Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6, 287-295
2024
Publisher: Learning Gate
DOI: 10.55214/25768484.v8i6.2062

The role of effect size and significance test in research design and analysis

Elsayir, H.A.[1*]
[1]Dept. of Mathematics, Al-Qunfudah University College, Umm Al- Qura University, Mecca, Saudi Arabia; hamusa@uqu.edu.sa Habibsayiroi@Yahoo.com (E.H.A.).

Abstract: Effect size is a straightforward method for measuring the difference between two groups and plays a critical role in research design and significance testing. This paper addresses the concept of effect size, explaining what it is, how it is calculated, and how it can be interpreted. It also explores related topics, such as statistical significance, meta-analysis, and power analysis. Effect sizes can be understood in the context of confidence intervals, which provide similar information to statistical significance tests but focus more on the magnitude of the effect rather than the size of the sample.

Keywords: *Cohen's d, confidence interval, Effect Size, Meta-analysis, Significance Test.*

1. Introduction

Statistical significance indicates whether a result is unlikely to have occurred due to random variations in the data. However, not all significant results correspond to effects with substantial impact; some may even represent phenomena that are not noticeable in everyday life. Statistical significance is largely influenced by factors such as sample size, data quality, and the power of the statistical methods used. In large datasets, such as those common in epidemiological studies or large-scale assessments, even very small effects can achieve statistical significance. To determine whether effects have a meaningful magnitude, effect sizes are used to quantify the strength of a phenomenon. The most widely known measure of effect size is Cohen's d,(Cohen(1988), though many others exist. Effect size, which is often conceptualized as a standardized difference, serves as a crucial tool for reporting and interpreting effectiveness and offers several advantages over relying solely on statistical significance tests. It is valuable for quantifying the effectiveness of an intervention relative to a comparison and helps researchers move beyond null hypothesis testing. This paper provides a brief overview of the basic methodologies for calculating and interpreting effect size.

2 Literature Review

The paper reviews various issues related to the topic, supported by numerical examples. The tables presented are compiled from different sources and verified using online software for calculating effect sizes (refer to the list of website references). Even though measures of effect size have been available for decades (Huberty, 2002), their use has primarily been limited to meta-analyses for combining and comparing estimates from different studies. Effect size, commonly represented by the symbol "d," is essentially the mean difference between groups in standardized score form. This concept originates from meta-analytic methods (see Biostat (2006), Poston & Hanson (2010), Aert & van Assen, 2016). Rosenthal and Rosnow (2000) provided a useful summary of effect size computations and their transformations for inferential statistics. R. Michael Fur (2008) also explored effect sizes and their connections to inferential statistics. Téllez et al. (2015) discussed the application of effect size, confidence intervals, and statistical power in psychological research. There has been considerable debate on the relationships between significance tests, p-values, and confidence intervals (see Sullivan (2012),

* Correspondence: hamusa@uqu.edu.sa

Pernet, (2016), Aert & van Assen (2016), Dunkler et al (2020), Carpenter (2020)). Recent studies have focused heavily on the use of effect sizes in psychological research (see Szucs & Ioannidis (2017), Lovakov & Agadullina(2017), Schäfer & Schwarz(2019), Lenhard & Lenhard(2022)).

2.1. The Relationship Between Effect Size, Significance and Meta-Analysis

Effect size is a name given to a set of indices that measure the magnitude of a treatment effect. Unlike significance tests, these indices are independent of sample size. Effect size measures are the common currency of meta-analysis studies that summarize the findings from a specific area of research. Effect size quantifies the size of the difference between two groups and may therefore be said to be a true measure of the significance of the difference. However, in statistics the word 'significance' is often used to mean 'statistical significance', which is the likelihood that the difference between the two groups could just be an accident of sampling. Another use of effect size is its use in performing power analysis, (see Buchner and Faul,F(2009) . Researcher designers use power analysis to minimize the likelihood of both false positives and false negatives (Type I and Type II errors, respectively) ,Richard A. Zeller and Yan Yan (2007).

Several statistics are sometimes proposed as alternative measures of effect size, other than the 'standardized mean difference'. One of these is the Proportion of variance accounted for, the R^2 which represents the proportion of the variance in each that is 'accounted for' by the other. It can be shown that the interpretation of the 'standardized mean difference' measure of effect size is very sensitive to violations of the assumption of normality. For this reason, several more robust (non-parametric) alternatives have been suggested. There are also effect size measures for multivariate outcomes. A detailed explanation can be found in Olejnik and Algina (2000).

