



## What does effect size tell you?

[Saul McLeod](#) (2019)

Statistical significance is the least interesting thing about experimental results.

We should interpret experimental results with measures of magnitude – for example, not just, whether a treatment affects people, but how much it affects them.

### What is effect size?

Effect size is a quantitative measure of the study's effect. The larger the effect size, the more powerful the study.

You can look at the effect size when comparing two groups to see how substantially different they are. In this case, the effect size is a quantification of the difference between two group means.

You can also look at the effect size when examining the correlation between two variables. In this case, the effect size is a quantification of the strength of the association between two variables.

# Cohen's *d*

Cohen's *d* is an appropriate effect size measure for comparing two group means. For example, Cohen's *d* can accompany the reporting of *t*-test and ANOVA results. Cohen's *d* is also widely used in meta-analysis, which is the practice of combining data from multiple studies.

To calculate Cohen's *d*, which is the standardized mean difference between two groups, subtract the mean of one group from the mean of the other group ( $M1 - M2$ ) and divide the result by the pooled standard deviation (SD).

Relative size	Effect size	% of control group below the mean of experimental group
	0.0	50%
Small	0.2	58%
Medium	0.5	69%
Large	0.8	79%
	1.4	92%

A Cohen's *d* of 1.000 indicates that the means of the two groups differ by 1.000 pooled standard deviation (or one z-score). A Cohen's *d* of 2.00 indicates that the means of two groups differ by 2.000 pooled standard deviations, and so on.

Cohen suggested that a Cohen's *d* of 0.200 be considered a 'small' effect size, a Cohen's *d* of 0.500 be considered a 'medium' effect size, and a Cohen's *d* of 0.800 be considered a 'large' effect size. Therefore, if two groups' means don't differ by more than 0.200 standard deviations, the difference is considered trivial, even if it is statistically significant.

# Why report effect sizes?

## The $p$ -value is not enough

A lower  $p$ -value is sometimes interpreted as meaning there is a stronger relationship between two variables. However, [statistical significance](#) (at the  $p < .05$  level) means that it is unlikely that the null hypothesis is true (less than 5%).

Therefore, a significant  $p$ -value might suggest whether an intervention works, but an effect size tells us how powerfully it works.

Emphasizing the size of an effect promotes a more scientific approach, because, unlike significance tests, effect size is independent of sample size.

## To compare the results of studies done in different settings

Unlike a  $p$ -value, an effect size can be used to quantitatively compare the results of studies conducted in different labs by different researchers at different times. A quantitative measure of effect size is widely used in meta-analysis, which is the practice of combining data from multiple studies.