The Effect Size

- The effect size (ES) makes meta-analysis possible.
- The ES encodes the selected research findings on a numeric scale.
- There are many different types of ES measures, each suited to different research situations.
- Each ES type may also have multiple methods of computation.

Examples of Different Types of Effect Sizes: The Major Leagues

- Standardized Mean Difference
  - group contrast research
    - treatment groups
    - naturally occurring groups
  - inherently continuous construct
- Odds-Ratio
  - group contrast research
    - treatment groups
    - naturally occurring groups
  - inherently dichotomous construct
- Correlation Coefficient
  - association between variables research

Examples of Different Types of Effect Sizes: Two from the Minor Leagues

- Proportion
  - central tendency research
    - HIV/AIDS prevalence rates
    - Proportion of homeless persons found to be alcohol abusers
- Standardized Gain Score
  - gain or change between two measurement points on the same variable
    - reading speed before and after a reading improvement class

What Makes Something an Effect Size for Meta-Analytic Purposes

- The type of ES must be comparable across the collection of studies of interest.
- This is generally accomplished through standardization.
- Must be able to calculate a standard error for that type of ES
  - the standard error is needed to calculate the ES weights, called inverse variance weights (more on this latter)
  - all meta-analytic analyses are weighted

The Standardized Mean Difference

$$\overline{ES} = \frac{\overline{X}_{G1} - \overline{X}_{G2}}{s_{pooled}} \qquad s_{pooled} = \sqrt{\frac{s_1^2(n_1 - 1) + s_2^2(n_2 - 1)}{n_1 + n_2 - 2}}$$

- Represents a standardized group contrast on an *inherently continuous* measure.
- Uses the pooled standard deviation (some situations use control group standard deviation).
- Commonly called “d” or occasionally “g”.

The Correlation Coefficient

$$\overline{ES} = r$$

- Represents the strength of association between two *inherently continuous* measures.
- Generally reported directly as “r” (the Pearson product moment coefficient).

The Odds-Ratio

- The Odds-Ratio is based on a 2 by 2 contingency table, such as the one below.

	Frequencies	
	Success	Failure
Treatment Group	*a*	*b*
Control Group	*c*	*d*

$$\overline{ES} = \frac{ad}{bc}$$

- The Odds-Ratio is the odds of success in the treatment group relative to the odds of success in the control group.

Methods of Calculating the Standardized Mean Difference

- The standardized mean difference probably has more methods of calculation than any other effect size type.

The different formulas represent degrees of approximation to the ES value that would be obtained based on the means and standard deviations

Great
- direct calculation based on means and standard deviations
- algebraically equivalent formulas (t-test)
- exact probability value for a t-test
- approximations based on continuous data (correlation coefficient)

Good
- estimates of the mean difference (adjusted means, regression B weight, gain score means)
- estimates of the pooled standard deviation (gain score standard deviation, one-way ANOVA with 3 or more groups, ANCOVA)

Poor
- approximations based on dichotomous data

Effect Size Overheads 9

Methods of Calculating the Standardized Mean Difference

Direction Calculation Method

$$ES = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{s_1^2(n_1-1) + s_2^2(n_2-1)}{n_1+n_2-2}}} = \frac{\overline{X}_1 - \overline{X}_2}{s_{pooled}}$$

Effect Size Overheads 10

Methods of Calculating the Standardized Mean Difference

Algebraically Equivalent Formulas:

$$ES = t\sqrt{\frac{n_1+n_2}{n_1 n_2}}$$ independent t-test

$$ES = \sqrt{\frac{F(n_1+n_2)}{n_1 n_2}}$$ two-group one-way ANOVA

exact *p-values* from a *t*-test or *F*-ratio can be converted into *t*-value and the above formula applied

Effect Size Overheads 11

Methods of Calculating the Standardized Mean Difference

A study may report a grouped frequency distribution from which you can calculate means and standard deviations and apply to direct calculation method.