Calculating effect size is important when testing the goodness fit ,or contingency test. For this test, the effect size symbol is w. In either case, two distributions over a certain number of categories may be compared . Once effect size is known, this information can be used to calculate the number of participants needed and the critical chi-square value (for sample size rules (see Aguinis, H. & Harden, E. E. (2009)),(and see the effect of sample size on effect size in Slavin, R., & Smith, D. (2008).Eta squared η^2 (part of the *r* family of effect sizes, and an extension of r^2 that can be used for more than two sets of observations) measures the proportion of the variation in Y that is associated with membership of the different groups defined by X, or the sum of squares of the effect divided by the total sum of squares.

3. Mathematical Formula Formulas

The developed formulas for effect size calculation vary depending on whether the researcher plans to use analysis of variance(ANOVA), t test, regression or correlation**,** (see Morris and DeShon's (2002))..Formulas used to measure effect size can be computed in either a standardized difference between two means, or in the correlation between the independent variable classification and the individual scores on the dependent variable, which is called the "effect size correlation" (Rosnow & Rosenthal(1996).

Effect size for differences in means Figure 1, is given by Cohen's" d"Cohen, J. (1988) , is defined in terms of population means (μs) and standard deviation (σ), as shown below:

$$d = \frac{|\mu_1 - \mu_2|}{\sigma} \qquad (1)$$

There are several different ways that one could estimate σ from sample data which leads to multiple variants within the Cohen's d family.(see Karl L. Wuensch(2010)).

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 287-295, 2024
DOI: 10.55214/25768484.v8i6.2062

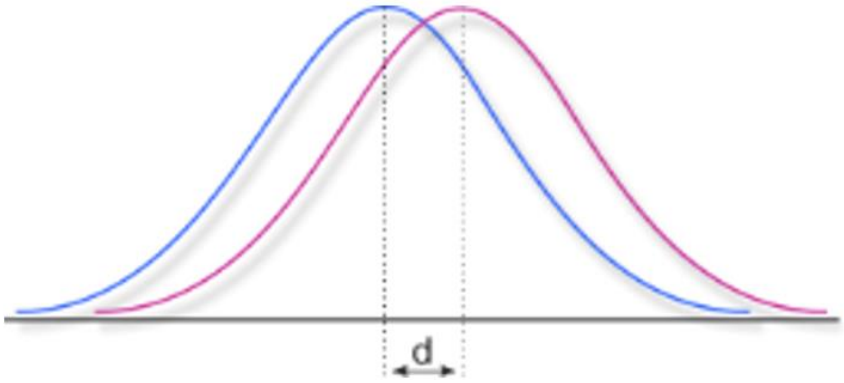

Figure 1.
The effect size " d "

When using the root mean square standard deviation, the" d " is given as:

$$d = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{s^2{}_1 + s^2{}_2}{2}}} \quad (2)$$

A version of Cohen's d uses the pooled standard deviation and is also known as Hedges' g(Zach, (2021) ,Glen (2022), is :

$$d = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{(n_1 - 1)S^2{}_1 + (n_2 - 1)S^2{}_2}{n_1 + n_2 - 2}}} \quad (3)$$

The value can be obtained from an ANOVA program by taking the square root of the mean square error which is also known as the root mean square error.

Another model of Cohen's " d" using the standard deviation for the control group is also known as Glass' Δ (see Karl L. Wuensch(2010)),where:

$$d = \frac{|\bar{x}_1 - \bar{x}_2|}{s_{cont}} \quad (4)$$

The control group's standard deviation is used because it is not affected by the treatment .It is suggested to use a pooled within group standard deviation because it has less sampling error than the control group standard deviation such that equal size constrain is adopted. When there are more than two groups , the difference between the largest and smallest means divided by the square root of the mean square error will be used ,i.e:

$$d = \frac{\bar{x}_{largest} - \bar{x}_{smallest}}{\sqrt{mse}} \quad (5)$$

As For OLS regression the measure of effects size is F which is defined by Cohen as follows:

$$f^2 = \frac{\rho^2}{1 - \rho^2} \quad (6)$$

Or, as usually computed by taking the square root of f^2

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 287-295, 2024
DOI: 10.55214/25768484.v8i6.2062

Once again there are several ways in which the effect size can be computed from sample data. It can be noted that η^2 is another name for R^2,the coefficient of determination, where: (see Karl L. Wuensch(2010)).

$$f^2 = \frac{R^2}{1-R^2} = \frac{eta^2}{1-eta^2} \quad (7)$$

An η^2 of 0.13 means that 13% of the total variance can be accounted for by group membership. Although η^2 is an efficient way to compare the sizes of effects *within* a study (given that every effect is interpreted in relation to the total variance, all η^2 from a single study sum to 100%), eta squared cannot easily be compared *between* studies, because the total variability in a study (SS_{total}) depends on the design of a study, and increases when additional variables are manipulated.

