

# Interpreting Cohen's $d$ effect size an interactive visualization

Created by Kristoffer Magnusson (<http://rpsychologist.com>)

Follow @krstoffr

[in Kristoffer's LinkedIn profile \(https://se.linkedin.com/pub/kristoffer-magnusson/b5/133/1b6\)](https://se.linkedin.com/pub/kristoffer-magnusson/b5/133/1b6)

Tweet

Gilla 1,5t

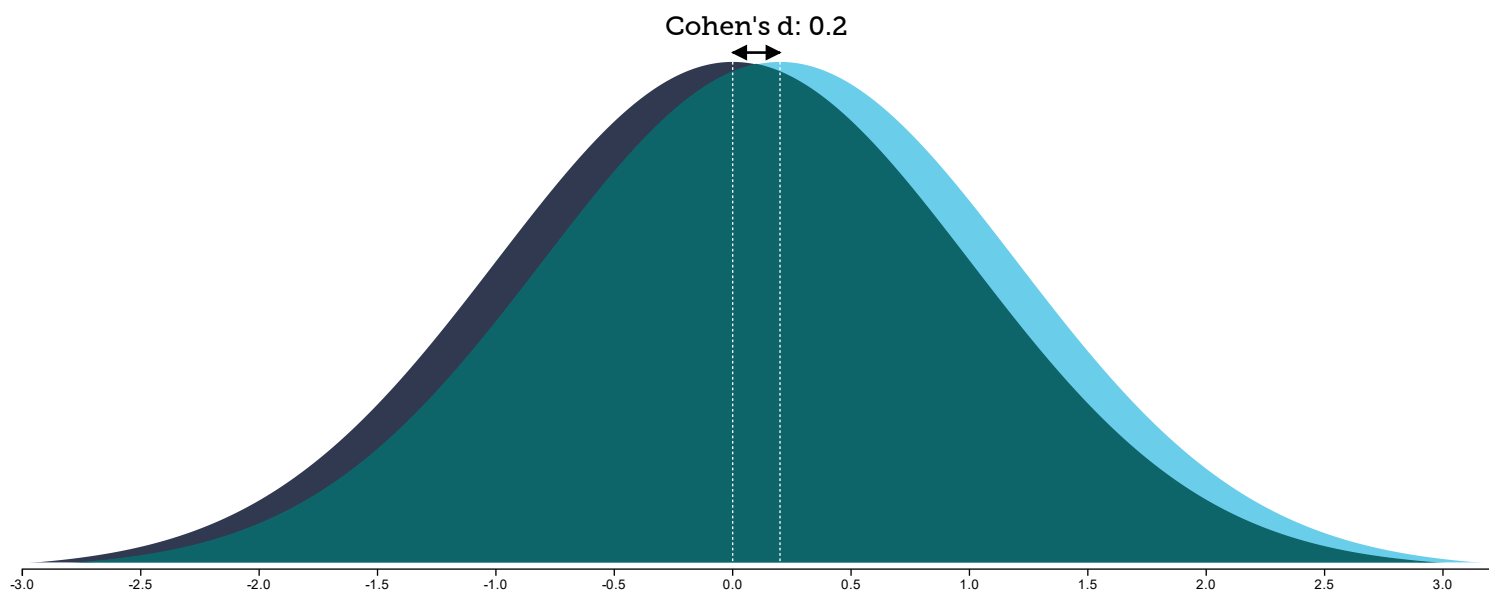
Dela

The Cohen's  $d$  effect size is immensely popular in psychology. However, its interpretation is not straightforward for clinicians and laypersons, as it requires prior knowledge about what a standard deviation is. Even practicing scientists often turn to general guidelines, such as small (0.2), medium (0.5) and large (0.8) when interpreting the effect of an intervention.

