Interpretation of Effect Size Statistics

Marc J. Diener, PhD

Learning Objectives

- Understand the difference between statistical significance and practical significance
- Identify the two basic types of effect sizes
- Understand two general principles for interpreting effect sizes
- Identify Cohen's (1988) benchmarks for interpreting effect sizes
- Understand the development of empirically-derived benchmarks for interpreting effect sizes

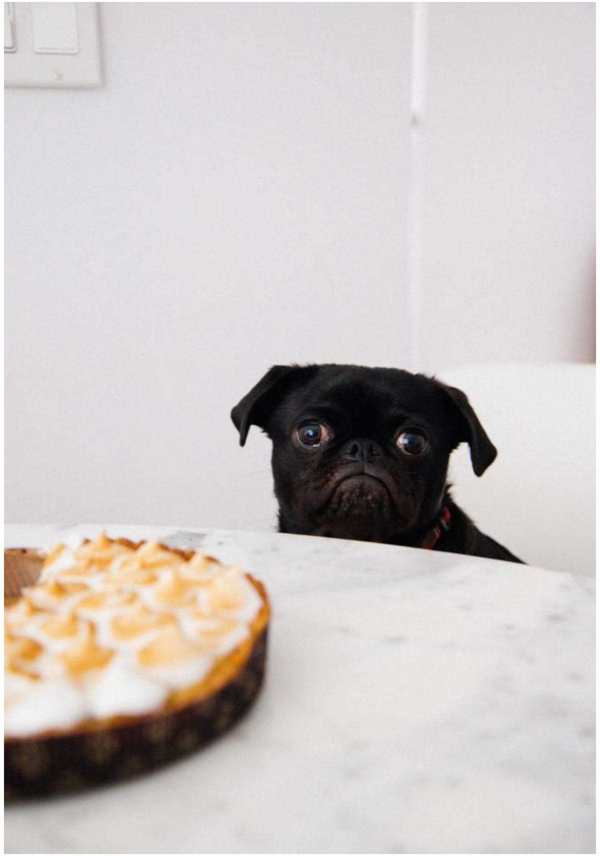
Ultra-Brief Overview of Effect Size (ES): Concepts and Basic Effect Size Types

Conceptual Underpinning of Effect Sizes (ESs): *p*-values and Effect Sizes Answer Difference Questions

(Field, 2013; Lipsey & Wilson, 2001; Urdan, 2005)

- ***p*-value =** "The probability of obtaining a statistic of a given size by chance, or due to random error" (Urdan, 2005, p.73)

- **Effect Size** = A statistic that indicates the (a) *direction* **and** (b) *magnitude* of quantitative research findings (Lipsey & Wilson, 2001)



Problems with *p*-values

(Field, 2013; Rosenthal, 1984)

- Lead to common errors in interpreting results, e.g.,
  - Statistically significant results mean that the findings are *important*
  - Results that are not statistically significant mean that the null hypothesis should be accepted
  - Statistically significant results indicate that the null hypothesis is false

Advantages of ESs

(Borenstein et al., 2009; Field, 2013)

- Allow direct measurement of the size of the effect
- They are usually standardized--> permits comparison of ESs across different studies that have
  - Different variables

    OR
  - Different scales of measurement
- Necessary for conducting meta-analyses

Global vs. Local Effect Sizes

(Peugh, 2010)

- Global Effect Size
  - Measures variance in the dependent variable that can be explained using ***all*** predictor variables
- Local Effect Size
  - Measures variance in the dependent variable that can be explained using an ***individual*** variable

Basic ES Types

(Ellis, 2010)

- Group differences: The *d* family
  - Group differences on dichotomous outcomes (e.g., odds ratio)
  - Group differences on continuous outcomes (e.g., Cohen's *d* or Hedges's *g*)

- Strength of association between variables: The *r* family
  - Correlation indexes (e.g., r or r_s)
  - Proportion of variance indexes (e.g., r^2 or R^2)

