WORK-LEARNING RESEARCH

How to calculate effect sizes from published research: A simplified methodology

Will Thalheimer
Samantha Cook

How to calculate effect sizes
from published research articles:
A simplified methodology

Will Thalheimer   Samantha Cook
Work-Learning Research   *Harvard University*

Overview

This article provides a simplified methodology for calculating Cohen's *d* effect sizes from published experiments that use *t*-tests and F-tests. Accompanying this article is a Microsoft Excel Spreadsheet to speed your calculations. Both the spreadsheet and this article are available as free downloads at www.work-learning.com/effect_sizes.htm.

Why we use effect sizes

Whereas statistical tests of significance tell us the likelihood that experimental results differ from chance expectations, effect-size measurements tell us the relative magnitude of the experimental treatment. They tell us the *size* of the experimental *effect*. Effect sizes are especially important because they allow us to compare the magnitude of experimental treatments from one experiment to another. Although percent improvements can be used to compare experimental treatments to control treatments, such calculations are often difficult to interpret and are almost always impossible to use in fair comparisons across experimental paradigms.

A simple methodology

Although extensive articles have been written detailing methods for calculating effect sizes from published research articles (e.g., Rosnow & Rosenthal, 1996; Rosnow, Rosenthal, & Rubin, 2000), at least some of us—the first author included—require a simpler approach. This article provides a method to calculate Cohen's *d* from both t-tests and some F-tests of significance. Accompanying this article is a Microsoft Excel Spreadsheet that can be used to compute Cohen's *d* from published data.

Cohen's *d* has two advantages over other effect-size measurements. First, its burgeoning popularity is making it the standard. Thus, its calculation enables immediate comparison

to increasingly larger numbers of published studies. Second, Cohen's (1992) suggestion that effect sizes of .20 are small, .50 are medium, and .80 are large enables us to compare an experiment's effect-size results to known benchmarks. The simple methodology offered below is not new but is drawn from previously published articles, most notably Rosnow and Rosenthal (1996) and Rosnow, Rosenthal, and Rubin (2000). We have simplified the methodology not by changing the formulas and calculations but by discarding as much as possible the jargon and computational rationales typically included in articles written for research audiences. This article is an attempt to provide a practical methodology to enable the calculation of effect sizes.

What is an effect size?

In essence, an effect size is the difference between two means (e.g., treatment minus control) divided by the standard deviation of the two conditions. It is the division by the standard deviation that enables us to compare effect sizes across experiments. Because t-tests and F-tests utilize different measures of standard deviation, two separate calculations are required. You will find it useful to keep this distinction in mind as you read this document and utilize the accompanying spreadsheet.

Table of Contents

Calculating Cohen's *d* from t-tests

(1) $$d = \frac{\overline{x}_t - \overline{x}_c}{s_{pooled}}$$

Key to symbols:
d = Cohen's *d* effect size
$\overline{x}$ = mean (average of treatment or comparison conditions)
s = standard deviation
Subscripts: *t* refers to the treatment condition and *c* refers to the comparison condition (or control condition).

How to calculate:
The article should list the means ($\overline{x}$) of the treatment condition and the comparison condition. Use those numbers in the formula and calculate the pooled standard deviation by using Formula 1a below. After you use Formula 1a, simply finish calculating Formula 1 to get Cohen's *d*.

(1a) $$s_{pooled} = \sqrt{\frac{(n_t - 1)s_t^{\,2} + (n_c - 1)s_c^{\,2}}{n_t + n_c}}$$

Key to symbols:
s = standard deviation
n = number of subjects
Subscripts: *t* refers to the treatment condition and *c* refers to the comparison condition (or control condition).

How to calculate:
The article should list the number of subjects (*n*) and the standard deviations (*s*) of the treatment condition and the comparison condition. Use those numbers to make your calculations. If the article does not list the standard deviations, use either Formula 2 or Formula 3 below if possible.

4

Calculating Cohen's d from t-tests:
When you don't have standard deviations or standard errors.

When an experiment that uses a t-test does not list standard deviations, you can calculate Cohen's **d** as follows using the t statistic:

$$(2) \qquad d = t \sqrt{\left(\frac{n_t + n_c}{n_t\, n_c}\right)\left(\frac{n_t + n_c}{n_t + n_c - 2}\right)}$$

Key to symbols:
  d = Cohen's **d** effect size
  t = **t** statistic
  n = number of subjects
  Subscripts: **t** refers to the treatment condition and **c** refers to the comparison condition (or control condition).