These cut-offs were introduced by Cohen himself, but with a strong caution that "this is an operation fraught with many dangers" (Cohen, 1977 ([http://www.amazon.com/gp/product/0805802835/ref=as\\_li\\_tl?ie=UTF8&camp=1789&creative=9325&creativeASIN=0805802835&linkCode=as2&tag=rpsyc-20&linkId=WA44H7YKKL5BDIFK](http://www.amazon.com/gp/product/0805802835/ref=as_li_tl?ie=UTF8&camp=1789&creative=9325&creativeASIN=0805802835&linkCode=as2&tag=rpsyc-20&linkId=WA44H7YKKL5BDIFK))). Just like  $p$ -values, these arbitrary cut-offs seem to be used mindlessly today. I believe that such "canned effect sizes" (Baguley, 2009, p. 613) should be avoided. Findings from studies need to be interpreted by their practical and clinical significance. Factors like the quality of the study, the uncertainty of the estimate and results from previous work in the field need to be appraised before declaring an effect "large".

In order to aid the interpretation of Cohen's  $d$  this visualization offers these different representations of Cohen's  $d$ : Visually, Cohen's  $U_3$ , Probability of superiority, Percentage of overlap and Number needed to treat.

Slide me ⚙



## Interpretation

57.93 %

Cohen's  $U_3$

92.03 %

% Overlap

55.62 %

Probability of Superiority

16.5

Number Needed to Treat<sup>1</sup>

## A Common Language Explanation

With a Cohen's  $d$  of 0.2, 58 % of the treatment group will be above the mean of the control group (Cohen's  $U_3$ ), 92 % of the two groups will overlap, and there is a 56 % chance that a person picked at random from the treatment group will have a higher score than a person picked at random from the control group (probability of superiority). Moreover, in order to have one more favorable outcome in the treatment group compared to the control group we need to treat 16.5 people. This means that if 100 people go through the treatment, 6.1 more people will have a favorable outcome compared to if they had received the control treatment<sup>1</sup>.

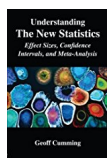
<sup>1</sup> It is assumed that 20 % of the control group have "favorable outcomes", i.e. improve below some predefined cut-off. Change this by pressing the ⚙ symbol above the slider. Go to the formula section for more information.

## Recommended books

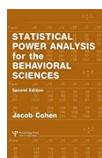
Here are some books that I have found useful both in understanding effect sizes and calculating them.

- **Statistical Power Analysis for the Behavioral Sciences (2nd Edition)** ([http://www.amazon.com/gp/product/0805802835/ref=as\\_li\\_tl?ie=UTF8&camp=1789&creative=9325&creativeASIN=0805802835&linkCode=as2&tag=rpsyc-20&linkId=7L2DJGEOMK65FIND](http://www.amazon.com/gp/product/0805802835/ref=as_li_tl?ie=UTF8&camp=1789&creative=9325&creativeASIN=0805802835&linkCode=as2&tag=rpsyc-20&linkId=7L2DJGEOMK65FIND))
- **The Essential Guide to Effect Sizes: Statistical Power, Meta-Analysis, and the Interpretation of Research Results** ([http://www.amazon.com/gp/product/0521142466/ref=as\\_li\\_tl?ie=UTF8&camp=1789&creative=9325&creativeASIN=0521142466&linkCode=as2&tag=rpsyc-20&linkId=MTEJ3MVIASDIWILB](http://www.amazon.com/gp/product/0521142466/ref=as_li_tl?ie=UTF8&camp=1789&creative=9325&creativeASIN=0521142466&linkCode=as2&tag=rpsyc-20&linkId=MTEJ3MVIASDIWILB))
- **Understanding The New Statistics: Effect Sizes, Confidence Intervals, and Meta-Analysis (Multivariate Applications Series)** ([http://www.amazon.com/gp/product/041587968X/ref=as\\_li\\_tl?ie=UTF8&camp=1789&creative=9325&creativeASIN=041587968X&linkCode=as2&tag=rpsyc-20&linkId=M4M464I22Q2U6JPC](http://www.amazon.com/gp/product/041587968X/ref=as_li_tl?ie=UTF8&camp=1789&creative=9325&creativeASIN=041587968X&linkCode=as2&tag=rpsyc-20&linkId=M4M464I22Q2U6JPC))
- **Effect Sizes for Research: Univariate and Multivariate Applications, Second Edition** ([http://www.amazon.com/gp/product/0415877695/ref=as\\_li\\_tl?ie=UTF8&camp=1789&creative=9325&creativeASIN=0415877695&linkCode=as2&tag=rpsyc-20&linkId=4OIBKHQMZYNA7IFK](http://www.amazon.com/gp/product/0415877695/ref=as_li_tl?ie=UTF8&camp=1789&creative=9325&creativeASIN=0415877695&linkCode=as2&tag=rpsyc-20&linkId=4OIBKHQMZYNA7IFK))
- **Introduction to Meta-Analysis** ([http://www.amazon.com/gp/product/0470057246/ref=as\\_li\\_tl?ie=UTF8&camp=1789&creative=9325&creativeASIN=0470057246&linkCode=as2&tag=rpsyc-20&linkId=WXCESYAT3KJRT5IQ](http://www.amazon.com/gp/product/0470057246/ref=as_li_tl?ie=UTF8&camp=1789&creative=9325&creativeASIN=0470057246&linkCode=as2&tag=rpsyc-20&linkId=WXCESYAT3KJRT5IQ))
- **Practical Meta-Analysis (Applied Social Research Methods)** ([http://www.amazon.com/gp/product/0761921680/ref=as\\_li\\_tl?ie=UTF8&camp=1789&creative=9325&creativeASIN=0761921680&linkCode=as2&tag=rpsyc-20&linkId=FI8DZ3BHN7KZ3HJP](http://www.amazon.com/gp/product/0761921680/ref=as_li_tl?ie=UTF8&camp=1789&creative=9325&creativeASIN=0761921680&linkCode=as2&tag=rpsyc-20&linkId=FI8DZ3BHN7KZ3HJP))

