```
course = “Estimating the credibility of past research”


lesson_iteration = 12
lesson_title = “Intuitions for effect sizes”


auth = "Ian Hussey"
dept = "Psychology of Digitalisation"
```

Understanding Cohen's d effect size

Differences between groups

Small / Medium / Large

Cohen's d

Standardized Mean Difference

Understanding Cohen's d effect size
Differences between groups

Differences in what?

Cohen's d

Standardized Mean Difference

Understanding Cohen’s d effect size

Differences between groups

Differences in what?

$$d = \frac{M_1 - M_2}{SD_{pooled}}$$

$$SD_{pooled} = \sqrt{\frac{(SD_1^2 + SD_2^2)}{2}}$$

Cohen’s d

Standardized Mean Difference

Understanding Cohen's d effect size
Differences between groups

Differences in what?

Approximately for the sake of understanding:

$$d = \frac{M_1 - M_2}{mean(SD_1, SD_2)}$$

Cohen's d
Standardized Mean Difference

Cohen's d
Standardized Mean Difference

Understanding Cohen's d effect size

Differences between groups

Differences in what?

Approximately for the sake of understanding:

$$d = \frac{M_1 - M_2}{mean(SD_1, SD_2)}$$

How can this go wrong?

Cohen's d

Standardized Mean Difference

How Cohen's d can go wrong

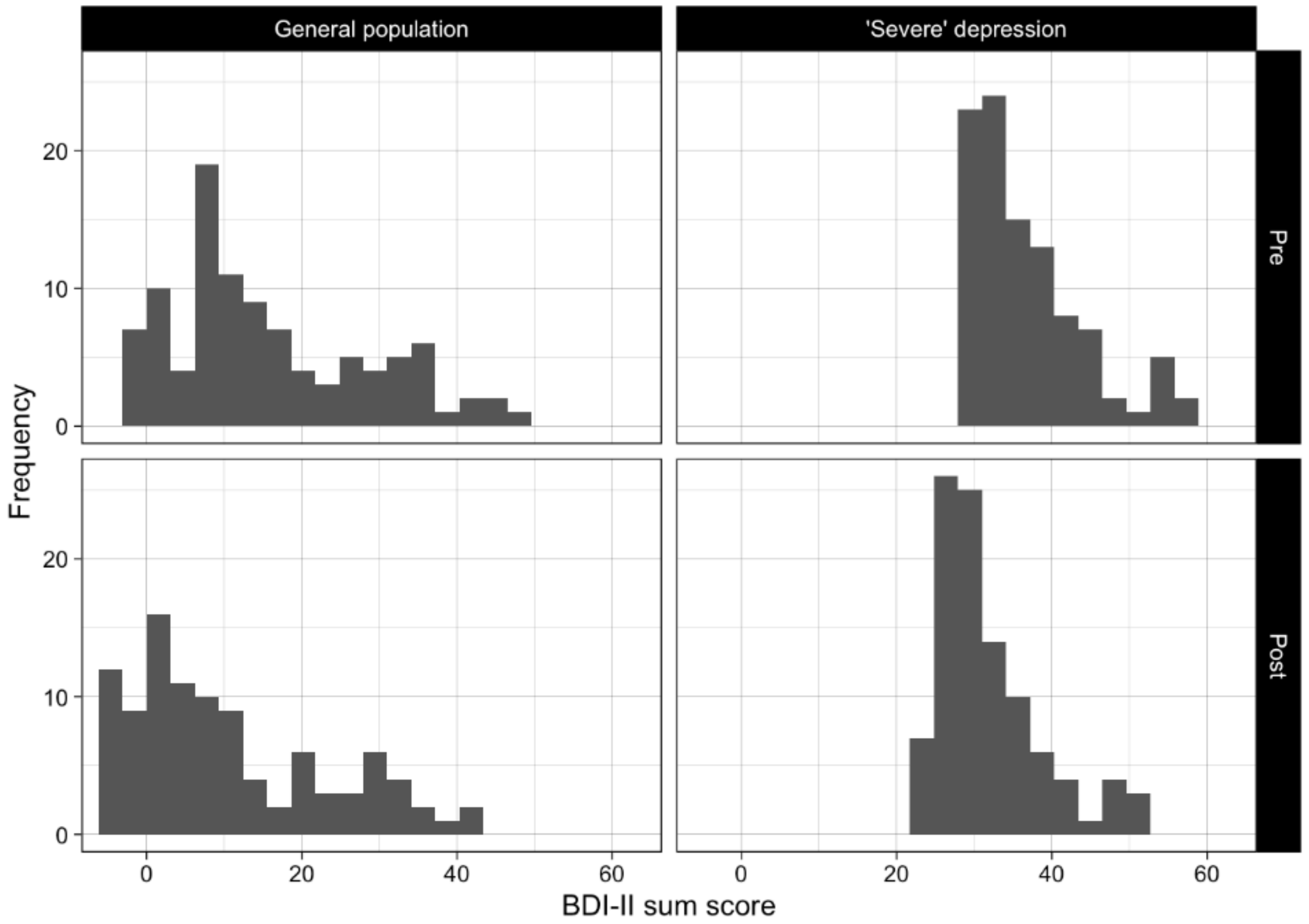
Approximately for the sake of understanding:

$$d = \frac{M_1 - M_2}{mean(SD_1, SD_2)}$$

- Differences in mean as a proportion of Standard Deviation
- “The Standard Error error”