Effect Size Overheads 12

Methods of Calculating the Standardized Mean Difference

Close Approximation Based on Continuous Data -- Point-Biserial Correlation. For example, the correlation between treatment/no treatment and outcome measured on a continuous scale.

$$ES = \frac{2r}{\sqrt{1-r^2}}$$

Methods of Calculating the Standardized Mean Difference

Estimates of the Numerator of ES -- The Mean Difference

-- difference between gain scores

-- difference between covariance adjusted means

-- unstandardized regression coefficient for group membership

Methods of Calculating the Standardized Mean Difference

Estimates of the Denominator of ES -- Pooled Standard Deviation

$s_{pooled} = se\sqrt{n-1}$ standard error of the mean

Methods of Calculating the Standardized Mean Difference

Estimates of the Denominator of ES -- Pooled Standard Deviation

$$s_{pooled} = \sqrt{\frac{MS_{between}}{F}}$$ one-way ANOVA >2 groups

$$MS_{between} = \frac{\sum \bar{X}_j^2 n_j - \frac{(\sum \bar{X}_j n_j)^2}{\sum n_j}}{k-1}$$

Methods of Calculating the Standardized Mean Difference

Estimates of the Denominator of ES --
Pooled Standard Deviation

$$s_{pooled} = \frac{s_{gain}}{\sqrt{2(1-r)}}$$

standard deviation of gain scores, where r is the correlation between pretest and posttest scores

Methods of Calculating the Standardized Mean Difference

Estimates of the Denominator of ES --
Pooled Standard Deviation

$$s_{pooled} = \sqrt{\frac{MS_{error}}{1-r^2} \cdot \frac{df_{error}-1}{df_{error}-2}}$$

ANCOVA, where r is the correlation between the covariate and the DV

Methods of Calculating the Standardized Mean Difference

Estimates of the Denominator of ES --
Pooled Standard Deviation

$$s_{pooled} = \sqrt{\frac{SS_B + SS_{AB} + SS_W}{df_B + df_{AB} + df_W}}$$

A two-way factorial ANOVA where B is the irrelevant factor and AB is the interaction between the irrelevant factor and group membership (factor A).

Methods of Calculating the Standardized Mean Difference

Approximations Based on Dichotomous Data

$$ES = probit(p_{group_1}) - probit(p_{group_2})$$

the difference between the probits transformation of the proportion successful in each group

converts proportion into a z-value

Methods of Calculating the Standardized Mean Difference

Approximations Based on Dichotomous Data

$$ES = 2\sqrt{\frac{c^2}{N - c^2}}$$

chi-square must be based on a 2 by 2 contingency table (i.e., have only 1 df)

$$ES = \frac{2r}{\sqrt{1 - r^2}}$$

phi coefficient

Data to Code Along with the ES

- The Effect Size
  - may want to code the data from which the ES is calculated
  - confidence in ES calculation
  - method of calculation
  - any additional data needed for calculation of the inverse variance weight
- Sample Size
- ES specific attrition
- Construct measured
- Point in time when variable measured
- Reliability of measure
- Type of statistical test used

Interpreting Effect Size Results

- **Cohen's "Rules-of-Thumb"**
  - **standardized mean difference effect size**
    - **small = 0.20**
    - **medium = 0.50**
    - **large = 0.80**
  - **correlation coefficient**
    - **small = 0.10**
    - **medium = 0.25**
    - **large = 0.40**
  - **odds-ratio**
    - **small = 1.50**
    - **medium = 2.50**
    - **large = 4.30**

Interpreting Effect Size Results

- Rules-of-Thumb do not take into account the context of the intervention
  - a "small" effect may be highly meaningful for an intervention that requires few resources and imposes little on the participants
  - small effects may be more meaningful for serious and fairly intractable problems
- Cohen's Rules-of-Thumb do, however, correspond to the distribution of effects across meta-analyses found by Lipsey and Wilson (1993)

Translation of Effect Sizes

- **Original metric**
- **Success Rates (Rosenthal and Rubin's BESD)**
  - **Proportion of "successes" in the treatment and comparison groups assuming an overall success rate of 50%**
  - **Can be adapted to alternative overall success rates**
- **Example using the sex offender data**
  - **Assuming a comparison group recidivism rate of 15%, the effect size of 0.45 for the cognitive-behavioral treatments translates into a recidivism rate for the treatment group of 7%**

Methodological Adequacy of Research Base

- Findings must be interpreted within the bounds of the methodological quality of the research base synthesized.
- Studies often cannot simply be grouped into "good" and "bad" studies.
- Some methodological weaknesses may bias the overall findings, others may merely add "noise" to the distribution.

Confounding of Study Features

- Relative comparisons of effect sizes across studies are inherently correlational!
- Important study features are often confounding, obscuring the interpretive meaning of observed differences
- If the confounding is not severe and you have a sufficient number of studies, you can model "out" the influence of method features to clarify substantive differences

Concluding Comments

- Meta-analysis is a replicable and defensible method of synthesizing findings across studies
- Meta-analysis often points out gaps in the research literature, providing a solid foundation for the next generation of research on that topic
- Meta-analysis illustrates the importance of replication
- Meta-analysis facilitates generalization of the knowledge gain through individual evaluations