A Scoping Review Protocol of Methodological Heterogeneity in Psychological Reviews and Meta-Analyses

B.A Edmar

The Question

The main goal of this review is to clarify the concept of methodological heterogeneity in systematic reviews and and meta-analyses in psychology. Specifically, we want to examine the problem of comparability across pooled effects and how researchers deal with problematic methodological heterogeneity. Another way to phrase this question could be: at what point are qualitative differences between scientific inquiries large enough that we conclude that the effects we are measuring are different, regardless of whether a measurable difference is present? In essence, we want to know what causes the exclusion of a study from a meta-analysis due to perceived methodological heterogeneity.

In order to answer this question a scoping review of the literature is needed. The literature in this case being recently published articles within the field of psychological science. Given that our goal is to clarify the concept of methodological heterogeneity, identify gaps in our knowledge, and asses how the execusion of research is impacted by the presence of methodological heterogeneity, a scoping review is preferable to a systematic review(Munn et al. 2018). An alternative to a scoping review could be a rapid review, that being a feasibility constrained systematic review, however, it is not the feasibility of our investigation that informs our method - it is the underlying research question. Since we are mainly concerned with the qualia of methodological heterogeneity, a quantitative assessment of the literature does not target the question.

Introduction

An often overlooked problem in meta-analytic studies within psychology is the handling of heterogeneity (Linden and Hönekopp, 2021). Heterogeneity occurs when the pooled effects contain more variation than expected from random sampling. This is often expressed with the

τ^2 statistic, which represent the between-study variation observed in the sample. The potential cause of the observed heterogeneity can be attributed to the fact that multiple ''true" effects are observed in the data and therefore more variation than expected is present within the sample. The presence of multiple true effects can be caused by the presence of subgroups of effects that remain unadjusted for through meta-regression or subgroup analyses. These subgroups of effects can be caused by a multitude of things such as differing research methodologies, populations and other design based matters.

However, statistical heterogeneity is not *necessarily* present even if studies are methodologically diverse or have fundamentally different measurements. In those cases we make the decision to exclude effects from the analysis based on some qualitative metric. The framework for these decisions are are constructed through the establishment of exclusion criteria a priori to screening the data (McKenzie et al. 2019). With the exclusion criteria researchers define how to decide wheat studies that are comparable, and what studies that are not comparable. However, sometimes the presence of statistical heterogeneity can cause a reevaluation whether studies are comparable or not, and in these cases an additional qualitative assessment of comparability is needed.

Inclusion Criteria

To be included in this systematic scoping review, any document needs to be a systematic review or meta-analysis within any of the following PsychInfo classification categories: Physiological Psychology and neuroscience (2500), Developmental Psychology (2800), Social Psychology (3000), Personality Psychology (3100), Health & Mental Health Treatment & Prevention (3300), Organizational Psychology (3600), Cognitive psychology and Intelligent systems (4100), Forensic psychology & Legal Issues (4200) published after 2016 to capture recently published studies. These categories are selected to ensure that a representative sample of different areas of psychology are selected. We want to capture as much of the context of psychological research as possible and stratifying our sample to be representative is our way of promoting external validity. Since this is meta-research the participants we are targeting are published research papers, thus the main source of evidence will be published papers. The concept of interest within these participants is the treatment of methodological heterogeneity in systematic reviews and meta-analyses.

Search Strategy

While the JBL guidelines detail a specific three-stage search strategy(*JBI Manual for Evidence Synthesis* 2020), following this strategy is not feasible in our case due to the breadth of the literature we are interested in. Since the field we are interested in is psychology, searching databases outside psychology will have to include field limitations that would not be needed

in psychology specific data bases such as psychInfo. Therefore, we will only search psychInfo since that database provides us with a natural exclusion of documents outside the field of psychology.

Taking inspiration from the sampling procedure by Linden and Hönekopp (2021), a stratified sample of studies from the PsychInfo classification categories will be collected and screened using the software ASReview. Studies will then be included in the data-set as long as the data-set is not information saturated, meaning that additional studies provide additional qualitative information to the data-set. An alternative approach would be to set a fixed level of studies to include from each sub-field.

The following Ovid search terms will be stratified across: (meta-analy*.mp. or exp Meta Analy/* or systematic review.mp.) and heterogeneity.mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures, mesh word].

This search yields 4885 results as of 28th of September 2023. When the results were filtered through the selected APA categories 441 results remained.

Including ''heterogeneity" in the search terms is to ensure finding studies that in some form explicitly mentions heterogeneity. The inclusion of this term reduces the number of results by 38420 results. Whether or not the inclusion of the term will bias the results away from studies that neglect covering heterogeneity is a concern, since only 55% of meta-analyses in psychology addressees heterogeneity (Linden and Hönekopp 2021).

Data Extraction

An excel code book for what elements to extract from each study will be constructed. This code book will cover three main themes: general study characteristics, the presence of methodological heterogeneity in the study, and how the authors dealt with the observed heterogeneity. Across these three themes we will include notation of the aims of the study, the sub-field of the study, the population examined, the methodology used, outcome variables, heterogeneity measures and adjustment, exclusion criteria, findings/conclusion of the study, etc.

Data Analysis

The analysis of the data will consist of summarizing the motivation for the various choices made in the individual data points. These will then be compared across sub-fields to see how the practice of adjusting for heterogeneity varies dependent on the sub field. These will be presented with simple descriptive statistics but also with a qualitative content analysis of how authors reason about their analysis choices and exclusion criteria with regards to methodological heterogeneity. This content analysis will aim to be descriptive and will not seek to find any emergent or latent themes within the data.

Questions for meeting Monday October 2nd

- What should our search strategy be?
- Should we pick a fixed number of studies or sample until we are comfortable with making inferences from the sample?
- What do you think about analyzing exclusion criteria?
- What should the code book consist of?
- What typeology of heterogeneity do we like? What exactly do we mean by methodological heterogeneity?

References

JBI Manual for Evidence Synthesis. 2020. JBI. https://doi.org/10.46658/JBIMES-20-01.

Linden, Audrey Helen, and Johannes Hönekopp. 2021. "Heterogeneity of Research Results: A New Perspective From Which to Assess and Promote Progress in Psychological Science." *Perspectives on Psychological Science* 16 (2): 358–76. https://doi.org/10.1177/1745691620964193.

McKenzie, Joanne E, Sue E Brennan, Rebecca E Ryan, Hilary J Thomson, Renea V Johnston, and James Thomas. 2019. "Defining the Criteria for Including Studies and How They Will Be Grouped for the Synthesis." In *Cochrane Handbook for Systematic Reviews of Interventions*, edited by Julian P. T. Higgins, James Thomas, Jacqueline Chandler, Miranda Cumpston, Tianjing Li, Matthew J. Page, and Vivian A. Welch, 1st ed., 33–65. Wiley. https://doi.org/10.1002/9781119536604.ch3.

Munn, Zachary, Micah D. J. Peters, Cindy Stern, Catalin Tufanaru, Alexa McArthur, and Edoardo Aromataris. 2018. "Systematic Review or Scoping Review? Guidance for Authors When Choosing Between a Systematic or Scoping Review Approach." *BMC Medical Research Methodology* 18 (1): 143. https://doi.org/10.1186/s12874-018-0611-x.