Same difference in means, different Cohen's d

recruitment	n	mean_pre	sd_pre	mean_post	sd_post	mean_diff	cohens_d
'Severe' depression	100	37.0	7.0	32.0	7.0	-5	-0.7
General population	100	15.7	12.6	10.7	12.6	-5	-0.4



Cohen's d

Standardized Mean Difference

aut = "Ian Hussey";

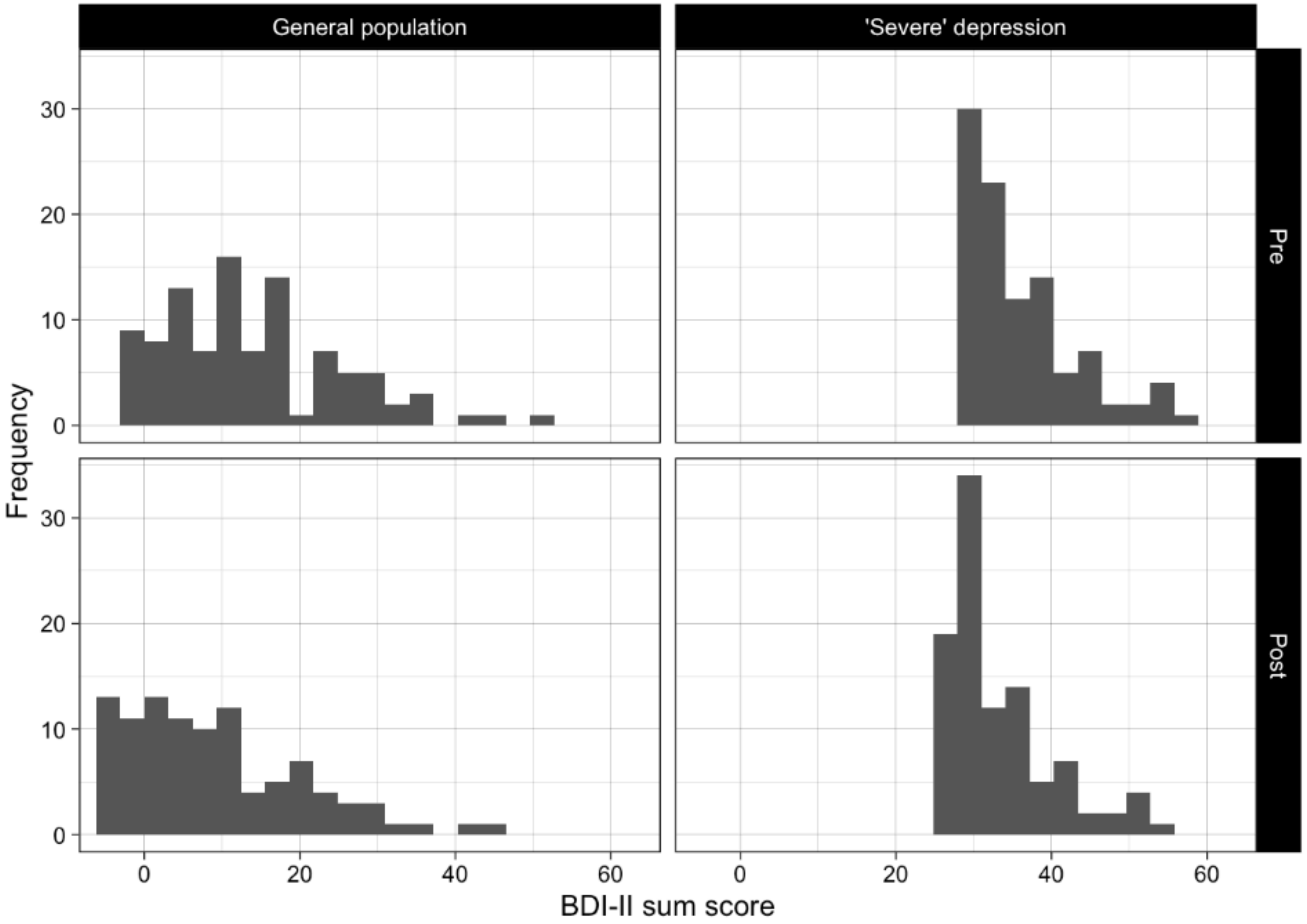
gy"