### Shop Related Products



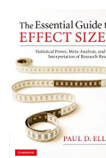
Understanding The New Statistics: Effect Sizes, Confidence Intervals, and Met...  
(<http://aax-us-east.amazon-adsystem.com/x/c/QiK39xloOlesxO3ejp...>)  
New-Statistics-Meta-Analysis-Multivariate/dp/041587968X/ref=sm\_n\_...  
sigts=1498650344498&sig=f6a5c01c52...  
20&linkCode=w41&ref=...  
refURL=http%3A%2F%2Frpsychologist...  
rR56xOspbg)



Statistical Power Analysis for the Behavioral Sciences (2nd Edition)  
(<http://aax-us-east.amazon-adsystem.com/x/c/QiK39xloOlesxO3ejp...>)  
Power-Analysis-Behavioral-Sciences/dp/0805802835/ref=sm\_n\_au...  
sigts=1498650344498&sig=5de4febb4...  
20&linkCode=w41&ref=...  
refURL=http%3A%2F%2Frpsychologist...  
rR56xOspbg)



Introduction to Meta-Analysis (<http://aax-us-east.amazon-adsystem.com/x/c/QiK39xloOlesxO3ejp...>)  
Meta-Analysis-Michael-Borenstein/dp/0470057246/ref=sm\_n\_a...  
sigts=1498650344498&sig=a1af9d350...  
20&linkCode=w41&ref=...  
refURL=http%3A%2F%2Frpsychologist...  
rR56xOspbg)



The Essential Guide to Effect Sizes: Statistical Power, Meta-Analysis, and the...  
(<http://aax-us-east.amazon-adsystem.com/x/c/QiK39xloOlesxO3ejp...>)  
Guide-Effect-Sizes-Interpretation/dp/0521142466/ref=sm\_n\_...  
sigts=1498650344498&sig=a1af9d350...  
20&linkCode=w41&ref=...  
refURL=http%3A%2F%2Frpsychologist...  
rR56xOspbg)

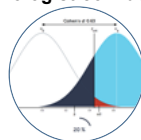
## More visualizations



### Interpreting Correlations

Interactive visualization of correlations

(<http://rpsychologist.com/d3/correlation/>)



### NHST

Interactive visualization of statistical power and significance testing

(<http://rpsychologist.com/d3/NHST/>)



### Confidence Intervals

Interactive visualization of confidence intervals

(<http://rpsychologist.com/d3/CI/>)

## About the visualization

The visualization shows two independent distributions, both with a standard deviation equal to 1. Hence, Cohen's  $d$  is just the difference of their means. In clinical psychology we usually compare two groups at some endpoint. Thus, this visualization can be seen to show the amount of separation of the treatment group and the control group at some endpoint of a study.

