Psychology
2011. Vol.2, No.6, 631-632

DOI:10.4236/psych.2011.26096

The Effect-Size: A Simple Methodology for Determining and Evaluating the "Effect-Size"

Marcel Lourel[1], Nicolas Gueguen[2], Alexandra Pascual[3], Farida Mouda[4]
[1]Lecturer in Psychology—Occupational Health, University of Artois, IUFM Nord-Pas de Calais, France;
[2]Lecturer in Social and Cognitive Psychology, University of Bretagne-Sud, IUT de Vannes, France;
[3]Reader in Social Psychology, University of Bordeaux 2, Bordeaux, France;
[4]Psychologist, CREPS Rouen, France.
Email: marcel.lourel@lille.iufm.fr, nicolas.gueguen@univ-ubs.fr,
alexandre.pascual@u-bordeaux2.fr, farida.mouda@yahoo.fr
Received May 1st, 2011; revised July 8th, 2011; accepted August 16th, 2011.

Effect-size measurement is a practice that is gradually encouraged indeed required by psychological and social behavior reviews in addition to the classical test of statistical significance. This paper is written as a methodological note that presents the conceptual interest of the effect-size, the main measurement indicators and their interpretation.

Keywords: Effect-Size, Test of Significance, Statistical Evaluation

The Effect-Size Principle

When a statistical test is carried out such as a test comparing 2 means or an χ^2 independence test, we base our analysis and our interpretation on the probability of the value produced by the test (e.g. the value for the Student *t* for the purpose of a test comparing 2 means or the χ^2 for an independence test). In fact, if the probability seems below or equal to a reference probability well known to any social sciences researcher – the famous p = .05- we conclude that there is a "statistically significant" effect. Whether or not the result is significant will subsequently determine our way of interpreting our data which will consequently directly affect the way in which we theoretically explain this effect. Furthermore, this significance or lack of significance will also, as a consequence, affect our subsequent research work. Most of the articles published are based on works that demonstrate statistically significant effects. In fact, the famous ".05" probability determines what is offered to the scientific community since we rarely come across publications that do not provide the main results with this level of signifyscance. Accordingly, the impact made by this test interpretation value is far from being insignificant.

Why Do We Need to Measure the Effect-Size?

If we query why it is necessary to quantify the effect-size, this requires 2 major parameters to be taken into account for the research: the size of the samples used to quantify the phenomena that we study and the extent of the differences between the variables subject of our analysis. However, in the case of inferential statistics, sample size dependence is the same as for small samples; major differences between variables are required if a statistical difference is to be expressed to within .05 whereas in the case of excessive numbers, minute deviations will be enough to produce differences. Then, effect-size indicators become completely relevant by reinstating their full meaning to the amplitude of the deviations. In effect, we will intuitively understand the force of conviction of a methodology that allows us to increase the baccalaureat pass rate for a small lycée made up of a few classes of students rather than report a significant 0.7% difference between two schools. The effect-size once again gives meaning to the deviations. Furthermore, quantifying the importance of an effect frequently constitutes the first stage in the collation of data that will be used for a meta-analysis. However, once again, we have to accept that meta-analysis is a way of synthetically analysing literature that has unseated the classic review of questions especially given its predictive capability and its capacity for testing the validity of a hypothesis using a wide corpus of data produced by various research works (Lipsey & Wilson, 2010).

For the purpose of illustrating our argument, we shall take an actual example involving a submission technique well-known in the literature: The foot in the door. This technique consists in putting a preliminary request before proceeding with the final request: by so doing, the final request is more easily accepted than if it had been submitted first. Talking of which, the experiment carried out in 1999 (Guéguen & Fischer-Lokou, 1999). involving a sample of 3280 people stopped in the street (two groups of 1640 people) revealed a 28.3% acceptance rate under control conditions and a 42.5% acceptance rate using the foot in the door approach. This produced the following spread:

The χ^2 calculated here is: $\chi^2(1, 3280) = 72.38$, $p < .0001$.
The same conclusion is reached in the precept study (Harris, 1972): the difference is significant in the probabilistic sense of the word; however, in the Harris experiment (op.cit., 1972), the value of χ^2 was found to be lower than that produced by the experiment carried out by the authors (Guéguen & Fischer-

Table 1.
Distribution of subjects by condition (experimental vs control).

	Condition addressed	
Subject response	Foot-in-the-Door	Control
Request accepted	697	464
Request rejected	943	1176

Lokou, 1999). whereas there were greater deviations between the two groups (See Table 1): 33.3% (44.4% - 11.1) for one (Harris, 1972) compared with 14.2% (42.5 - 28.3) for the others (op.cit., 1999). Measuring the effect-size merely consists in translating the deviation and no longer the probability of obtaining said devia- tion on the basis of the size of the sample.