Same Cohen’s d, different difference in means

Cohen’s d

Standardized Mean Difference

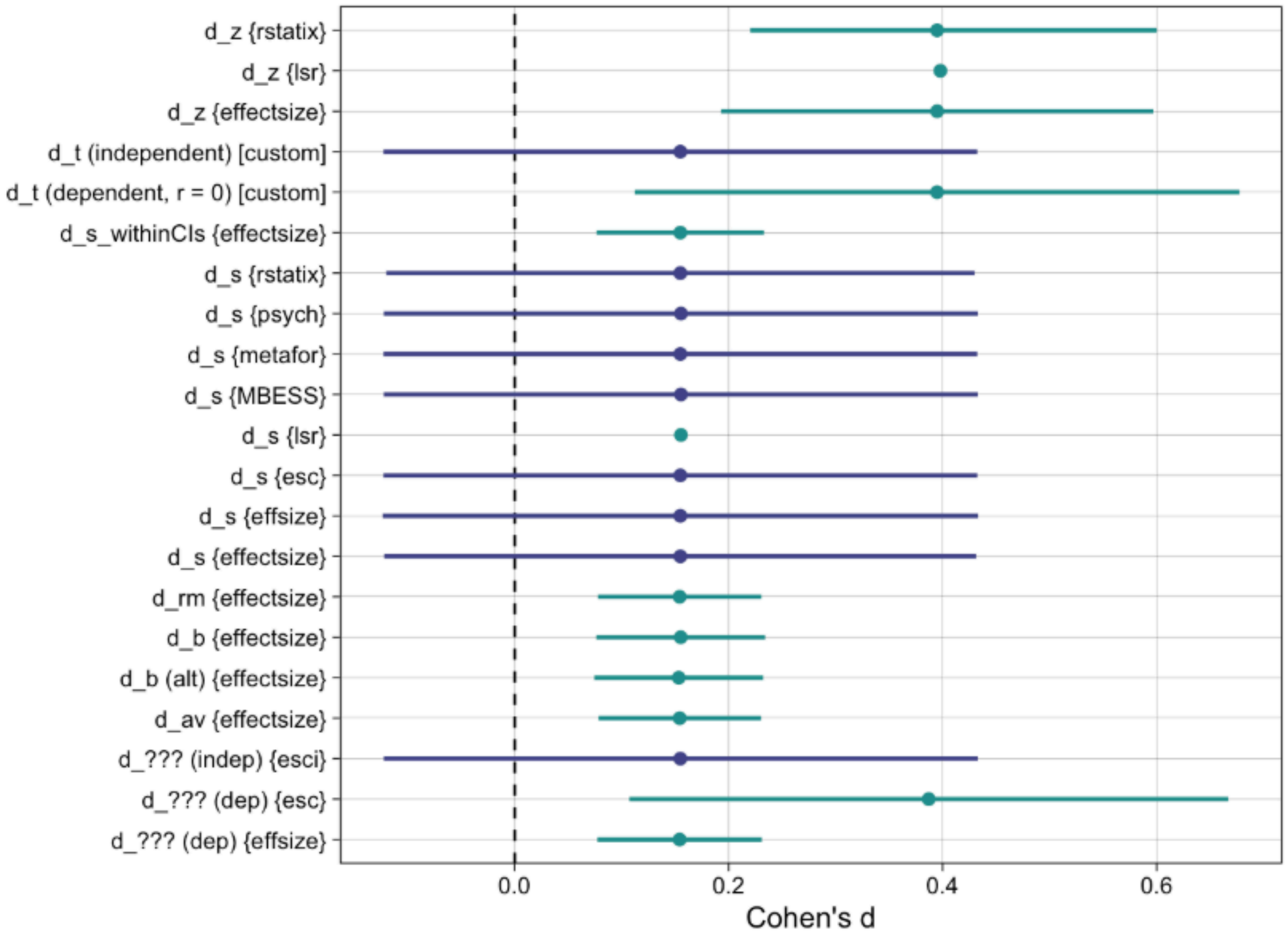
recruitment	n	mean_pre	sd_pre	mean_post	sd_post	mean_diff	cohens_d
‘Severe’ depression	100	36.2	6.7	33.2	6.7	-3	-0.4
General population	100	14.2	11.2	9.2	11.2	-5	-0.4



