

BINOMIAL EFFECT SIZE DISPLAY

Authored by
mohammad looti

November 6, 2025

RECOMMENDED CITATION

mohammad looti (2025). *BINOMIAL EFFECT SIZE DISPLAY*. PSYCHOLOGICAL SCALES.
Retrieved from <https://scales.arabpsychology.com/?p=66788>

BINOMIAL EFFECT SIZE DISPLAY

Primary Disciplinary Field(s): Statistics, Quantitative Psychology, Research Methods, Meta-Analysis

1. Core Definition

The Binomial Effect Size Display (BED) is a powerful statistical technique and pedagogical tool designed primarily to translate correlation coefficients into readily interpretable practical outcomes, typically expressed as percentages or frequencies in a 2x2 contingency table. Developed primarily by Robert Rosenthal and Donald B. Rubin, the method addresses the common difficulty researchers and lay audiences face in grasping the substantive importance of a standardized effect size, such as Pearson's r . While correlation coefficients mathematically quantify the strength and direction of linear association between two variables, their meaning in real-world terms--especially regarding predictive validity or treatment efficacy--is often obscured by statistical abstraction.

The fundamental premise of the BED is to array experimental or correlational results into a matrix that clearly displays the effect of a treatment or predictor variable on a dichotomous outcome. This display typically assumes a hypothetical scenario where both the independent variable (e.g., treatment versus control) and the dependent variable (e.g., success versus failure, survival versus mortality) have been artificially dichotomized at their respective medians. By forcing this binary representation, the BED highlights the difference in success rates between the two groups, effectively demonstrating the variance explained in terms of tangible rates of improvement or survival, thereby making the effect size immediately accessible and understandable.

For example, if a study examines the effect of a new intervention, the BED takes the calculated correlation (r) between intervention usage and recovery and translates it into the percentage of people who recovered in the intervention group versus the control group. This shift from a technical statistic (r) to intuitive success rates (percentages) is crucial for communicating the **practical significance** of research findings, especially in applied fields like medicine, social work, and organizational psychology, where stakeholders require clear metrics of efficacy.

2. Statistical Foundation and Calculation

The mathematical transformation underlying the BED allows for the conversion of the continuous effect size r into the frequencies necessary for the 2x2 table, provided that the marginal distributions are assumed to be 50% for both the predictor and the outcome variables. This crucial standardization ensures that the base rate for success in the population is set at 50%, a simplifying assumption that maximizes the displayed difference relative to the effect size and ensures a clear representation of the effect size itself, independent of potential real-world base rate complications.

The calculation is straightforward. If r represents the correlation coefficient, the proportion of successes in the treatment group (PT) is calculated as $PT = 0.50 + r/2$, and the proportion of successes in the control group (PC) is calculated as $PC = 0.50 - r/2$. The resulting difference between these two success rates (PT - PC) is mathematically equivalent to the correlation coefficient r . Consequently, a correlation of $r = 0.40$ results in a success rate of 70% ($0.50 + 0.20$) for the treatment group and 30% ($0.50 - 0.20$) for the control group, representing a 40 percentage point difference attributable to the effect.

The BED is primarily a descriptive measure and is not intended to replace standard inferential statistics or measures of variance explained, such as R-squared (which is r^2). Instead, it serves as a powerful complementary tool. While R-squared quantifies the proportion of variance in the outcome accounted for by the predictor, the BED quantifies the magnitude of the effect in terms of differential success rates. This dual approach is essential for bridging the gap between statistical significance and meaningful effect size estimation, a challenge frequently encountered in research reporting and interpretation.

3. Key Characteristics and Assumptions

The Binomial Effect Size Display possesses several defining characteristics that dictate its utility and application in research communication. Its primary strength lies in its ability to standardize the presentation of effects across diverse studies, provided the effect size used is the correlation coefficient or a mathematically equivalent metric that can be converted to r .

Dichotomization of Variables: The BED necessitates the assumption that both the independent and dependent variables have been dichotomized, usually at the median. This simplification converts continuous data into binary outcomes (e.g., high performance/low performance, survival/mortality), which streamlines the creation of a clear 2x2 contingency table structure.

Standardized Base Rate (50%): The methodology requires the strict assumption of a 50% base rate for both the predictor and the outcome variables. This assumption is crucial because it ensures that the resulting displayed difference is purely reflective of the correlation coefficient r itself, isolating the effect size magnitude from the confounding influence of skewed marginal distributions that might exist in the real data.

Focus on Practical Magnitude: Unlike measures of statistical significance (p-values) which test against the null hypothesis, the BED is purely a measure of **effect size magnitude**. It provides an answer to the critical research question: "How much difference does this intervention make?" in terms of easily comprehensible success percentages, thereby informing practical decision-making.

Independence from Scale: Because the BED operates solely based on the standardized correlation coefficient r , its result is independent of the original scales of measurement used in the study. This makes it an ideal tool for comparing effect sizes across vastly different studies, such as comparing the effectiveness of a teaching method (measured via test scores) with the

effectiveness of a medical treatment (measured via survival time).

4. Historical Development and Pedagogical Value

The Binomial Effect Size Display was introduced by Robert Rosenthal and Donald B. Rubin in 1982, in response to the pervasive tendency among researchers and journals to overemphasize statistical significance (p -values) while minimizing the importance of reporting and interpreting effect size. They observed that even when effect sizes were reported, metrics like R -squared were often misinterpreted as trivial simply because they seemed small (e.g., an R -squared of 0.04 or 4% of variance explained).

