STATISTICAL HORIZONS

Meta-Analysis

David Wilson, Ph.D.

Upcoming Seminar:
October 20-21, 2017, Philadelphia, Pennsylvania

Meta-Analysis Workshop

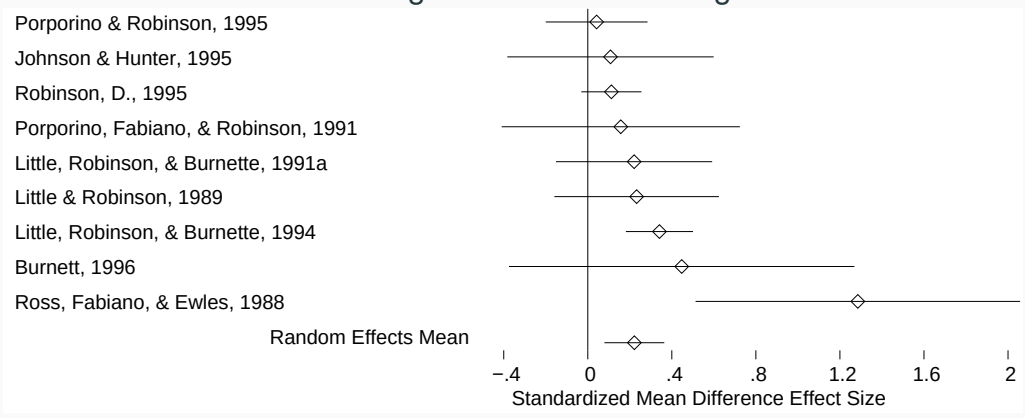
David B. Wilson, PhD

September 16, 2016

George Mason University
Department of Criminology, Law & Society
dwilsonb@gmu.edu

The End-Game

Forest-Plot of Standardized Mean Differences and 95% Confidence Intervals for the Effects of Cognitive-Behavioral Programs on Recidivism



Overview

- Historical background
- Logic of Meta-analysis
- Effect sizes
  - Common types
  - Computing standardized mean difference effect sizes
  - Computing odds ratio effect sizes
- Basic meta-analysis methods
- Random-effects versus fixed-effect model
- Moderator analysis
  - Analog to the ANOVA
  - Meta-analytic regression
- Forest plots
- Publication bias
- Cutting edge methods
  - Network meta-analysis
  - Robust standard errors for statistically dependent effect sizes
  - Regression coefficient and fully multivariate models

Historical Background

A Great Debate

- Eysenck 1952: Psychotherapy doesn't work
- Dizzying array of mixed results followed
- Glass (with Smith) average results from 375 studies
- Glass coined the term **meta-analysis**

Deep Roots

- Pearson (1904): averaged correlations between inoculation for typhoid fever and mortality
- Fisher (1944): independent studies individually may not be significant, yet the aggregate seem improbable
- W. G. Cochran (1953): developed methods of averaging means across studies
- A. Wicker (1967) average correlations between attitudes and behavior
- Concurrent with Smith and Glass (1977) were
  - Hunter and Schmidt (1977) *Validity generalization*
  - Rosenthal and Rubin (1978) *Interpersonal expectancy effects*

Logic of Meta-analysis

Logic of Meta-analysis

- Narrative review methods:
  - Focuses on statistical significance
  - Lacks transparency and replicability
- Weakness of statistical significance:
  - Significant effect is a strong conclusion
  - Non-significant effect is a weak conclusion
  - How do you balance a collection of significant and non-significant effects?

Logic of Meta-analysis

- Meta-analysis:
  - Focuses on **direction** and **magnitude** of effect
  - Approaches task as a research endeavor
  - Examines pattern of evidence across studies
    - Average effect
    - Consistency of effects
    - Relationship between study features and effects

Research Suitable to Meta-analysis

Forms of Research Findings Suitable to Meta-analysis

- Central tendency research
- Prevalence rates
- Pre-post contrasts
- Growth rates
- Group contrasts
- Experimentally created groups
- Comparison of outcomes between treatment and comparison groups
- Comparison of two naturally occurring groups

Forms of Research Findings Suitable to Meta-analysis

- Association between variables
  - Measurement research
  - Validity generalization
  - Individual differences research
  - Correlation between personality constructs
  - Regression models (can be done but challenging)