Different R packages produce different results for Cohen's d 😬

Cohen's d

Standardized Mean Difference

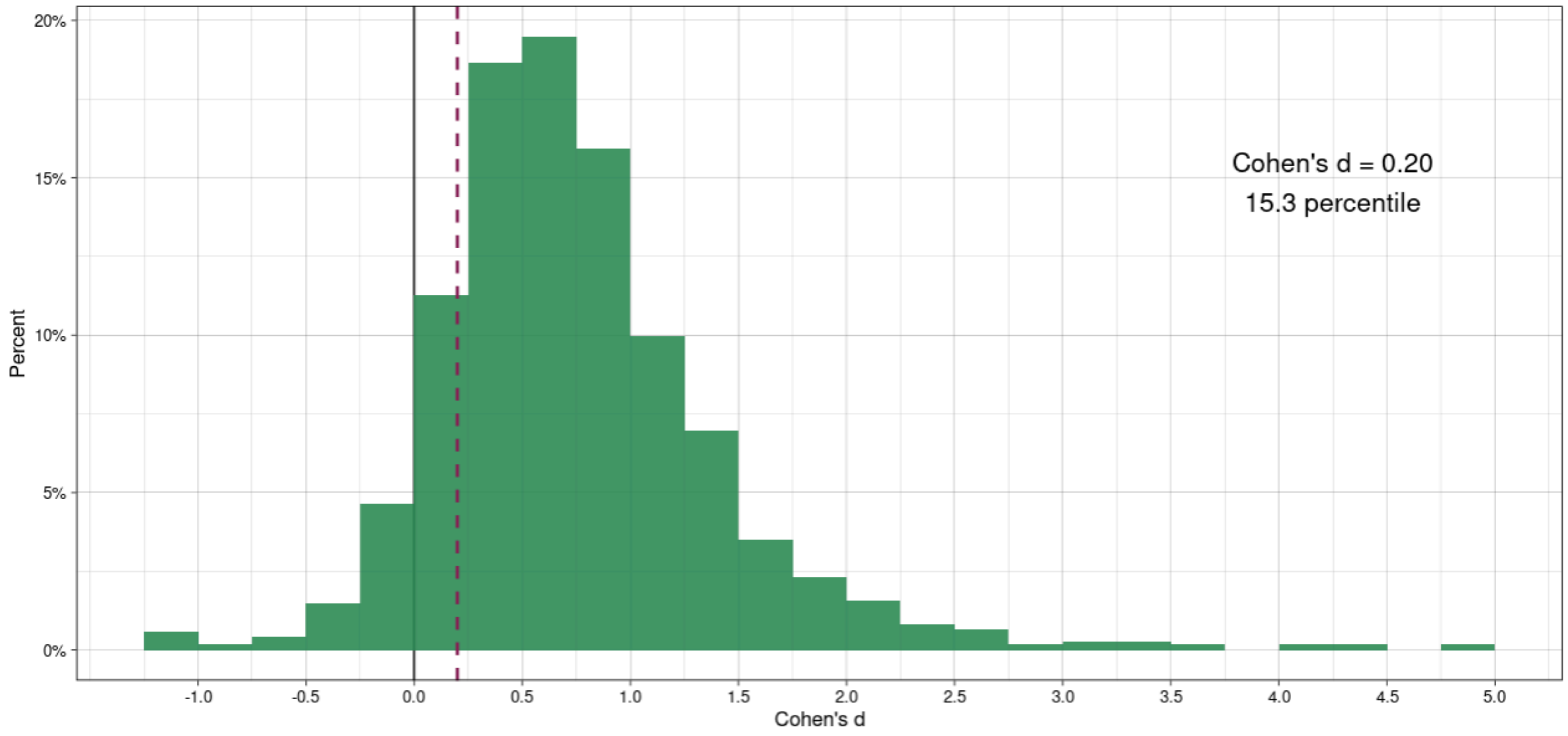


Cohen's d effect sizes in psychotherapy for depression

https://errors.shinyapps.io/effect_size_percentiles/

Cohen's d

Standardized Mean Difference



aut = "Ian Hussey";