Vexing Questions about ESs

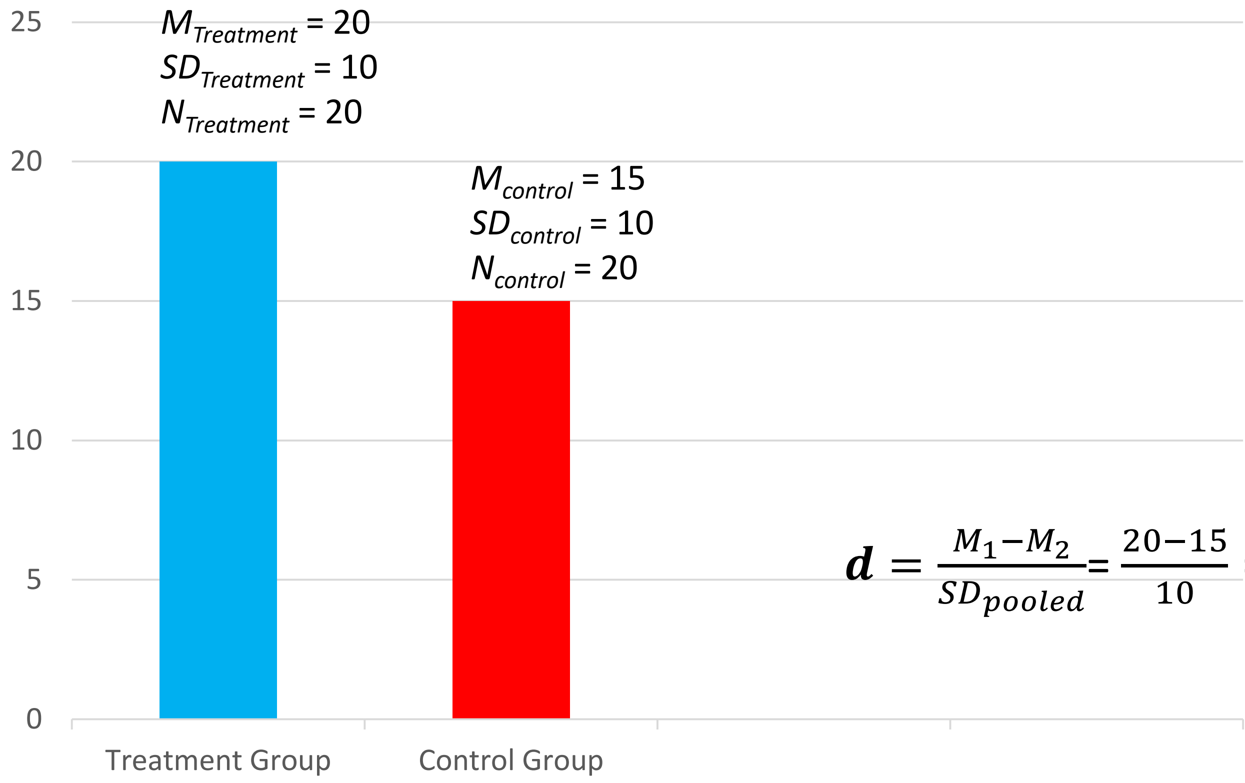
- What do they actually mean?

- How do you interpret them?

Plain English Explanation of the *d* Family

(Ellis, 2010; Lipsey & Wilson, 2001)

- Group comparisons of continuous outcomes
  - $\boldsymbol{d} = \frac{M_1 - M_2}{SD_{pooled}}$
    - Use when standard deviations for the two groups are similar
  - **Hedges' *g*** : Corrects *d* for potential bias of small sample sizes
    - Use when sample sizes of the two groups are quite different
  - **Glass's △** : Standardizes the mean difference score by the control group
    - Use when the standard deviations for the two groups are different



$$\boldsymbol{d} = \frac{M_1 - M_2}{SD_{pooled}} = \frac{20-15}{10} = \frac{5}{10} = 0.5$$

Plain English Explanation of the *r* Family

(Ellis, 2010; Lipsey & Wilson, 2001)

- Correlation Indexes
  - r : Used when both X and Y are continuous variables
  - r_s : Used when both X and Y are measured with ranked scales
  - r_{pb} : Used when X is dichotomous and Y is continuous
  - Phi coefficient: Used when both X and Y are dichotomous variables