Although η^2_p is more useful when the goal is to compare effect sizes across studies, it is not perfect, because η^2_p differs when the same two means are compared in a within-subjects design or a between-subjects design. In a within-subjects ANOVA, the error sum of squares can be calculated around the mean of each measurement, but also around the mean of each individual when the measurements are averaged across individuals. Consequently, whenever the two groups of observations are positively correlated, η^2_p will be larger in a within-subjects design than in a between-subjects design. This is also the reason a within-subjects ANOVA typically has a higher statistical power than a between-subjects ANOVA. The effect size used in analysis of variance is defined by the ratio of population standard deviations:

$$f = \frac{\sigma_{means}}{\sigma} \quad (8)$$

Based on definitional formula in terms of population values., effect size w can be viewed as the square root of the standardized chi-square statistic.

$$w = \sqrt{\sum \frac{(\pi_0 - \pi_1)^2}{\pi_0}} \quad (9)$$

And *w* is computed using sample data by the formula:

$$w = \sqrt{\sum \frac{(P_0 - P_1)^2}{P_0}} \quad (10)$$

According to Poston &Hanson(2010),when a study reports a hit rate (percentage of success after taking the treatment or no treatment),the following formula can be used:

d= arcsine(p1)+ arcsine(p2)

Where p1 and p2 are the hit rates of the two groups.

If the effect size estimate from the sample is *d*, then it is normally distributed, with standard deviation:

$$\sigma(d) = \sqrt{\frac{N_{exp} + N_{cont}}{(N_{\text{exp}})(N_{cont})} + \frac{d^2}{2(N_{\text{exp}} + N_{cont})}} \quad (11)$$

(Where N_{exp} and N_{cont} are the numbers in the experimental and control groups, respectively.)

The control group will provide the best estimate of standard deviation, since it consists of a representative group of the population who have not been affected by the experimental intervention..However, in studies where there is not a true 'control' group then it may be an arbitrary decision which group's standard deviation to use to for the estimate of effect size.

Therefore, it is often better to use a 'pooled' estimate of standard deviation, which is given by:

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 287-295, 2024
DOI: 10.55214/25768484.v8i6.2062

$$SD(pooled) = \sqrt{\frac{(N_{exp}-1)\ SD^2{}_{exp} + (N_{cont}-1)\ SD^2{}_{cont}}{(N_{\exp})+(N_{cont})-2}} \qquad (12)$$

(Where N_{exp} and N_{cont} are the numbers in the experimental and control groups, respectively, and $SD^2{}_{exp}$ and $SD^2{}_{cont}$ are their variances.). The experimental and control group standard deviations differ only because of sampling variation. Where this assumption cannot be made), then a pooled estimate should not be used.

Based on the above formula's values, the larger the effect size, the greater is the impact of an intervention. Cohen suggested that a correlation of 0.5 is large, 0.3 is moderate, and 0.1 is small Cohen defined .40 as the medium effect size because it was close to the average observed effect size(Aguinis, & Harden(2009). The usual interpretation of this statement is that anything greater than 0.5 is large, 0.5-0.3 is moderate, 0.3-0.1 is small, and anything smaller than 0.1 is trivial.

4. Effect Size Computation

The interpretations of effect-sizes given in Table 1 in which a suggested values for low, medium and high effects is given, depend on the assumption that both control and experimental groups have a 'normal' distribution, otherwise, it may be difficult to make a fair comparison between an effect-size based on normal distributions and one based on non-normal distributions. In practice, the values for large effects may be exceeded with values Cohen's d greater than 1.0 not uncommon. However, using very large effect sizes in prospective power analysis is probably not a good idea as it could lead to underpowered studies.

Table 1.
Calculating effect size for different tests.