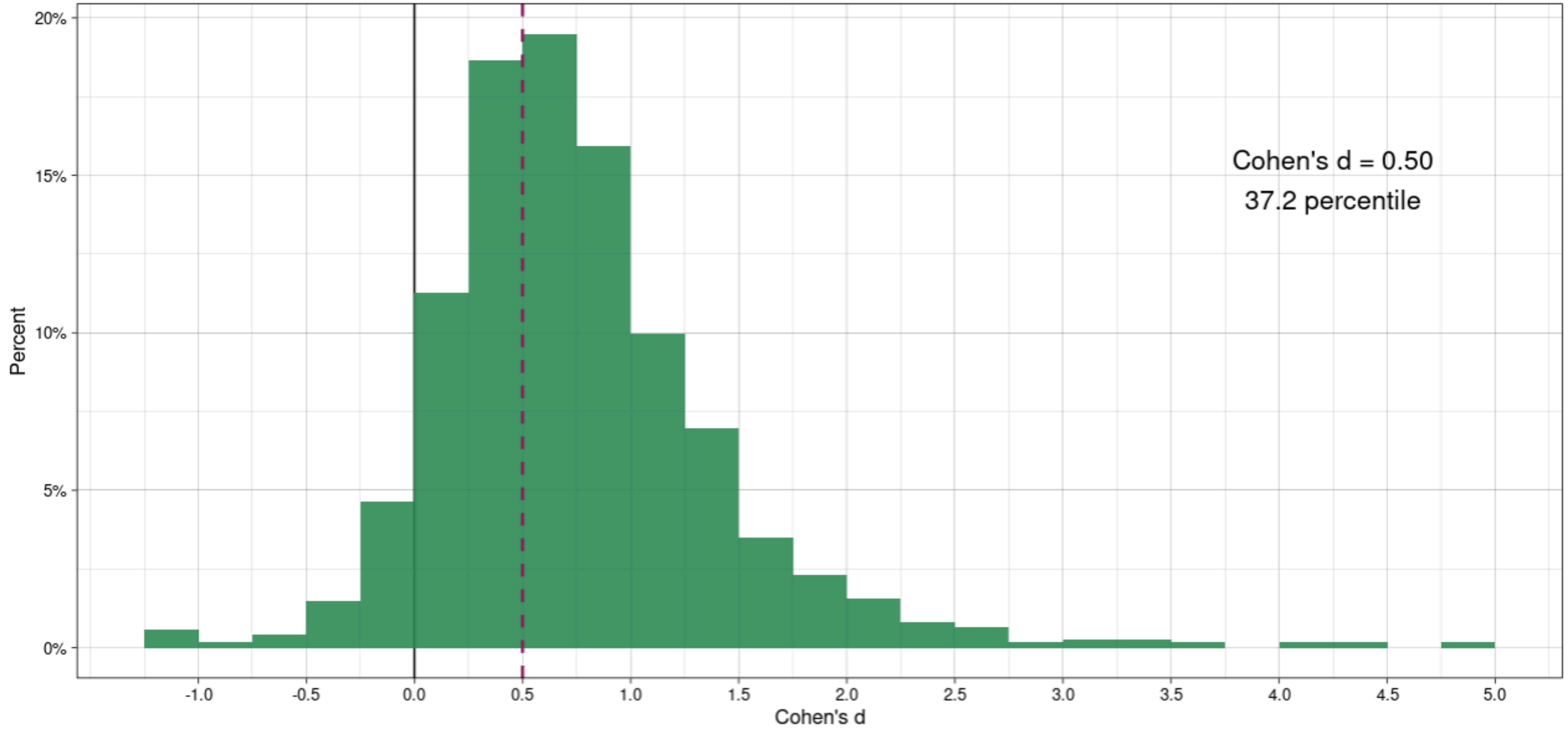
dept = "Psychology of Digitalisation || Digitalisation of Psychology"

