Effect sizes are Strongly Influenced by Changes in Variability

Andrew Brand (King's College London, United Kingdom)
Michael T. Bradley (University of New Brunswick, Saint John, Canada)
Luke MacNeill (University of New Brunswick, Saint John, Canada)

Standardized Effect Sizes

Standardized effect sizes such as Cohen's d are commonly calculated and reported.

- Due to recommendations of the Wilkinson & APA Task Force on Statistical Inference (1999)
- Used in Power Analyses and Meta Analyses
- Are implicit in the calculation of p values

Cohen's d effect size

Cohen's d = $\frac{\text{Mean difference (Signal)}}{\text{SD pooled (Noise)}}$

where $SD_{pooled} = \sqrt{\frac{(n_1 - 1)SD_1^2 + (n_2 - 1)SD_2^2}{n_1 + n_2}}$

- Signal to noise ratio
- Metric free
- Effect is measured in units of SD pooled

Cohen’s d effect size

So…

Cohen’s d will increase if the SD pooled decreases.

Cohen’s d will decrease if the SD pooled increases.

But by how much?

Should we be concerned?

Simulation - Method

Small Effect (d = 0.20)

	Mean	SD
Control Group:	10	2
Experimental Group:	10.40	Varied

Medium Effect (d = 0.50)

	Mean	SD
Control Group:	10	2
Experimental Group:	11.00	Varied

Large Effect (d = 0.80)

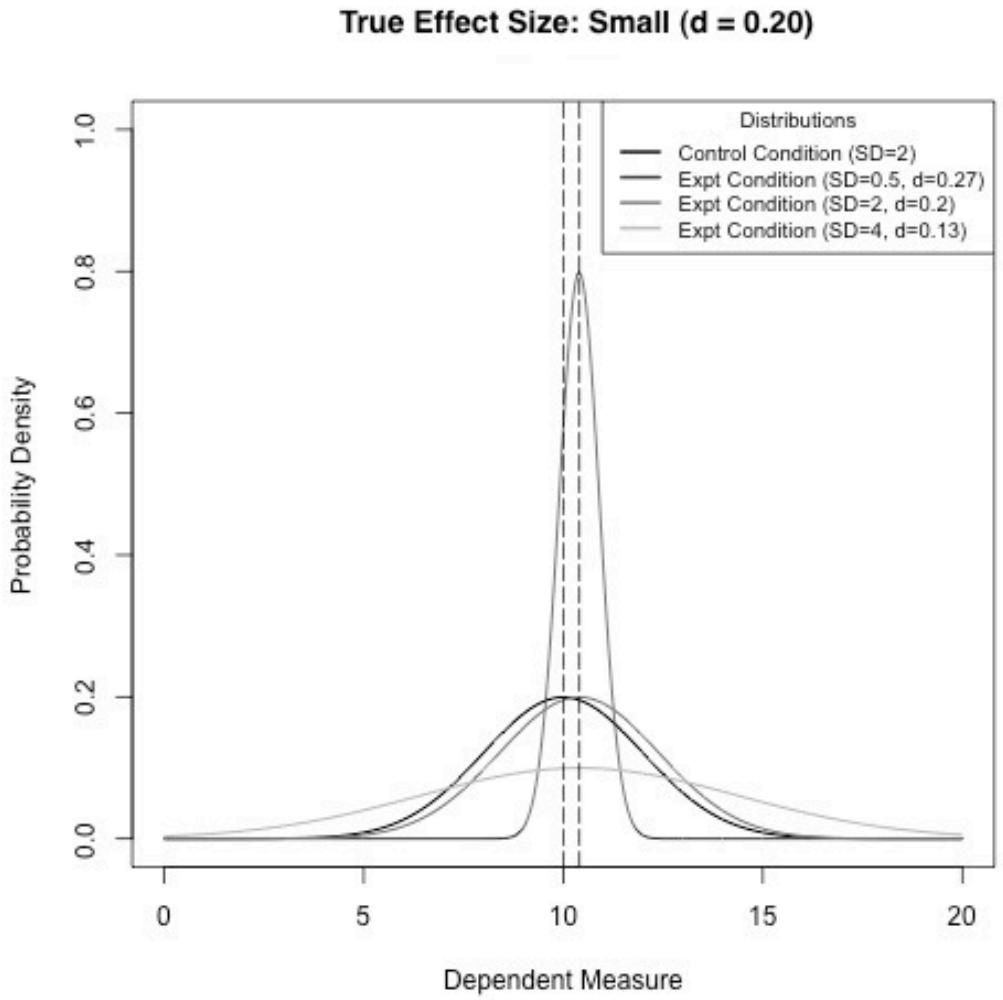
	Mean	SD
Control Group:	10	2
Experimental Group:	11.60	Varied

