A Scoping Review Protocol of Heterogeneity in Meta-Analyses within Psychology

B.A Edmar

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 eligibility 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 re-evaluation whether studies are comparable or not, and in these cases an additional qualitative assessment of comparability is needed.

The Question

The main goal of this review is to clarify the concept of methodological heterogeneity in systematic reviews and meta-analyses in psychology. Specifically, we want to examine the problem of comparability across pooled effects and how researchers deal with problematic methodological diversity. 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 execution 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.

Eligibility Criteria

To be included in this systematic scoping review, any document needs to be a systematic review or meta-analysis within psychology published after 2021 to capture recently published studies. Additionally, documents must have an applied or theoretical research question that requires the synthesis of evidence across multiple independent studies to be included. This means that studies posing meta-research questions will be excluded. An example of such a question would be an evaluation of the prevalence of open-science practices or replications within a given field. The data within the documents must have been generated through systematic search of the literature, this excludes sequential studies that aggregate their findings using an ''internal meta-analysis".

Scoping reviews investigate a population and a concept in a specific context. For this review the population under investigation are published systematic reviews and meta-analyses in psychology. The concept that we want to investigate is the handling of methodological diversity across studies. The context we are interested in is recently published documents. We are not interested how this has been done historically but how it is currently being done.

Methods

This review will follow the guidelines provided by JBI to the greatest extent possible for the research question at hand. The manuscript will follow the Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) (Tricco et al. 2018).

Search Strategy

The JBI(*JBI Manual for Evidence Synthesis* 2020) guidelines state that at least two databases should be subject to an initial search, thereafter a second search using all identified keywords should be conducted, and then the reference list of all identified sources should be consulted for additional documents. While these guidelines are sound, following this strategy is not feasible in our case due to the breadth of the literature we are interested in. Therefore we deviate from the JBI manual in this area.

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. As an extension of this strategy is to identify PsychInfo classification categories that can provide a representative sample of papers across the many sub-fields within psychology. To this end we limit our search to papers within the following categories: Developmental Psychology (2800), Social Psychology (3000), Personality Psychology (3100), Organizational Psychology (3600), Cognitive psychology and Intelligent systems (4100), Forensic psychology & Legal Issues (4200). These categories are selected to ensure that a representative sample of different areas of psychology are searched for while limiting the number of search results. Note that health and clinical categories are excluded, this is due to the heavy overlap in medical and psychological research in these areas.

Search Terms

The search was conducted through the data base Ovid in three steps detailed in table 1.

Table 1

	Search Terms
1	(meta-analy* or meta analy*).mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures, mesh word]
2	limit 1 to (2800 developmental psychology or 3000 social psychology or 3100 personality psychology or 3600 organizational psychology & human resources or 4100 cognitive psychology & intelligent systems or 4200 forensic psychology & legal issues)
3	limit 2 to yr="2022 -Current"

The initial search resulted in 53 041 hits, the first limitation reduced this to 2015 hits, and the final limitation produced 283 documents.

Document screening

The identified documents will be screened using the software tool rayyan.ai. The screening will be conducted in a two stage fashion where the abstracts from the initial search are screened for eligibility and labeld according to the sub-field of psychology they cover, after which full text screening will be conducted on a subset of the identified studies. As of Thursday the 12th of October 2023 the initial screening of the abstracts is complete, resulting in 165 documents. The next step is to formalise a procedure for how to select studies to analyse from this pool of documents. Since the goal is to produce generalisable results, a diverse and representative sample is required, but given that the analysis is mostly qualitative a smaller purposefully selected sample of documents is needed.

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 16th

- How should we select studies? I think sampling until saturation across sub-fields sounds like a good idea.

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.

Tricco, Andrea C., Erin Lillie, Wasifa Zarin, Kelly K. O'Brien, Heather Colquhoun, Danielle Levac, David Moher, et al. 2018. "PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation." *Annals of Internal Medicine* 169 (7): 467–73. https://doi.org/10.7326/M18-0850.