Cohen’s d effect sizes in psychotherapy for depression

https://errors.shinyapps.io/effect_size_percentiles/

Cohen’s d

Standardized Mean Difference

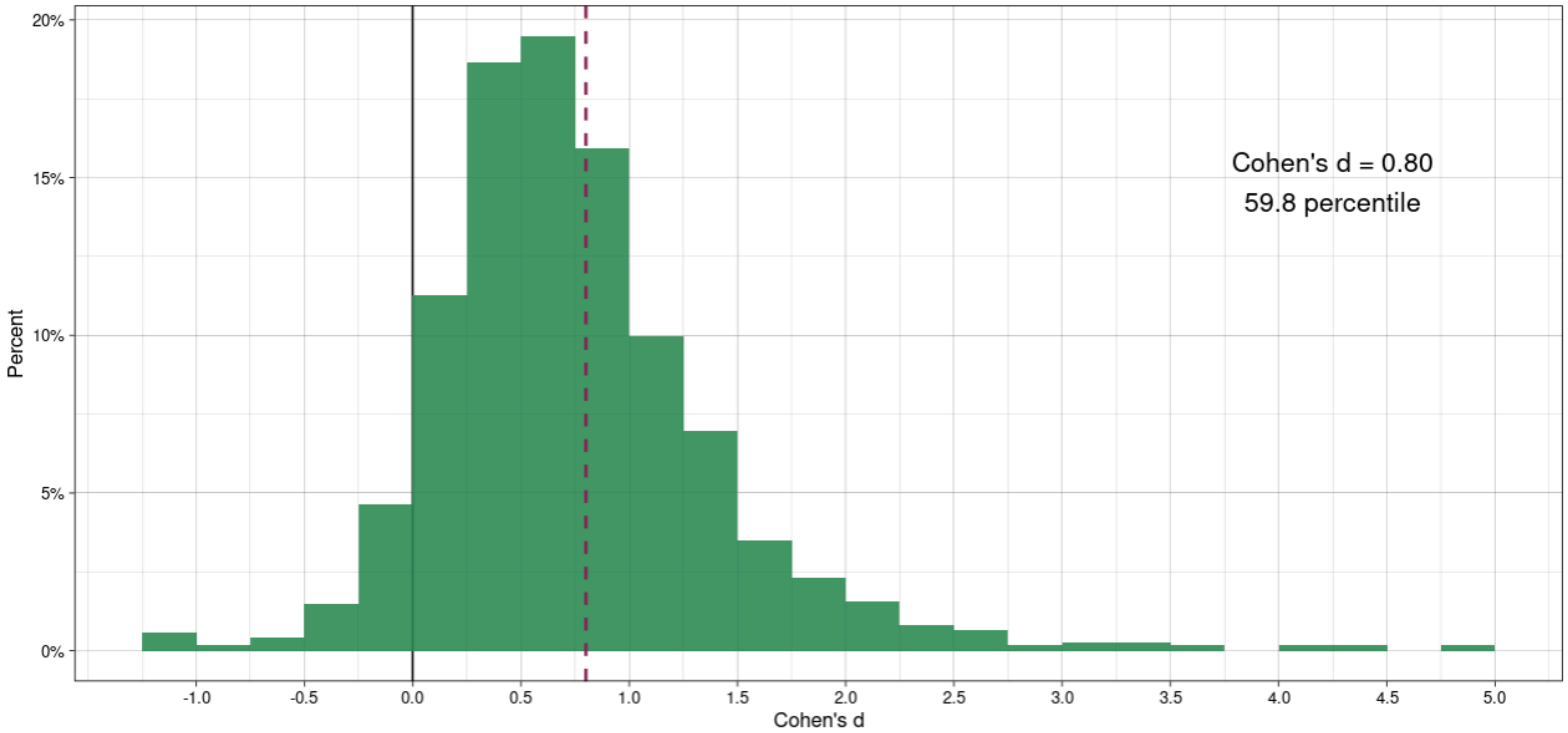


Cohen’s d effect sizes in psychotherapy for depression

https://errors.shinyapps.io/effect_size_percentiles/

Cohen’s d

Standardized Mean Difference



Cohen’s d > 2 are suspicious

Cohen’s d > 5 are silly

Intuitions for Cohen's d effect size

Balcetis & Dunning (2010) Study 3b pretest

Chocolates are more desirable than feces:

Cohen's d = 4.5, 95% CI [3.3, 5.7]

Pretest

In a pretest, we confirmed that chocolates were more desirable than feces. A separate group of participants ($n = 20$) rated how appealing, positive, likeable, attractive, and interesting both objects were. It should come as no surprise that averages across these qualities indicated that the chocolates ($M = 5.5$) were more desirable than the feces ($M = 2.1$), paired $t(19) = 17.44$, $p_{rep} = .99$, $p < .001$, $d = 4.52$.

Intuitions for Cohen’s d effect size

Streck & Kessels (2024) Gender stereotypes in children

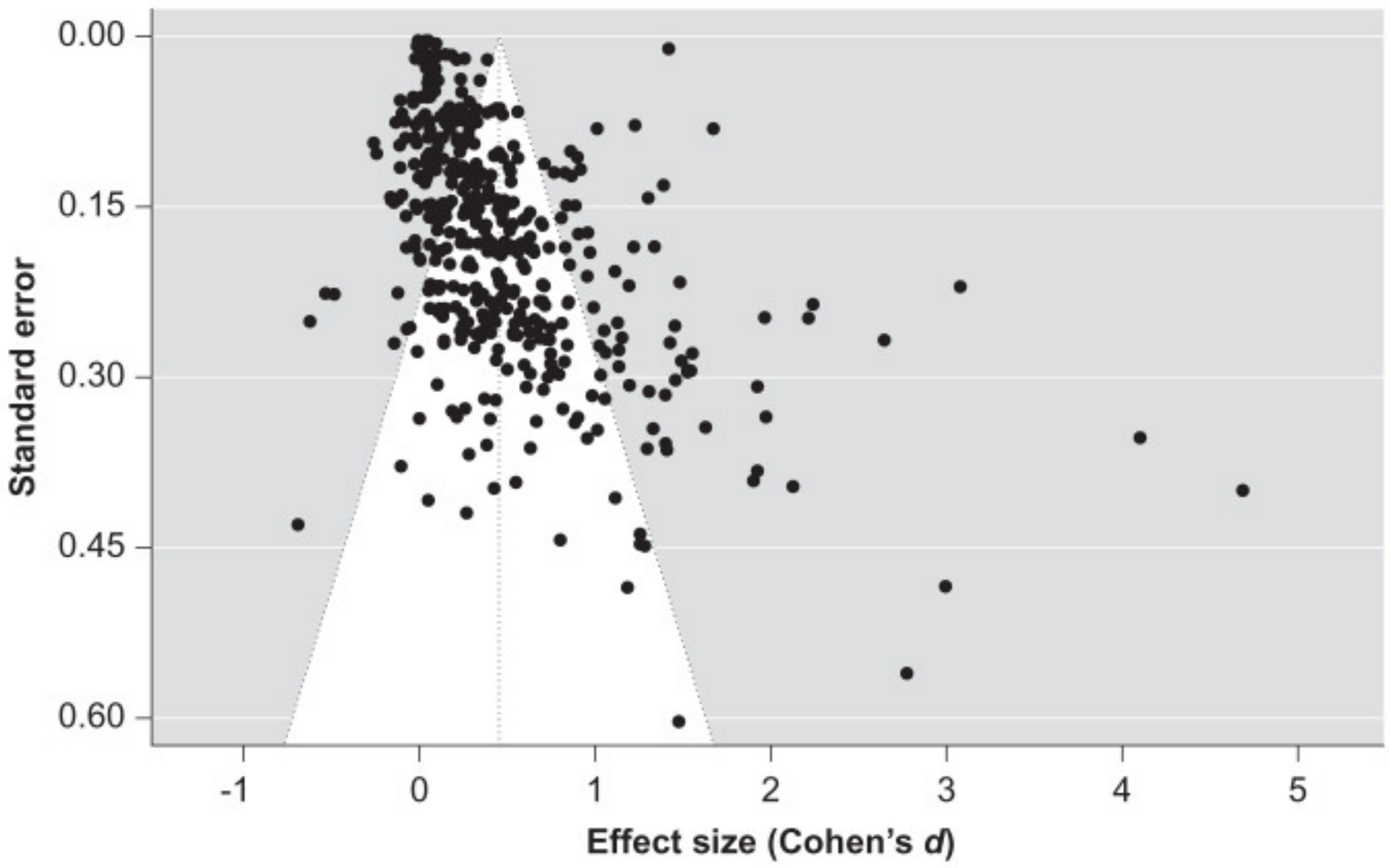
Asking children
“Boys tend to wear skirts” vs. “Girls tend to wear skirts”:

