Weak, moderate, or strong effect?

E-mail distributed on 16-09-2022

Dear all,

Reporting an effect size with a statistical test is frequently recommended. The question then often arises how to interpret such effect sizes qualitatively, e.g., in terms of a weak, moderate or strong effect. Below I provide a handy reference to make such interpretations for many effect sizes. Before doing so, however, it is important to establish the type of effect size you're dealing with. Generally, these come in a number of "families" (non-exhaustive):

- Difference of means (e.g., Cohen's d, Glass G)
- Correlational (e.g., Pearson correlation, Kendall's tau, standardized regression coefficients, odds ratio)
- Proportion-of-variance explained (e.g., eta-squared, omega-squared, and their partial analogues)
- Relative information (e.g., Bayes factors, information criteria)

All of these are standardized or normalized in order to be scale-free. They are insensitive to sample size and should be comparable quantitatively/qualitatively between different samples and study designs. Their relative abstractness is also a drawback, however, and in fact I strongly recommend **avoiding abstract effect sizes** if your raw effect can be expressed in **psychologically meaningful units** (e.g., items sold, reaction time, willingness-to-pay, etc)! Among effect sizes, I also recommend choosing a statistic that you understand. A standardized mean difference for example is less abstract than a Bayes factor.

For effect sizes, agreed-upon rules are necessary to decide what qualifies as a small, moderate or strong effect. The R package **effectsize** has collected such rules (and their source reference) for a wide variety of effect sizes and can generate an interpretation using its "interpret_X" functions (e.g., interpret_r). In fact, you **do not even even need R** or to install the package to use these rules. You can simply consult the help pages of the functions directly, which list the rules of thumb.* The package vignette has an excellent introduction on the interpretation of effect sizes with examples for pearson correlation, Cohen's d, and the odds ratio. This comes with the important caution that these rules of thumb are somewhat arbitrary, and may change depending on the subject or field of study (e.g., psychology versus physics)!

Best,

Ben

* Some functions are not in the main help directory but listed under their parent function (e.g., epsilon squared is under omega squared)

--
Ben Meuleman, Ph.D.
Statistician
Swiss Center for Affective Sciences
University of Geneva | Campus Biotech
Chemin des Mines 9 | CH-1202 Genève
ben.meuleman@unige.ch | +41 (0)22 379 09 79