		Small	Medium	Large
t-test for means	d	0.20	0.50	0.80
t-test for correlation	r	0.10	0.30	0.50
F-test for regression	f^2	0.02	0.15	0.35
F-test for ANOVA	f	0.10	0.25	0.40
chi-square	w	0.10	0.30	0.50

Considering Table 1 and Table 2, it can be noted that, *d* can be converted to *r* and vice versa. For example, the *d* value of 0.8 corresponds to an *r* value of 0.371. The square of the r-value is the percentage of variance in the dependent variable that is accounted for by the effect in the explanatory variable groups.

Table 2.
Comparable effect sizes when there are two groups.

d	r	r^2	f	f^2
2	0.707	0.49985	0.999698	0.999396
1.8	0.669	0.44756	0.900086	0.810155
1.6	0.625	0.39063	0.800641	0.641026
1.4	0.573	0.32833	0.699160	0.488824
1.2	0.514	0.26420	0.599214	0.359058
1.0	0.447	0.19981	0.499702	0.249702
0 .8	0.371	0.13764	039951 0	0.159610
0 .6	0.287	0.08237	0.299604	0.089763

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 287-295, 2024
DOI: 10.55214/25768484.v8i6.2062

0 .4	0.196	0.03842	0.199877	0.039951
0 .2	0.10	0.01000	0.100504	0.010100
0.1	0.05	0.0025	0.050063	0.002506
0	0	0	0	0

Note: *Notice the relationship between d, r, and r^2

For a *d* value of 0.8, the amount of variance in the dependent variable by membership in the treatment and control groups is 13.8%. T-tests are used to evaluate the null hypothesis that a product moment correlation in the population is zero (r = 0). For this test, the effect size symbol is r. If the desired effect size is known, statistical power and needed sample size can be calculated. For instance, if the target is to find how many elements are need in a study for a medium effect size (r = 0.30) with an alpha of .05. and power of 0.95, this information can be used to find the answer.

For ANOVA, the effect size index f is used, and the effect size index from the group means can then be computed.

Table 3.
Effect size, sample size &power (Alpha=0.05; Power=0.95).*

Effect size	Delta	Critical t	Total sample size	Actual power
0.001	3.605	1.960	51978840	0.950
0.100	3.606	1.960	5200	0.950
0.200	3.608	1.962	1302	0.950
0.300	3.613	1.964	580	0.950
0.400	3.622	1.967	328	0.951
0.500	3.623	1.971	210	0.950
0.600	3.650	1.976	148	0.952
0.700	3.671	1.982	110	0.953
0.800	3.666	1.989	84	0.952
0.900	3.711	1.997	68	0.955
10.00	10.000	4.303	4	0.993

Note: *Power depends on the effect size, the sample size and the significance level.

Power is the chance that if "d" exists in the real world, one gets a statistically significant difference in the data .if the power level is taken to be 80%,there is an 80% chance to discover an existing difference in the sample. Alpha is the chance that one would conclude that an effect difference "d", has been discovered, while in fact this difference or effect does not exist. If alpha is set at 5%, this means that in 5%, or one in twenty, the data indicate that "something" exists, while in fact it does not.In table (3),consider that: power = 1-β = p (HA is accepted/HA is true).Set α ,the probability of false rejecting Ho,equal to some small value .Then ,considering the alternative hypothesis HA,choose a region of rejection such that the probability of observing a sample value in that region is less than or equal to α when Ho is true.

If the value of sample statistic falls within the rejection region, the decision is made to reject the null hypothesis. Typically, is set at 0.05, and critical t values are specified. The calculation works as follows: Entering α=0.05, power=0.95, effect size specified as in column (1), we find the needed elements (sample size (column 4)) and so on.

The effect size is seen in Table 3 Column (1).

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 287-295, 2024
DOI: 10.55214/25768484.v8i6.2062

Table 4.
Calculate d and r using means and st.ds for two groups.

Group I				Group II			
M1	**SD1**	**Cohen's d**	**Effect size r**	**M2**	**SD2**	**Cohen's d**	**Effect size r**
1	1	0	0	1	1	0	0
2	5	0.505-	0.245-	6	10	0.505-	0.245-
5	10	0.632-	0.302-	10	5	0.632-	0.302-
5	10	0.5	0.243	0	10	0.5	0.243
15	50	0.1-	0.049-	20	50	0.1-	0.049-
20	50	0	0	20	10	0	0.5
50	100	0.380	0.186	20	50	0.380	0.186
50	100	-0.280	0.139-	50	100	-0.280	0.139-

Note: d and r are positive if the mean difference is in the predicted direction.

$$Cohen's\ d\ = \frac{m_1\_m_2}{\sigma_{pooled}} \quad , \sigma_{pooled} = (\ \sigma_1^2 + \sigma_2^2\ )/2$$

The effect size conventions are small =0.20, medium=0.50, large=0.80. Calculate d and r using t values and df (separate groups t test) calculate the value of Cohen's d and the effect size correlation r ,using the t test value for a between subjects t test and the degrees of freedom. Results are shown in table (4),while in Table 5 ,d and r are calculated using t values and df.