Cohen’s d = 5.5, 95% CI [4.9, 6.1]

Cohen’s d
Standardized Mean Difference

Intuitions for Cohen’s d effect size

Mertens et al. (2021) The effectiveness of nudging



Cohen’s d

Standardized Mean Difference

Intuitions for Cohen's d effect size

Mertens et al. (2021) The effectiveness of nudging

83. *A Geier, B Wansink, P Rozin, Red potato chips: Segmentation cues can substantially decrease food intake. *Heal. Psychol.* **31**, 398–401 (2012).

191. *M Shimizu, CR Payne, B Wansink, When snacks become meals: How hunger and environmental cues bias food intake. *Int. J. Behav. Nutr. Phys. Activity* **7**, 1–6 (2010).

217. *E van Kleef, M Shimizu, B Wansink, Serving bowl selection biases the amount of food served. *J. Nutr. Educ. Behav.* **44**, 66–70 (2012).

220. *B Wansink, AS Hanks, Slim by design: Serving healthy foods first in buffet lines improves overall meal selection. *PLoS ONE* **8**, 1–5 (2013).
221. *B Wansink, J Kim, Bad popcorn in big buckets: Portion size can influence intake as much as taste. *J. Nutr. Educ. Behav.* **37**, 242–245 (2005).
222. *B Wansink, K van Ittersum, Bottoms up! the influence of elongation on pouring and consumption volume. *J. Consumer Res.* **30**, 455–463 (2003).
223. *B Wansink, A Cardello, J North, Fluid consumption and the potential role of canteen shape in minimizing dehydration. *Mil. Medicine* **170**, 871–873 (2005).
224. *B Wansink, JE Painter, J North, Bottomless bowls: Why visual cues of portion size may influence intake. *Obes. Res.* **13**, 93–100 (2005).
225. *B Wansink, D Soman, KC Herbst, Larger partitions lead to larger sales: Divided grocery carts alter purchase norms and increase sales. *J. Bus. Res.* **75**, 202–209 (2017).

226. *B Wansink, K van Ittersum, JE Painter, Ice cream illusions: Bowls, spoons, and self-served portion sizes. *Am. J. Prev. Medicine* **31**, 240–243 (2006).
227. *B Wansink, K van Ittersum, CR Payne, Larger bowl size increases the amount of cereal children request, consume, and waste. *The J. Pediatr.* **164**, 323–326 (2014).

Intuitions for Cohen’s d effect size

Mertens et al. (2021) The effectiveness of nudging

83. *A Geier, B Wansink, P Rozin, Red potato chips: Segmentation cues can substantially decrease food intake. *Heal. Psychol.* **31**, 398–401 (2012).

191. *M Shimizu, CR Payne, B Wansink, When snacks become meals: How hunger and environmental cues bias food intake. *Int. J. Behav. Nutr. Phys. Activity* **7**, 1–6 (2010).

217. *E van Kleef, M Shimizu, B Wansink, Serving bowl selection biases the amount of food served. *J. Nutr. Educ. Behav.* **44**, 66–70 (2012).

220. *B Wansink, AS Hanks, Slim by design: Serving healthy foods first in buffet lines improves overall meal selection. *PLoS ONE* **8**, 1–5 (2013).
221. *B Wansink, J Kim, Bad popcorn in big buckets: Portion size can influence intake as much as taste. *J. Nutr. Educ. Behav.* **37**, 242–245 (2005).
222. *B Wansink, K van Ittersum, Bottoms up! the influence of elongation on pouring and consumption volume. *J. Consumer Res.* **30**, 455–463 (2003).
223. *B Wansink, A Cardello, J North, Fluid consumption and the potential role of canteen shape in minimizing dehydration. *Mil. Medicine* **170**, 871–873 (2005).
224. *B Wansink, JE Painter, J North, Bottomless bowls: Why visual cues of portion size may influence intake. *Obes. Res.* **13**, 93–100 (2005).
225. *B Wansink, D Soman, KC Herbst, Larger partitions lead to larger sales: Divided grocery carts alter purchase norms and increase sales. *J. Bus. Res.* **75**, 202–209 (2017).

