Methodologist’s Corner

A Visitor’s Guide to Effect Sizes *

Statistical Significance Versus Practical (Clinical) Importance of Research Findings

MOHAMMADREZA HOJAT1* and GANG XU2
1Center for Research in Medical Education and Health Care and Department of Psychiatry and Human Behavior, Jefferson Medical College of Thomas Jefferson University; 2The International Cultural Exchange Institute, Eastern Region of Anhui University of China (*Author for correspondence, 1025 Walnut Street, Jefferson Medical College, Philadelphia, PA 19107, USA; E-mail: Mohammadreza.Hojat@Jefferson.edu)

Abstract. Effect Sizes (ES) are an increasingly important index used to quantify the degree of practical significance of study results. This paper gives an introduction to the computation and interpretation of effect sizes from the perspective of the consumer of the research literature. The key points made are:
1. ES is a useful indicator of the practical (clinical) importance of research results that can be operationally defined from being “negligible” to “moderate”, to “important”.
2. The ES has two advantages over statistical significance testing: (a) it is independent of the size of the sample; (b) it is a scale-free index. Therefore, ES can be uniformly interpreted in different studies regardless of the sample size and the original scales of the variables.
3. Calculations of the ES are illustrated by using examples of comparisons between two means, correlation coefficients, chi-square tests and two proportions, along with appropriate formulas.
4. Operational definitions for the ESs are given, along with numerical examples for the purpose of illustration.

Key words: clinical significance, effect size, medical education research, practical importance

I. Estimates of Effect Sizes

Effect size can be considered an index of the extent to which the research hypothesis is true, or the degree to which the findings have practical significance in the study population. In other words, effect size is an index that quantifies the degree

* Based on our experiences in teaching statistics and research methodology to medical and other health professions students, we believe that the first step in evolving to a practitioner of research is to become an informed consumer of research. A better-informed individual can accept or reject the research findings with a better critical view. In this article, we will describe, in a non-technical language, the procedures for calculating the effect size estimate and determining the practical (clinical) significance of research findings as opposed to the regularly reported statistical significance of findings. The conceptual and technical details in this article are not certainly sufficient for the practitioners of research, but may help the consumers of research to better understand and hopefully enable them to critically evaluate the research findings.
to which the study results should be considered negligible, or important, regardless of the size of the study sample.

For the purpose of operational definitions of the magnitudes of treatment differences (usually one or more experimental groups compared to a control group), or group differences in non-experimental designs, and for comparability of findings in different studies (e.g., meta analyses) it is desirable to have a “scale-free” effect size index. Estimates of effect sizes are calculated as standardized differences to serve that purpose. Therefore, effect size estimates can be used to compare treatment effects for different variables in the same study, or for the same or different variables across different studies, regardless of the study sample size and the original scales of the variables. These are important characteristics of any estimate of effect size that also have important implications in meta-analytic studies.

Because of these two important advantages of the effect size estimates (independent of sample size, and scale-free characteristic), some professional research journals recently began to recommend, and some require, that the authors report the effect size estimates of the findings in their submitted articles. In the latest publication manual of the American Psychological Association (APA, 2001), for instance, authors are encouraged to report effect size values in any empirical study.

Effect size estimates can be calculated for many different statistical indices. We have chosen the following three topics because of their frequent use in medical education research: mean differences (t-test), measures of association (correlation coefficient, chi-square), and the difference between two proportions.

1. **Effect size estimate for mean differences**

In this section, we will discuss effect size estimates for the differences between two means. We shall use **ES** to represent effect size estimate from this point forward. The typical inferential statistical method used to examine the statistical significance of the difference between two means is the *t*-test. Three different cases are discussed below.

*a. Comparing Two Independent Samples*

Comparisons of the means for two independent groups, such as an experimental and control groups, are commonly reported in the literature. Calculation of the **ES** in this case is simply the difference between the means of experimental (\(M_e\)) and control (\(M_c\)) groups divided by the standard deviation for the control group (\(\sigma_c\)).

\[
ES = \frac{|M_e - M_c|}{\sigma_c}
\]

This **ES** is sometimes referred to as the “Glass’s effect”. We should point out that there is a dispute among experts about the appropriate denominator in the formula used in experimental designs (see Morris and DeShon, 2002).\(^1\)

One point deserves attention. For the sake of simplicity, we report the absolute value of the **ES** throughout this article, without making any distinction between