How to calculate:
  The article should list the **t** statistic, which it will usually do, for example, with the following notation: **t** (29) = 3.12, where 29 is the degrees of freedom and 3.12 is the **t** statistic. The article should also list the number of subjects (**n**) within each condition. Use those numbers to make your calculations. If the article does not list the number of subjects in each condition but does list the total number of subjects—and if you can assume that both conditions have roughly equal numbers of subjects—you can estimate Cohen's **d** by using Formula 2a below.

  Warning: Some studies using repeated-measure designs (where each subject is measured several times within the same condition) incorrectly use experimental trials, instead of subjects, as the units of analysis. The formulas on this page cannot be used for these studies because the t-statistic is not relevant to the number of subjects (**n**) in the study. These studies are often easy to spot because they have outrageously high degrees of freedom.

$$(2a) \qquad d \approx \frac{2t}{\sqrt{n-2}}$$

Calculating Cohen's *d* from t-tests:
When you have standard errors instead of standard deviations.

When an experiment that uses a t-test does not list standard deviations but does list standard errors (SE), you can calculate the standard deviations as follows and then use the resulting numbers in Formula 1a:

$$(3) \qquad s = SE\sqrt{n}$$

Key to symbols:
  s = standard deviation
  SE = standard error
  n = number of subjects

How to calculate:
  This formula assumes that the article lists the standard error (SE) and number of subjects (n) within each condition. Use those numbers to make your calculations.

Calculating Cohen's *d* from F-tests

(4)
$$d = \frac{\overline{x}_t - \overline{x}_c}{\sqrt{MSE\left(\dfrac{n_t + n_c - 2}{n_t + n_c}\right)}}$$

Key to symbols:
d = Cohen's *d* effect size
$\overline{x}$ = mean (average of treatment or comparison condition)
n = number of subjects
MSE = mean squared error
Subscripts: *t* refers to the treatment condition and *c* refers to the comparison condition (or control condition).

How to calculate:
If standard deviations are available, use Formulas 1 and 1a above because *MSE*'s will not produce a precise Cohen's *d* when the F-test is a comparison among more than two conditions. Otherwise, continue.

The article should list the means ($\overline{x}$) of the treatment condition and the comparison condition, and the mean squared error (*MSE*). Use those numbers in the formula to get Cohen's *d*. Be careful to select the correct *MSE* if many are listed. Note that only when the F-test numerator degrees of freedom are equal to 1—when the F-test compares one condition to one other condition—will the *MSE* produce an exact Cohen's *d* effect size. In this case, the F-test is equivalent to a t-test. Selecting other *MSE*'s may not produce valid results.

Calculating Cohen's *d* from F-tests:
When you don't have MSE's.

When an experiment that uses an F-test does not list the MSE, you can calculate Cohen's *d* as follows using the F statistic. This calculation should only be used when the F-test compares one condition to one other condition.

$$(5) \qquad d = \sqrt{F\left(\frac{n_t + n_c}{n_t \, n_c}\right)\left(\frac{n_t + n_c}{n_t + n_c - 2}\right)}$$

Key to symbols:
  d = Cohen's *d* effect size
  F = F statistic
  n = number of subjects
  Subscripts: *t* refers to the treatment condition and *c* refers to the comparison condition (or control condition).

How to calculate:
  This formula can ONLY be used when the F-test compares two conditions (when the first degrees of freedom is equal to one). The article should list the *F* statistic, which it will usually do, for example, with the following notation: $F(1,39) = 3.12$, where 1 is the degrees of freedom based on the number of conditions, and 39 is the degrees of freedom based on the number of subjects. The article should also list the number of subjects *(n)* within each condition.

References

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*, 155-159.

Rosnow, R. L., & Rosenthal, R. (1996). Computing contrasts, effect sizes, and counternulls on other people's published data: General procedures for research consumers. *Psychological Methods, 1*, 331-340.

Rosnow, R. L., Rosenthal, R., & Rubin, D. B. (2000). Contrasts and correlations in effect-size estimation. *Psychological Science, 11*, 446-453.

How to cite this article

Thalheimer, W., & Cook, S. (2002, August). *How to calculate effect sizes from published research articles: A simplified methodology.* Retrieved November 31, 2002 from http://work-learning.com/effect_sizes.htm.

(NOTE: You should replace the fictional November 31 date with the date on which the article was downloaded.)

Acknowledgements