Plain English Explanation of the *r* Family

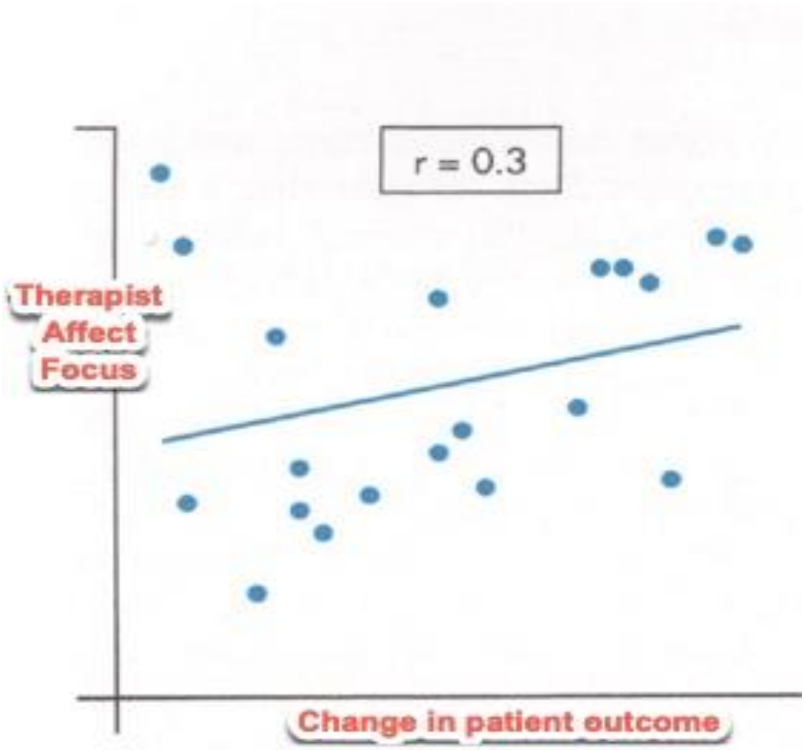
(cont.; Ellis, 2010; Lipsey & Wilson, 2001)

- Proportion of shared variance Indexes
  - r^2 : Coefficient of determination
  - R^2 : Coefficient of multiple determination
  - f (Cohen's f): dispersion of means in more than 2 groups
  - f^2 (Cohen's f-squared): Alternative to R^2 or $\triangle R^2$
  - Eta-squared: Correlation ratio
  - Omega-squared: Corrects eta-squared for potential inflation

Interpretation of Effect Sizes: General Principles

(Cooper, 2017; Rosenthal, 1984)

- Interpret based on context
  - Compare ESs to those obtained by previous meta-analytic research
  - Compare ESs to those obtained by previous researchers in primary studies



Adapted from Figure 13.5A in *Biostatistics: The bare essentials* (4th ed.; p. 146), by G. R. Norman and D. L. Streiner, 2014, Shelton, CT: People's Medical Publishing House.


evian

Interpretation of Effect Sizes: General Principles

(cont.; Cooper, 2017; Rosenthal, 1984)

- Consider the nature of the dependent variable

Cohen's Benchmarks

(Cohen, 1988; Cooper, 2017)

- *d*
  - 0.20 = Small
  - 0.50 = Medium
  - 0.80 = Large
- *r*
  - .10 = Small
  - .30 = Medium
  - .50 = Large

STOP!

DON'T USE THESE UNLESS YOU HAVE NO OTHER BETTER OPTION!!