Concept of Effect Size and Essential Features

Effect Size: The Key to Meta-analysis

- The effect size makes meta-analysis possible
- It is the “dependent variable”
- It standardizes findings across studies such that they can be directly compared

Effect Size: The Key to Meta-analysis

- Any standardized index can be an “effect size” (e.g., standardized mean difference, correlation coefficient, odds-ratio) as long as it meets the following
  - Is comparable across studies (generally requires standardization)
  - Represents the magnitude and direction of the relationship of interest
  - Is independent of sample size
- Different meta-analyses may use different effect size indices

Study Inclusion

The Replication Continuum

You must be able to argue that the collection of studies you are meta-analyzing examine the same relationship. This may be at a broad level of abstraction, such as the relationship between criminal justice interventions and recidivism or between school-based prevention programs and problem behavior. Alternatively it may be at a narrow level of abstraction and represent pure replications.

The closer to pure replications your collection of studies, the easier it is to argue comparability.

Which Studies to Include?

- It is critical to have an explicit inclusion and exclusion criteria
- The broader the research domain, the more detailed they tend to become
- Refine criteria as you interact with the literature
- Components of a detailed criteria
  - distinguishing substantive features
  - research design or designs
  - participants
  - key variables
  - cultural and linguistic range
  - time frame
- A comment about publication types

Methodological Quality Dilemma

- Include or exclude low quality studies?
- The findings of all studies are potentially in error (methodological quality is a continuum, not a dichotomy)
- Being too restrictive may restrict ability to generalize
- Being too inclusive may weaken the confidence that can be placed in the findings
- Methodological quality is often in the “eye-of-the-beholder”
- Balance you strike must fit the purpose of the meta-analysis

Searching Far and Wide

- The “we only included published studies because they have been peer-reviewed” argument
- Significant findings are more likely to be published than nonsignificant findings
- Critical to try to identify and retrieve all studies that meet your eligibility criteria

Strengths and Weaknesses of Meta-analysis

Strengths of Meta-analysis

- Imposes a discipline on the process of summing up research findings
- Represents findings in a more differentiated and sophisticated manner than conventional reviews
- Capable of finding relationships across studies that are obscured in other approaches
- Protects against over-interpreting differences across studies
- Can handle a large numbers of studies (this would overwhelm traditional approaches to review)

Weaknesses of Meta-analysis

- Requires a good deal of effort
- Mechanical aspects don't lend themselves to capturing more qualitative distinctions between studies
- "Apples and oranges" criticism
- Most meta-analyses include "blemished" studies to one degree or another
- Publication and outcome reporting bias posses a continual threat
  - Negative and null finding studies that you were unable to find
  - Outcomes for which there were negative or null findings that were not reported
- Analysis of between study differences is fundamentally correlational

Effect Sizes

The Heart and Soul of Meta-analysis: The Effect Size

- Meta-analysis shifts focus from statistical significance to the *direction* and *magnitude* of effect
- Key to this is the effect size
- It is the dependent variable of meta-analysis
- Encodes research findings on a numerical scale
- Different types of effect sizes for different research situations
- Each type may have multiple methods of computation

Overview

- Main types of effect sizes
- Logic of the standardized mean difference
- Methods of computing the standardized mean difference
- Logic of the odds ratio and risk ratio
- Methods of computing the odds ratio
- Adjustments, such as for baseline differences
- Issues related to the variance estimate

Most Common Effect Size Indexes

- Standardized mean difference (*d* or *g*)
  - Group contrast (e.g., treatment versus control)
  - Inherently continuous outcome construct
- Odds ratio and Risk ratio (*OR* and *RR*)
  - Group contrast (e.g., treatment versus control)
  - Inherently dichotomous (binary) outcome construct
- Correlation coefficient (*r*)
  - Two inherently continuous constructs

Less Common Effect Size Indexes

- Raw (unstandardized) mean difference
- Proportions
- Standardized gain score
- Standardized regression coefficient

Necessary Characteristics of Effect Sizes

- Numerical values produced must be comparable across studies
- Must be able to compute its standard error
- Must not be a direct function of sample size

The Standardized Mean Difference Effect Size