Determining Effect-Sizes

The ideal would consist in identifying an indicator that is independent of the sample size so that the surveys can be compared with each other and, above all, so that we can assess the effect-size observed. Indeed, sample size sensitivity is one of the first reasons that led statisticians to work on the concept of the effect-size.

The contingency coefficient Phi (written ϕ) is one of the indicators used to quantify this significance of deviations between our two proportions. It is easily applied because we only need to calculate $\sqrt{\left(\chi^2/n\right)}$ (n being the total number of individuals tested).

Accordingly, if we take the data provided by the two contingency tables above, ϕ equals $\sqrt{\left(72.38/3280\right)}$ or 0.15 in the case of the Guéguen and Fischer-Lokou experiment (1999).

Interpretation

What is the next step once this coefficient ϕ has been calculated and the effect-size assessed? Cohen (1988) put forward coefficient values used to measure the effect-size. Three categories were identified: the low effect (0.1); the medium effect (.03) and the high effect (0.5)

The effect is qualified as low/medium in the survey [2]. Thus, therefore, we can be allowed to question the scope of the analysis from the mere statistical significance viewpoint or as an assessment of the value for χ^2 (the opposite of the importance obtained).

The Various Effect-Size Indicators

In order to be able to calculate the various indicators used for measuring effect-size on the basis of the types of variable or of the analyses carried out, we have separated the headings so that you can quickly identify the appropriate procedure. We have used standard indicators for quantifying the effect-size (see Cohen, 1988 for a more in-depth review). As far as we are aware, there are many others (e.g. comparison of a mean with a standard; comparison of 2 means from separate samples; comparison of 2 means for linked samples; comparison of frequentcies in a contingency table etc.) (see Rosenthal & Rosnow, 1999 for a more in-depth review). In the case of some indicators, it is clear that the effect-size is easily calculated because the indicator had been produced from earlier analyses (e.g. the linear correlation coefficient). Furthermore, when the analysis is performed, most statistics processing software offer options allowing the user to access these indicators. We should also note that there are online statistical resources available on the Internet for calculated these various "effect-sizes" such as:
http://www.uccs.edu/~faculty/lbecker/
http://cognitiveflexibility.org/effectsize/

Conclusions

Measuring effect-size is an approach that the social psychology researcher must now include, whenever possible, into his data analysis. This becomes all the more important when extensive samples are used because they encourage the effect to be revealed even when these are limited. Foreign psychology reviews and especially the Anglo-Saxon reviews increasingly tend to demand that these indicators be presented on the same footing as the various strategic inferential tests used. Some research used to the meta-analysis computation based an adjusted variance and/or upon a pooled variance of effect size. Berk and Freedman (2003) are skeptical regarding the effecttiveness of the meta-analysis. The authors questioned the assumed independence of studies and to randomization for forced inclusion of studies. It's a very important problem for scientific research. Further, the authors are skeptical about the social dependence (and financial) between the some pool of peer-review journals and then taken to a subsequent meta-analysis by the scientific community. For authors: "In the present state of our science, invoking a formal relationship between random samples and populations is more likely to obscure than to clarify."

In this article, we have attempted to present the principle of this quantification and the way in which customary indicators are calculated. Obviously, there are presently a great many indicators that refer to specific analysis cases and that take various utilisation criteria into account. However, determining these indicators can help the social psychology researcher to break free from the classic inferential model used in statistical analysis and to opt for a method for assessing his data based on a more equitable assessment of the effects. Many researchers expose the imperialism of the inferential method and the .05 value as objectionable and recommend that these indicators be imposed (Thompson, in press). Therefore, if we use these indicators, we would be led to view some of our theoretical analyses and interpretations in a different light.

References

Berk, R. A., & Freedman, D. (2003). Statistical assumptions as empirical commitments. In T. G. Blomberg and S. Cohen (Eds.). *Law, Punishment, and Social Control: Essays in Honor of Sheldon Messinger* (2nd ed.) (pp. 235-254). New York, NY: Aldine.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Guéguen, N., & Fischer-Lokou, J. (1999). Sequential request strategy: Effect of donor generosity. *The Journal of Social Psychology, 139,* 669-671.

Harris, M. B. (1972). The effects of performing one altruistic act on the likelihood of performing anaother. *The Journal of Social Psychology, 88,* 65-73. doi:10.1080/00224545.1972.9922544

Lipsey, M. W., & Wilson, D. B. (2000). *Practical Metanalysis*. Thousand Oaks, CA: Sage Publications.

Rosenthal, R., & Rosnow, R. L. (1991). *Essentials of behavioral research: Methods and data analysis* (2nd ed.). New York: McGraw Hill.

Thompson, B. (in press) Research synthesis: Effect sizes. In J. Green, G. Camilli, & P.B. Elmore (Ed.), *Complementary methods for researchin education.* Washington, DC: American Educational Research Association.