Table 5.
Calculate d and r using t values and df.

T value	D f	Cohen's d	Effect size r
1	1	2	0.7071
1.5	2	2.1213	0.7276
2.0	5	1.7888	0.6666
2.0	10	1.2649	0.5345
2.5	30	0.9128	0.4152
3.0	30	1.0954	0.4803
3.0	50	0.8485	03905

Note: d and r are positive if the mean difference is in the predicted direction.

$$\text{Cohen's d} = \frac{2t}{dt^2} \quad , \quad r = (\frac{t^2}{t^2 + df})^2$$

5. Discussion and Conclusion

The effect size refers to the magnitude of the effect under the alternative hypothesis. It should represent the smallest difference that would be of significance. It varies from study to study. It is also variable from one statistical procedure to the other. It could be the difference in cure rates, or a standardized mean difference or a correlation coefficient. If the effect size is increased, the type II error decreases. Power is a function of an effect size and the sample size. For a given power, 'small effects' require larger sample size than 'large effects'. Power depends on (a) the effect size, (b) the sample size, and (c) the significance level. But if the researcher knew the size of the effect, there would be no reason to conduct the research.To estimate a sample size prior to doing the research, requires the postulation of an effect size, which might be related to a correlation, an f-value, or a non-parametric test. In the procedure implemented here ,'d' is the difference between two averages, or proportions. Effect size 'd' is

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 287-295, 2024
DOI: 10.55214/25768484.v8i6.2062

mostly subjective, it is the difference you want to discover as a researcher or practitioner, and it is a difference that you find relevant. However, if cost aspects are included ,'d' can be calculated objectively. The size of the difference in the response to be detected, which relates to underlying population, not to data from sample, is of importance since it measures the distance between the null hypothesis (HO) and specific value of the alternative hypothesis (HA).A desirable effect size is the degree of deviation from the null hypotheses that is considered large enough to attract the attention. The concept of small, medium, and large effect sizes can be a reasonable starting point if you do not have more precise information. (Note that an effect size should be stated in terms of a number in the actual units of the response, not a percent change such as 5% or 10 %.).

Statistical significance is the probability that the observed difference between two groups is due to chance. If the *P* value is larger than the alpha level chosen (eg, .05), any observed difference is assumed to be explained by sampling variability. With a sufficiently large sample, a statistical test will almost always demonstrate a significant difference, unless there is no effect whatsoever, that is, when the effect size is exactly zero; yet very small differences, even if significant, are often meaningless.

Plotting power as a function of effect size and sample size helps to avoid wasting resources and identify the optimal sample size. Generally, as either the effect size or the sample size increases, the power level also increases. Regarding the use of effect sizes, the following points can be summarized:

5.1. Standardized Measure

Effect size is a standardized measure that quantifies the magnitude of an intervention's impact, making it particularly useful for comparing the relative sizes of effects across different studies.

5.2. Assumptions Matter

When using effect size, it's crucial to consider assumptions such as normality and equality of variances between the 'control' and 'experimental' groups. Effect sizes can also be interpreted in terms of the degree of overlap between two distributions, reflected by percentiles or ranks.

5.3. Complement to Statistical Significance

Using an effect size with a confidence interval provides similar information to a test of statistical significance but focuses on the magnitude of the effect rather than just the sample size.

Given these points, it is important to encourage the use of effect sizes, their confidence intervals, and appropriate statistical power in research design and analysis. This approach offers a more comprehensive perspective on the preparation, analysis, and interpretation of statistical data.

Copyright:

References

[1] Aert, R. C., Wicherts, J. M., & van Assen, M. A. (2016). Conducting meta-analyses based on p values: Reservations and recommendations for applying p-uniform and p-curve. Perspectives on Psychological Science, 11, 713–729. doi:10.1177/1745691616650874

[2] Aguinis ,H.&Haren,E.E.(2009).Sample size rules of thumb: Evaluating three common practices. In Charles E. Lance and Robert J. Vandenberg (Eds.), Statistical and methodological myths and urban legends:Doctrine,verify and fable in the organizational and social sciences (pp.267-286).New York:Routledge .