Simulation - Results

True Effect = Small (d=0.20)								
	Standard Deviation of Group 2 as a % of Group 1							
	25% (0.5)	50% (1.0)	75% (1.5)	100% (2.0)	125% (2.5)	150% (3.0)	175% (3.5)	200% (4.0)
Effect size(d)	0.27	0.25	0.23	0.2	0.18	0.16	0.14	0.13
% of Distortion	35	25	15	0	-10	-20	-30	-35
True Effect = Medium (d=0.50)								
	Standard Deviation of Group 2 as a % of Group 1							
	25% (0.5)	50% (1.0)	75% (1.5)	100% (2.0)	125% (2.5)	150% (3.0)	175% (3.5)	200% (4.0)
Effect size(d)	0.69	0.63	0.57	0.5	0.44	0.39	0.35	0.32
% of Distortion	38	26	14	0	-12	-22	-30	-36
True Effect = Large (d=0.80)								
	Standard Deviation of Group 2 as a % of Group 1							
	25% (0.5)	50% (1.0)	75% (1.5)	100% (2.0)	125% (2.5)	150% (3.0)	175% (3.5)	200% (4.0)
Effect size(d)	1.1	1.01	0.91	0.8	0.71	0.63	0.56	0.51
% of Distortion	38	26	14	0	-11	-21	-30	-36

Note. Level of distortion approximately equal across the 3 different true effect sizes

Simulation - Results



Mean differences are equal but Cohen's ds are markedly different

Theoretical and Practical Implications

The difference between group variance **should** be taken into account when interpreting standardized effect sizes.

Is there a theoretical interpretation for the change / difference in group variance?

What are the practical implications for the change in group variance? Is the difference of practical value?

Theoretical and Practical Implications

Is a reduction in variance important or desirable?

E.g. an educational program may reduce the variance of the academic achievement for a class of students.

Theoretical and Practical Implications

Or are we solely interested in the the mean difference?

E.g. Ziliak & McCloskey (2008) - Weight Loss Pill Thought Experiment

Oomph Pill: Mean Dif = 20 pounds, SD pooled = 10, Cohen's d = 2

Precision Pill: Mean Dif = 5 pounds, SD pooled = 0.5, Cohen's d = 10

Which pill should we recommend to a person who wants to lose weight?

Quick Fixes

Report (the seldom reported) Glass's Delta effect size that is calculate using only the control group's standard deviation.

But sometimes it might not be clear which group should be treated as the control group.

Report the mean difference (a.k.a., simple effect size) as recommend by Baguley (2009).

But note CIs for the mean difference will be affected by group variances.

Overall Recommendations

Do not just report standardized effect sizes blindly.

Be mindful.

Look at the group Means and SDs as well.

If there is a sizable difference between the variance of the groups then consider its theoretical and practical implications.

References

Baguley, T. (2009). Standardized or simple effect size: What should be reported? *British Journal of Psychology, 100*, 603-617.

Wilkinson, L., & APA Task Force on Statistical Inference. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist, 54*, 594-604.

Ziliak, S. T., & McCloskey, D. N. (2008). *The cult of statistical significance: How the standard error costs us jobs, justice, and lives*. Ann Arbor, MI: University of Michigan Press.