



Meta-analysis

- Define
- Importance
- Key features
- Strengths & Weaknesses



Definition

- A synthesis of many trial results.
- An overview in which quantitative methods are used to summarize the results of several studies on a single topic.

Meta-analysis



Importance

“I had hoped to find research to support or to conclusively oppose my belief that quality integrated education is the most promising approach. For every study that contains a recommendation, there is another, equally well documented study, challenging the conclusions of the first... No one seems to agree with anyone else's approach. But more distressing, no one seems to know what works.”

~ Fritz Mondale

Meta-analysis



What problems for researchers does Mr. Mondale's statement identify?

- Educational research often produces contradictory results.
- Differences among studies in treatments, settings, measurement instruments, and research methods make research findings difficult to compare.
- Even frequent replications can prove inconclusive.
- Literature on a topic may be so extensive as to obscure trends with an overwhelming amount of information.

Meta-analysis



Steps in a Meta-Analysis

1. Identify studies
2. Define eligibility criteria
 - Quality
 - Sample Size (N)
 - Date
 - Randomized /non-randomized
3. Create structured formats to capture the data results of each study
4. Do statistical analysis to calibrate data results on a common scale

Meta-analysis

Key Features of Meta-Analysis

- The purpose (research question) is defined, and also differential factors affecting outcomes.
- Gathering the research data focuses on the magnitude and direction (+ /-) of “effect size” for each study.
- Statistical analysis gives more or less weight to each study based on defined criteria using mathematical formulas...

Meta-analysis

Yikes!

Statistical Analysis

It's all about weighting the effect size:

- The standard error is needed to calculate the effect size weights, called inverse variance weights
- All meta-analytic analyses are weighted
 - Homogeneity Analysis
 - Fixed effect variance, random effect variance
 - Inverse Variance: weighting studies with a larger sample size
 - The Odds-Ratio is the odds of success in the treatment group relative to the odds of success in the control group.
 - 95% Confidence Interval
 - Q factor



Standard Deviation

The standard deviation is a measure of the spread of the data. In most cases (and in a normal distribution), about 95% of the subjects will have a value within 2 standard deviations of the mean. Thus if the mean height of some patients is 5 foot 8 inches, and the standard deviation is 3 inches, this means that most of the patients (19 in every 20) will have heights between 5 foot 2 inches and 6 foot 2 inches.

<http://www.ubht.nhs.uk/R&D/RDSU/Advisory/glossary.htm>

Standard Error

The standard error of the mean is a measure of how certain we are about the value of this mean. A standard error always relates to some kind of statistical result, such as a mean or an odds ratio. Unlike the standard deviation, the value of the standard error depends on the size of the study. It is small when we have a large amount of data, and big when we have a tiny study. The standard error is used to generate the confidence interval.

95% Confidence Interval

If you repeated the study 100 times, you would expect the results to lie within this interval in 95% of the studies. Confidence interval values can be calculated for most statistical results such as odds ratios. They can be thought of as representing the probable range in which the result lies. A meta-analysis usually produces an estimate of the treatment effect size, and a confidence interval around that estimate.

<http://www.ubht.nhs.uk/R&D/RDSU/Advisory/glossary.htm>

Odds Ratio

Odds of having an event in one group compared with that of having the same event in a different group. Odds ratio of 2 for recovery in the treatment group roughly translates to twice as likely to recover in the treatment group compared with those in the untreated group.

<http://www.ubht.nhs.uk/R&D/RDSU/Advisory/glossary.htm>

p-value

The p value represents the probability of a result of a test of association having occurred by chance if there was actually no association between the variables.

Probability values lie between 0 and 1. If the p value is less than 0.05, it signifies that a result like this could only have appeared by chance 5% of the time if no association actually existed. We say that this is a statistically significant finding. Please note that if you test 20 different associations, you would expect one to be significant by chance. To avoid this error minimize the number of tests, or use more stringent criteria for significance such as $p < 0.01$.

<http://www.ubht.nhs.uk/R&D/RDSU/Advisory/glossary.htm>



Effect Size

- Used to measure the findings of each study, which makes meta-analysis possible.
- Puts results from different studies on a single uniform scale of effectiveness.
- The type of effect size must be comparable across the studies.
- Positive when the experimental group outperforms the control group and negative when the control group comes out on top.