The core innovation of the BED was the recognition that the correlation coefficient, when translated into binomial success rates under the 50% base rate assumption, provides an immensely intuitive measure of practical importance. For instance, an r of 0.20 yields an R -squared of only 4%. When viewed through the BED, this same effect translates into a success rate of 60% for the treatment group and 40% for the control group--a 20 percentage point difference that is rarely dismissed as trivial. Rosenthal and Rubin argued that this display was essential for improving statistical literacy among consumers of research.

The BED is highly valued as a teaching tool in quantitative methodology courses. It helps students and non-statisticians intuitively understand why a correlation coefficient that appears small in absolute magnitude can still represent a highly meaningful difference in practical terms. It serves to decouple the conceptual understanding of effect size magnitude from the narrow focus on variance accountability, thereby addressing a fundamental confusion in effect size interpretation that persisted throughout the latter half of the 20th century.

5. Application in Research Synthesis and Communication

The practical utility of the BED extends significantly into the realm of meta-analysis and the comparison of treatment effects across diverse bodies of literature. Because the BED provides a standardized method for translating various effect sizes into comparable differences in success rates, it allows researchers to synthesize findings from multiple studies that might have originally used differing outcome measures or scales, provided those measures can be converted to r .

In clinical trials and intervention research, the BED can powerfully convey the comparative efficacy of treatments to non-technical audiences, such as institutional review boards, government funders, and patient advocacy groups. If a meta-analysis finds an average correlation of $r = 0.15$ for Intervention A and $r = 0.50$ for Intervention B, the BED immediately translates these abstract numerical differences into clear policy implications. Intervention A results in a success rate spread of 57.5% versus 42.5% (a 15-point difference), whereas Intervention B results in a spread of 75% versus 25% (a 50-point difference). This immediate visual and numerical contrast aids

stakeholders in making informed decisions regarding resource allocation and treatment adoption.

Furthermore, the BED is often employed when researchers wish to dispel the myth that statistically significant findings necessarily imply large or practical effects, or conversely, that seemingly small correlations are trivial. It serves as a necessary check against the sole reliance on p-values, emphasizing that even a weak correlation applied across a massive population might still yield substantial practical benefits, a concept easily obscured when effect size is only assessed via variance explained.

6. Limitations and Criticisms

Despite its pedagogical and communicative advantages, the Binomial Effect Size Display is subject to several significant statistical and practical limitations, primarily stemming from its core simplifying assumption regarding base rates and data structure.

The most frequent criticism levied against the BED is its forced assumption of a 50% marginal distribution for both the predictor and the outcome variables. While this standardization is mathematically convenient for deriving r , it rarely reflects the true prevalence or base rate found in empirical research. For instance, if researchers are evaluating a screening tool for a rare disease where the actual prevalence (base rate) is 2%, presenting the results as if the success rate were 50% fundamentally misrepresents the population data and can lead to inflated expectations about the actual number of true positives achieved. Critics argue that while the BED accurately displays the magnitude of r , it fails to display the actual data of the study in question unless the base rates happen to be near 50%. This limitation means the BED must be communicated strictly as an index of **effect size magnitude**, rather than a prediction of real-world outcomes.

A secondary criticism relates to the artificial dichotomization of continuous variables, a process that inherently results in the loss of statistical power and information. By collapsing a continuous distribution (e.g., scores on a measurement scale, salary levels) into two categories (e.g., high/low), researchers discard the nuanced variation within the original data. Although this is done for visualization purposes, some statisticians argue that relying on the BED reinforces potentially misleading statistical practices. Researchers must take care to ensure that the audience understands that the BED is a highly stylized visualization of the correlation coefficient under standardized conditions, and not a replacement for full data analysis.

7. Comparison with Other Effect Size Measures

The BED occupies a unique position among standardized effect size metrics, offering interpretive advantages that metrics based on means or variance sometimes lack. It is often compared to the two most common measures used in quantitative research.

The primary comparative metric is **Cohen's d** , which quantifies the difference between two means in standard deviation units. While d is excellent for comparing group means in continuous outcomes, the BED is better suited for conveying probability differences in binary outcomes. The conversion between r (used in BED) and d is widely used, allowing researchers to choose the display format most appropriate for their audience. However, d maintains sensitivity to the standard deviation of the populations being compared, whereas the BED standardizes this context away by assuming a homogenous distribution.

Another important comparison is the **Number Needed to Treat (NNT)**, a measure widely used in epidemiology and clinical medicine. The NNT states the number of patients who must be treated to achieve one additional favorable outcome compared to a control group. While NNT provides a direct, practical measure for clinicians, it is highly sensitive to the actual base rate (prevalence) of the outcome in the population. Conversely, the BED, which is sensitive only to r , provides a measure of effect size that is conceptually independent of the base rate (under the fixed 50% assumption). Researchers frequently find that using both the BED (to show the universal magnitude of the effect) and the NNT (to show the contextual impact given the actual base rate) provides the most comprehensive picture for both researchers and practitioners.

8. Further Reading

[Binomial Effect Size Display \(Wikipedia\)](#)

[Rosenthal, R., & Rubin, D. B. \(1982\). A simple, general purpose display of magnitude of experimental effect. Journal of Educational Psychology, 74\(2\), 166-169.](#)

[Rosenthal, R., & Rubin, D. B. \(2019\). The Binomial Effect Size Display: A simple, general-purpose method for displaying the magnitude of experimental effects. Psychological Methods, 24\(1\), 1-7.](#)