[3] Carpenter, A. (2020). Effect size: Moving beyond the p-value. Towards Data Science. Retrieved from https://towardsdatascience.com/effect-size-d132b0cc8669. Accessed December 19, 2020.

[4] Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 287-295, 2024
DOI: 10.55214/25768484.v8i6.2062

[5] Dunkler, D., Haller, M., Oberbauer, R., &Heinze, G. (2020). To test or to estimate? P-values versus effect sizes. Transplant International: Official Journal of the European Society for Organ Transplantation, 33(1), 50–55. https://doi.org/10.1111/tri.13535
[6] Glen, S. (2022). Hedges' g: Definition, Formula. StatisticsHowTo.com. Retrieved from https://www.statisticshowto.com/hedges-g/. Accessed October 14, 2022.
[7] Huberty, C.J.. (2002) 'A history of effect size indices'. Educational and Psychological Measurement, 62, 2, 227-240.
[8] Lenhard, W. & Lenhard, A. (2022). Computation of effect sizes. Retrieved from: https://www.psychometrica.de/effect_size.html. Psychometrica. DOI: 10.13140/RG.2.2.17823.92329.
[9] Lovakov, A., & Agadullina, E. R. (2017). Empirically derived guidelines for interpreting effect size in social psychology. Retrieved from https://psyarxiv.com/2epc4 van.
[10] Mike Fur(2008) Effect sizes and significance Tests .Psychology Dept. Wake Forest University(.March,2008).
[11] Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. Psychological Methods, 7, 105-125.
[12] Olejnik, S. and Algina, J. (2000) 'Measures of Effect Size for Comparative Studies: Applications, Interpretations and Limitations.' Contemporary Educational Psychology, 25, 241-286.
[13] Pernet, C. (2016). Null hypothesis significance testing: A short tutorial. F1000Research, 4, Article 621. https://doi.org/10.12688/f1000research.6963.5 Publication Manual of the American Psychological Association. (7th ed.). (2020). American Psychological Association. https://doi.org/10.1037/0000165-000
[14] Poston,J.M,&Hanso,W.E.(2010).Meta-analysis of psychological assessment as a therapeutic intervention. Psychological Assessment ,22,203-12.
[15] Richard A. Zeller and Yan Yan (2007). Establishing the Optimum Sample Size .Epidemiology and Medical Statistics ,Handbook of Statistics Volume 27, Pages 656-678.
[16] Rosenthal, R. (1994) 'Parametric Measures of Effect Size' in H. Cooper and L.V. Hedges (Eds.), The Handbook of Research Synthesis. New York: Russell Sage Foundation.
[17] Rosenthal, R.Rosnow,R.L.&Rubin (2000)D.B. Contrasts and Effect sizes in behavioral research :A correlational approach. Cambridge .UK. Cambridge University Press.
[18] Schäfer, T., & Schwarz, M. A. (2019). The meaningfulness of effect sizes in psychological research: Differences between sub-disciplines and the impact of potential biases. Frontiers in Psychology, 10, 813. doi:10.3389/fpsyg.2019.00813.
[19] Slavin,R.,&Smith,D.(2008).Effects of sample size on effect size in systematic reviews in education .Paper presented at the annual meetings of the Society for Research on Effective Education, Crystal City ,VI.
[20] Sullivan GM, Feinn R. Using Effect Size-or Why the P Value Is Not Enough. J Grad Med Educ. 2012 Sep;4(3):279-82. doi: 10.4300/JGME-D-12-00156.1. PMID: 23997866; PMCID: PMC3444174.
[21] Szucs, D., & Ioannidis, J. P. (2017). Empirical assessment of published effect sizes and power in the recent cognitive neuroscience and psychology literature. PLoS Biology, 15, e2000797. doi:10.1371/journal.pbio.2000797.
[22] Téllez A., García C.H., Corral-Verdugo V. (2015). Effect size, confidence intervals, and statistical power in psychological research. Psychology in Russia: State of the Art, 8(3), 27-46. https://doi.org/10.11621/pir.2015.0303
[23] Zach, Statology. (2021). What is Hedges' g? (Definition & Example). Statology. Retrieved from https://www.statology.org/hedges-g. Accessed October 15, 2022.

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 287-295, 2024
DOI: 10.55214/25768484.v8i6.2062