

Understanding Effect Size Measures

Types of Effect Size Measures

At the simplest level, effect size measures come in at least two types which differ in purpose:

1. Correlational indices: Statistics that communicate the size of the relationship (i.e., correlation) between two variables.
2. Standardized difference indices: Statistics that communicate the size of the differences between two group averages.

Correlation-Based Effect Size Measures

Correlational indices of effect size use a statistic called a correlation (r) to describe the relationship between two variables. For example, if we wanted to know whether high school GPA is related to college GPA, we would calculate a correlation to indicate whether and to what extent these two are related to each other.

The correlation itself is already a standardized measure, meaning that its meaning and interpretation does not change based on how the variables were originally measured. Simply put, a correlation of zero means that the two variables are not related; conversely, any non-zero correlation means that the two variables are related.

The number describes both the direction and size/strength of the correlation. If the sign of the number is positive, then the two variables “go in the same direction” – as one increases so does the other. If the sign is negative, then the two variables “go in opposite directions” – they are inversely related. Generally speaking, correlations farther away from zero are considered stronger relationships.

In some cases, the correlation-based measure is squared before it is presented (like with r^2 or eta-squared). In these cases, the squared correlation is interpreted as the proportion of variance in one variable that is accounted for by the other variable. For example, we could say that high school GPA accounts for some percentage of college GPA. Note that the squared correlation can never be negative and ranges from 0 to 1 (or correspondingly from 0% to 100%).

Standardized Difference Effect Sizes

Standardized differences start with difference between two group means (i.e., averages), but alter that difference to make it more interpretable across various studies and situations. For example, a control group and an experimental group might differ on some variable by four points on average, but unless we know more about the issue under study, we really have no way of knowing if the four-point difference is big or small.

One simple solution to this problem is to divide the average difference (also known as a deviation) by something called the standard deviation (which is a measure of how spread out the scores tend to be). This general calculation procedure produces something typically referred to as Cohen’s d .

As a result, the new number reflects how different the groups are relative to how different people are generally. For example, imagine that the *individuals* in the study tend to differ in their scores by eight points. Therefore, a four-point difference between the *groups on average* is only half of the point difference observed among *individual scores* ($d = .50$).

Notice two important things about this calculation procedure. The resulting number itself can be either positive or negative, depending on the order in which the two means are subtracted; therefore, the meaning of the sign (positive or negative) depends on the study. Second, the farther the number is from zero, the larger the difference is between the two groups.

Interpreting Effect Sizes

Because the two types of measures are standardized, the size of the number can be used to compare different studies, regardless of how the study measured its variables. This point makes it desirable to derive some generally accepted guidelines for determining what is considered a small, medium, or large effect.

One way to determine if these labels are accurate is to look at what tends to happen in real psychological research. In the case of social psychological research, the following numbers of studies tend to fall in the various categories:

Label	Convention for Cohen's <i>d</i>	Convention for Correlation (<i>r</i>)	Convention for Squared Correlation (<i>r</i> ²)	Percentage of Social Psychology Studies
Small	0.20	0.10	0.01	30%
Medium	0.50	0.24	0.06	45%
Large	0.80	0.37	0.14	25%

Importantly, these guidelines may not be true for all fields or topics of study. That is, what constitutes a large effect in one area of study may be only a moderate effect in another. For example, even within social psychology, the average effect ranges from $r = .14$ (a relatively small to medium effect) to $r = .32$ (a relatively moderate to large effect) depending on the topic chosen.