## Formulas used

Here the visualization's underlying calculations are presented.

### Cohen's $d$

$$\delta = \frac{\mu_2 - \mu_1}{\sigma}$$

where  $\delta$  is the population parameter of Cohen's  $d$ . Where it is assumed that  $\sigma_1 = \sigma_2 = \sigma$ , i.e. homogeneous population variances. And  $\mu_i$  is the mean of the respective population.

### Cohen's $U_3$

Cohen (1977) defined  $U_3$  as a measure of non-overlap, where "we take the percentage of the A population which the upper half of the cases of the B population exceeds". Cohen's  $d$  can be converted to Cohen's  $U_3$  using the following formula

$$U_3 = \Phi(\delta)$$

where  $\Phi$  is the cumulative distribution function of the standard normal distribution, and  $\delta$  the population Cohen's  $d$ .

### Overlap

Generally called the overlapping coefficient (OVL). Cohen's  $d$  can be converted to OVL using the following formula (Reiser and Faraggi, 1999)

$$OVL = 2\Phi(-|\delta|/2)$$

where  $\Phi$  is the cumulative distribution function of the standard normal distribution, and  $\delta$  the population Cohen's  $d$ .

### Probability of superiority

This is effect size with many names: common language effect size (CL), Area under the receiver operating characteristics (AUC) or just A for its non-parametric version (Ruscio & Mullen, 2012). It is meant to be more intuitive for persons without any training in statistics. The effect size gives the probability that a person picked at random from the treatment group will have a higher score than a person picked at random from the control group. Cohen's  $d$  can be converted CL using the following formula (Ruscio, 2008)

$$CL = \Phi\left(\frac{\delta}{\sqrt{2}}\right)$$

where  $\Phi$  is the cumulative distribution function of the standard normal distribution, and  $\delta$  the population Cohen's  $d$ .

### Number Needed to Treat

NNT is the number of patients we would need to treat with the intervention to achieve one more favorable outcome compared to the control group. Furukawa and Leucht (2011) gives the following formula for converting Cohen's  $d$  into NNT

$$NNT = \frac{1}{\Phi(\delta + \Psi(CER)) - CER}$$

where  $\Phi$  is the cumulative distribution function of the standard normal distribution and  $\Psi$  its inverse, CER is the control group's event rate and  $\delta$  the population Cohen's  $d$ .

**N.B. CER is set to 20 % in the visualization above. You can change this by pressing the ⚙ symbol above the slider.** The definition of an "event" or a "response" is arbitrary and could be defined as the proportion of patients who are in remission, e.g. below some cut-off on a standardized questionnaire. It is possible to convert Cohen's  $d$  into a version of NNT that is invariant to the event rate of the control group. The interested reader should look at Furukawa and Leucht (2011) where a convincing argument is given to why this complicates the interpretation of NNT.

### R code to calculate NNT from Cohen's d

Since many have asked about R code for the formula above, here it is

```
CER <- 0.2
d <- 0.2
1 / (pnorm(d + qnorm(CER)) - CER)
```

## References

- Baguley, T. (2009). Standardized or simple effect size: what should be reported? *British journal of psychology*, 100(Pt 3), 603–17.
- Cohen, J. (1977). *Statistical power analysis for the behavioral sciences*. Routledge.
- Furukawa, T. A., & Leucht, S. (2011). How to obtain NNT from Cohen's d: comparison of two methods. *PloS one*, 6(4).
- Reiser, B., & Faraggi, D. (1999). Confidence intervals for the overlapping coefficient: the normal equal variance case. *Journal of the Royal Statistical Society*, 48(3), 413–418.
- Ruscio, J. (2008). A probability-based measure of effect size: robustness to base rates and other factors. *Psychological methods*, 13(1), 19–30.
- Ruscio, J., & Mullen, T. (2012). Confidence Intervals for the Probability of Superiority Effect Size Measure and the Area Under a Receiver Operating Characteristic Curve. *Multivariate Behavioral Research*, 47(2), 201–223.

## Changelog

Date	Changes
2014-02-03	Added "settings". Let the user change CER, step size and slider's max value
2014-01-13	Initial release