Empirically Derived Benchmarks

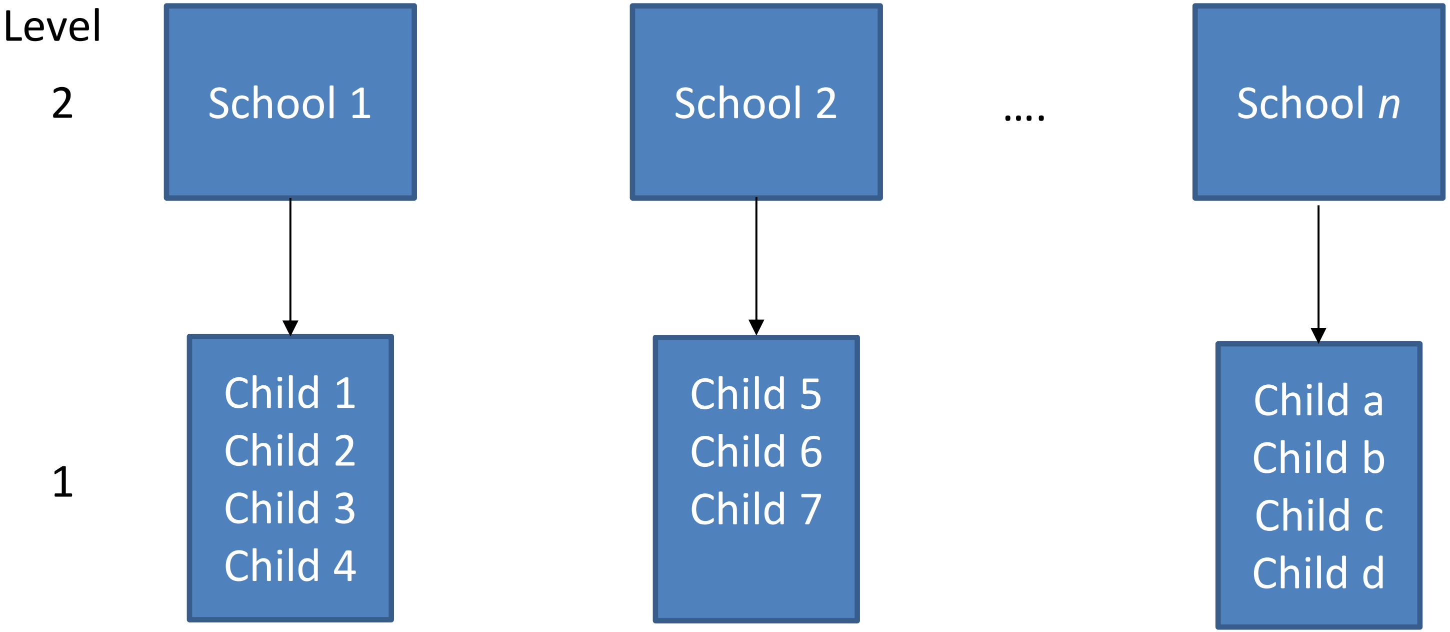
Psychological Assessment and Treatment

(Hemphill, 2003)

Distribution of correlation coefficients	Assessment Review	Treatment Review	Combined Reviews	Empirical Guidelines
Lower third	.02 to .21	-.08 to .17	-.08 to .17	< .20
Middle third	.21 to .33	.17 to .28	.18 to .29	.20 to .30
Upper third	.35 to .78	.29 to .60	.30 to .78	> .30

Advanced Statistical Issues: Effect Sizes in Multi-Level Models

(Peugh, 2010; Selya et al., 2012)



Effect Sizes for Multi-Level Models

(Peugh, 2010; Selya et al., 2012)

- Problems with using an effect size from the *d* family
  - The independent variable needs to be categorical
  - Global vs. local effect sizes

(cont.)

- Research Questions
  - Level 1: Effect of student SES on student science achievement
  - Level 2: Effect of student-to-teacher ratio on school-level science achievement
  - Interaction: Is the association between SES and science achievement moderated by student-teacher ratio

Parameters	**No Predictors**
Regression coefficients (fixed effects)	Coefficient (*SE*)
Intercept (γ_{00}) [Average science achievement score]	18.90 (.07)**
Variance Components (random effects)	
Residual (σ^2) [Variance in achievement across students within a school]	18.67 (.25)**
Intercept (τ_{00}) [Variance in achievement across schools]	4.18 (.28)**

Parameters	No Predictors	Level-1: random
Regression coefficients (fixed effects)	Coefficient (*SE*)	Coefficient (*SE*)
Intercept (γ_{00}) [Average science achievement score]	18.90 (.07)**	18.89 (.07)**
Student SES (γ_{10})	--	2.00 (.07)**
Variance Components (random effects)		
Residual (σ^2) [Variance in achievement across students within a school]	18.67 (.25)**	16.97 (.24)**
Intercept (τ_{00}) [Variance in achievement across schools]	4.18 (.28)**	4.45 (.28)**
Slope (τ_{11}) [Variance in effect of SES on achievement across schools]	--	.54 (.21)**