Meta-analysis



Effect Size

- It encodes the research findings on a numeric scale.
- Specifies the number of standard deviation units that separate outcomes of experimental and control groups.

$$\text{Effect Size} = \frac{\text{Treatment group mean} - \text{Control group mean}}{\text{Control group standard deviation}}$$

- **Effect sizes:**
 - around 0.2 is small
 - around 0.5 is moderate
 - around 0.8 is large
- Effect sizes above 0.25 are large enough to be considered educationally significant

Meta-analysis

(Wilson, n.d., Overview, Slide 7)

(Waxman, Lin, & Michko, 2003)

(Kulik, 2003)



An Example

A Meta-Analysis of the Effectiveness of Teaching and Learning With Technology on Student Outcomes.

Abstract

To estimate the effects of teaching and learning with technology on students' cognitive, affective, and behavioral outcomes of learning, 282 effect sizes were calculated using statistical data from 42 studies that contained a combined sample of approximately 7,000 students. The mean of the study-weighted effect sizes averaging across all outcomes was .410 ($p < .001$), with a 95-percent confidence interval (CI) of .175 to .644. This result indicates that teaching and learning with technology has a small, positive, significant ($p < .001$) effect on student outcomes when compared to traditional instruction. The mean study-weighted effect size for the 29 studies containing cognitive outcomes was .448, and the mean study-weighted effect size for the 10 comparisons that focused on student affective outcomes was .464. On the other hand, the mean study-weighted effect size for the 3 studies that contained behavioral outcomes was -.091, indicating that technology had a small, negative effect on students' behavioral outcomes. The overall study-weighted effects were constant across the categories of study characteristics, quality of study indicators, technology characteristics, and instructional/teaching characteristics.

Meta-analysis



Strengths of Meta-Analysis

- **Combination of multiple studies increases the strength of research findings**
- **An objective, reliable, repeatable, and measurable approach to summarizing research findings**
- **Meta-analysis can highlight “gaps in the research literature, providing a solid foundation for the next generation of research on that topic.”**
- **Where there are conflicting findings, meta-analysis can summarize studies and show trends**
- **Ability to examine differential effects of the treatment**

Meta-analysis

(Wilson, n.d. Overview, Slide 11)

(Bangert-Drowns, 1991)



Weaknesses

- Takes a lot of effort (mathematical analysis)
- Results are correlational, not inherently causal
- Quality of results depends on quality of studies
- Potential problems in choosing which studies to include:
 - Selection bias
 - Comparing apples to apples
 - Negative & null studies underreport
 - Inclusion of studies with minor problems
- Primarily quantitative
(loss of qualitative distinctions)

Meta-analysis



Questions:

- Suppose a study reports an effect size of $+0.29$
Is this significant?
- Suppose a study reports an effect size of -0.17
Is this significant?
- Suppose a study reports an effect size of $+0.64$
Is this significant?
- Does NCLB data compare “apples to apples”?

Meta-analysis



Definition revisited...

A statistical method designed to increase the objectivity, reliability, and precision of research by analyzing and synthesizing the results of all the available high quality evidence carried out on a specified topic. Weighted effect sizes are calculated for each study.

Meta-analysis



References

Bangert-Drowns, R. L. (1991). Meta-Analysis in Educational Research. Retrieved October 13, 2004 from Eric Digest web site:

http://www.ericfacility.net/databases/ERIC_Digests/ed339748.html

Basu, A. (n.d.) How to Conduct a Meta-Analysis. Retrieved October 12, 2004 from University of Pittsburg web site: <http://www.pitt.edu/~super1/lecture/lec1171/001.htm>

Kulik, J. A. (2003) Effects of Using Instructional Technology in Elementary and Secondary Schools: What Controlled Evaluation Studies Say. Retrieved 10/23/04 from SRI web site:

<http://www.sri.com/policy/csted/reports/sandt/it/>

Waxman, H. C., Lin, M., Michko, G. M. (2003). A Meta-Analysis of the Effectiveness of Teaching and Learning With Technology on Student Outcomes. Retrieved October 17, 2004 from North Central Regional Laboratory web site:

<http://www.ncrel.org/tech/effects2/>

Wilson, D. B. (n.d.). Meta-analysis stuff [PowerPoint Files]. Retrieved October 12, 2004 from George Mason University web site: <http://mason.gmu.edu/~dwilsonb/ma.html>

Meta-analysis