, reflecting a theoretical construct that is measured in all studies. Therefore, the parameter estimates do not have to be comparable by themselves, as long as the hypotheses under investigation in all studies represent an equivalent theoretical construct.

## 2 References

- Baker, M. (2016). Reproducibility crisis. *Nature*, 533(26), 353–366. <https://doi.org/10.1038/533452a>
- Brandt, M. J., IJzerman, H., Dijksterhuis, A., Farach, F. J., Geller, J., Giner-Sorolla, R., ... Van't Veer, A. (2014). The replication recipe: What makes for a convincing replication? *Journal of Experimental Social Psychology*, 50, 217–224. <https://doi.org/10.1016/j.jesp.2013.10.005>
- Cumming, G. (2014). The new statistics: Why and how. *Psychological Science*, 25(1), 7–29. <https://doi.org/10.1177/0956797613504966>
- Gelman, A., & Loken, E. (2014). The statistical crisis in science: Data-dependent analysis—a "garden of forking paths"—explains why many statistically significant comparisons don't hold up. *American Scientist*, 102(6), 460–466.
- Glasauer, S. (2019). Sequential bayesian updating as a model for human perception. In S. Ramat & A. G.

- Shaikh (Eds.), *Mathematical modelling in motor neuroscience: State of the art and translation to the clinic. Gaze orienting mechanisms and disease* (pp. 3–18). <https://doi.org/10.1016/bs.pbr.2019.04.025>
- Kuiper, R. M., Buskens, V., Raub, W., & Hoijtink, H. (2013). Combining statistical evidence from several studies: A method using bayesian updating and an example from research on trust problems in social and economic exchange. *Sociological Methods & Research*, 42(1), 60–81.
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. SAGE publications, Inc.
- Munafò, M. R., & Smith, G. D. (2018). Robust research needs many lines of evidence. *Nature*, 553(7689), 399–401. <https://doi.org/10.1038/d41586-018-01023-3>
- Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251). <https://doi.org/10.1126/science.aac4716>
- Sutton, A. J., & Abrams, K. R. (2001). Bayesian methods in meta-analysis and evidence synthesis. *Statistical Methods in Medical Research*, 10(4), 277–303. <https://doi.org/10.1177/096228020101000404>