Parameters	No Predictors	Level-1: random	Interaction
Regression coefficients (fixed effects)	Coefficient (*SE*)	Coefficient (*SE*)	Coefficient (*SE*)
Intercept (γ_{00}) [Average science achievement score]	18.90 (.07)**	18.89 (.07)**	18.90 (.07)**
Student SES (γ_{10})	--	2.00 (.07)**	2.00 (.07)**
Student-to-Teacher ratio (γ_{01})	--	--	-.10 (.01)**
Interaction (γ_{11})	--	--	-.04 (.02)**
Variance Components (random effects)			
Residual (σ^2) [Variance in achievement across students within a school]	18.67 (.25)**	16.97 (.24)**	16.97 (.24)**
Intercept (τ_{00}) [Variance in achievement across schools]	4.18 (.28)**	4.45 (.28)**	4.15 (.27)**
Slope (τ_{11}) [Variance in effect of SES on achievement across schools]	--	.54 (.21)**	.49 (.21)**

Global Effect Size: Pseudo-R^2

(Peugh, 2010)

- Use regression coefficients to obtain predicted science achievement score for each student
- Correlate the observed and predicted science achievement scores
-
-

Participant	Predicted Score	Actual Score
1	25	30
2	20	18
3	17	27
---	---	---

Local Effect Size: Proportional Reduction in Variance

(Peugh, 2010)

$$\text{PRV} = (var_{\text{NoPredictor}} - var_{\text{predictor}}) \,/\, var_{NoPredictor}$$

PRV from Adding SES (Peugh, 2010)

$$(18.67 - 16.97) / 18.67 = .09$$

Variance Components (random effects)			
Residual (σ^2) [Variance in achievement across students within a school]	18.67 (.25)**	16.97 (.24)**	16.97 (.24)**
Intercept (τ_{00}) [Variance in achievement across schools]	4.18 (.28)**	4.45 (.28)**	4.15 (.27)**
Slope (τ_{11}) [Variance in effect of SES on achievement across schools]	--	.54 (.21)**	.49 (.21)**

PRV from Adding Student-to-Teacher Ratio (Peugh, 2010)

$$(4.45 – 4.15) / 4.45 = .07$$

Variance Components (random effects)			
Residual (σ^2) [Variance in achievement across students within a school]	18.67 (.25)**	16.97 (.24)**	16.97 (.24)**
Intercept (τ_{00}) [Variance in achievement across schools]	4.18 (.28)**	4.45 (.28)**	4.15 (.27)**
Slope (τ_{11}) [Variance in effect of SES on achievement across schools]	--	.54 (.21)**	.49 (.21)**

PRV from Adding Interaction (Peugh, 2010)

$$(.54 - .49) / .54 = .09$$

Variance Components (random effects)			
Residual (σ^2) [Variance in achievement across students within a school]	18.67 (.25)**	16.97 (.24)**	16.97 (.24)**
Intercept (τ_{00}) [Variance in achievement across schools]	4.18 (.28)**	4.45 (.28)**	4.15 (.27)**
Slope (τ_{11}) [Variance in effect of SES on achievement across schools]	--	.54 (.21)**	.49 (.21)**

References

Borenstein, M., Hedges, L.V., Higgins, J.P.T., & Rothstein, H.R. (2009). *Introduction to meta-analysis*. Chichester, UK: John Wiley & Sons, Ltd.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cooper, H. (2017). *Research synthesis and meta-analysis: A step-by-step approach* (5th ed.). Thousand Oaks, CA: Sage publications.

Ellis, P. D. (2010). *The essential guide to effect sizes: Statistical power, meta-analysis, and the interpretation of research results*. Cambridge University Press.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th ed.). Thousand Oaks, CA: Sage publications.

Hemphill, J. F. (2003). Interpreting the magnitudes of correlation coefficients. *American Psychologist*, *58*, 78–79. https://doi.org/10.1037/0003-066X.58.1.78

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage publications.

Peugh, J. L. (2010). A practical guide to multilevel modeling. *Journal of School Psychology*, 48, 85-112.

Rosenthal, R. (1984). *Meta-analytic procedures for social research*. Beverly Hills, CA: Sage publications.

Selya, A. S., Rose, J. S., Dierker, L. C., Hedeker, D., & Mermelstein, R. J. (2012). A practical guide to calculating Cohen's f^2, a measure of local effect size, from PROC MIXED. *Frontiers in Psychology*, *3*, 1-6.

Urdan, T. C. (2005). *Statistics in plain English* (2nd ed.). Mahwah, NJ: Erlbaum.