226. *B Wansink, K van Ittersum, JE Painter, Ice cream illusions: Bowls, spoons, and self-served portion sizes. *Am. J. Prev. Medicine* **31**, 240–243 (2006).
227. *B Wansink, K van Ittersum, CR Payne, Larger bowl size increases the amount of cereal children request, consume, and waste. *The J. Pediatr.* **164**, 323–326 (2014).

CAREERS COMMENTARY JOURNALS Science

SCIENCEINSIDER | PEOPLE & EVENTS

Cornell nutrition scientist resigns after retractions and research misconduct finding

Misreporting and sloppy statistics undermine flashy findings on unconscious eating habits

21 SEP 2018 • BY KELLY SERVICK

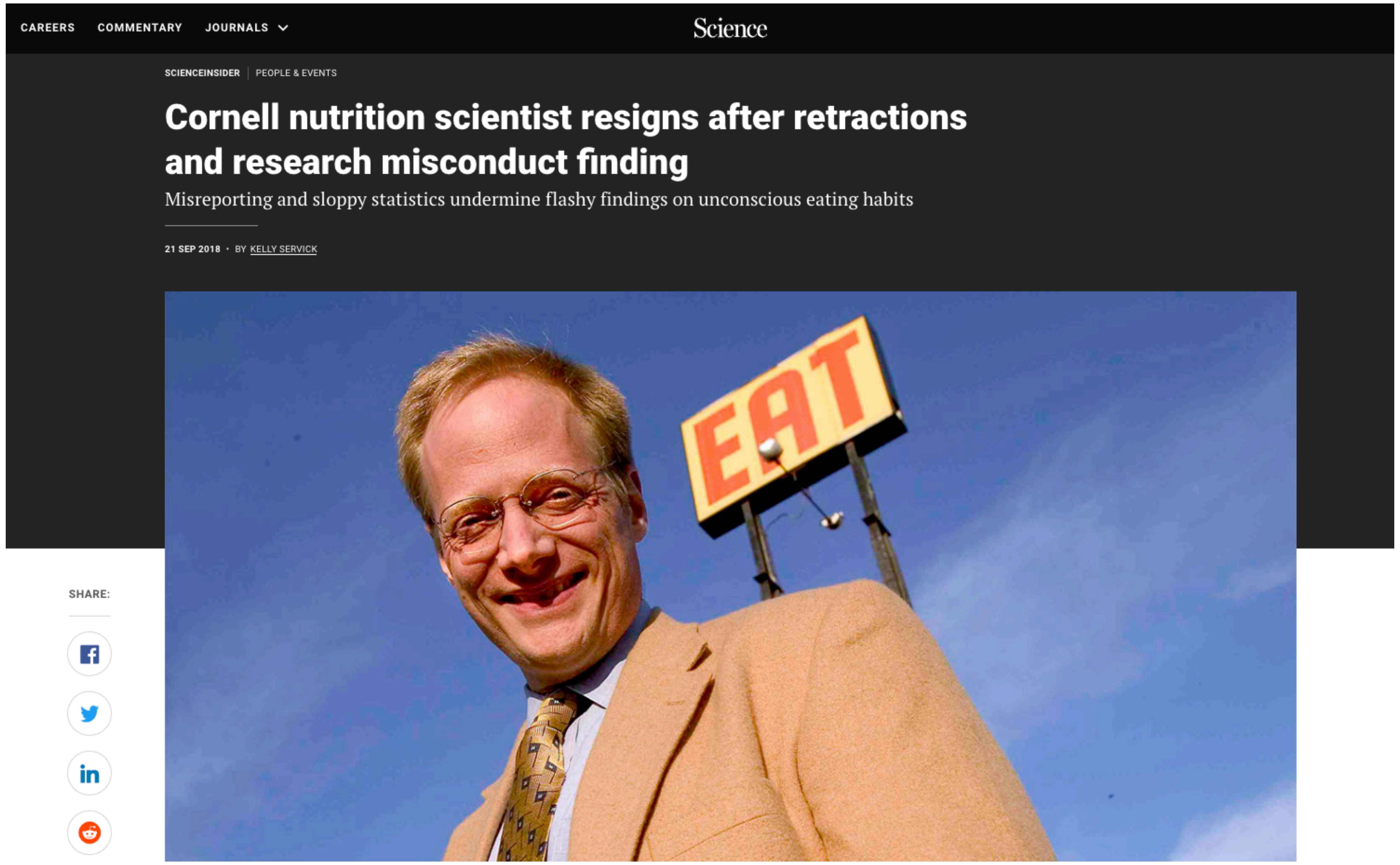
SHARE:

Intuitions for Cohen's d effect size

Mertens et al. (2021) The effectiveness of nudging

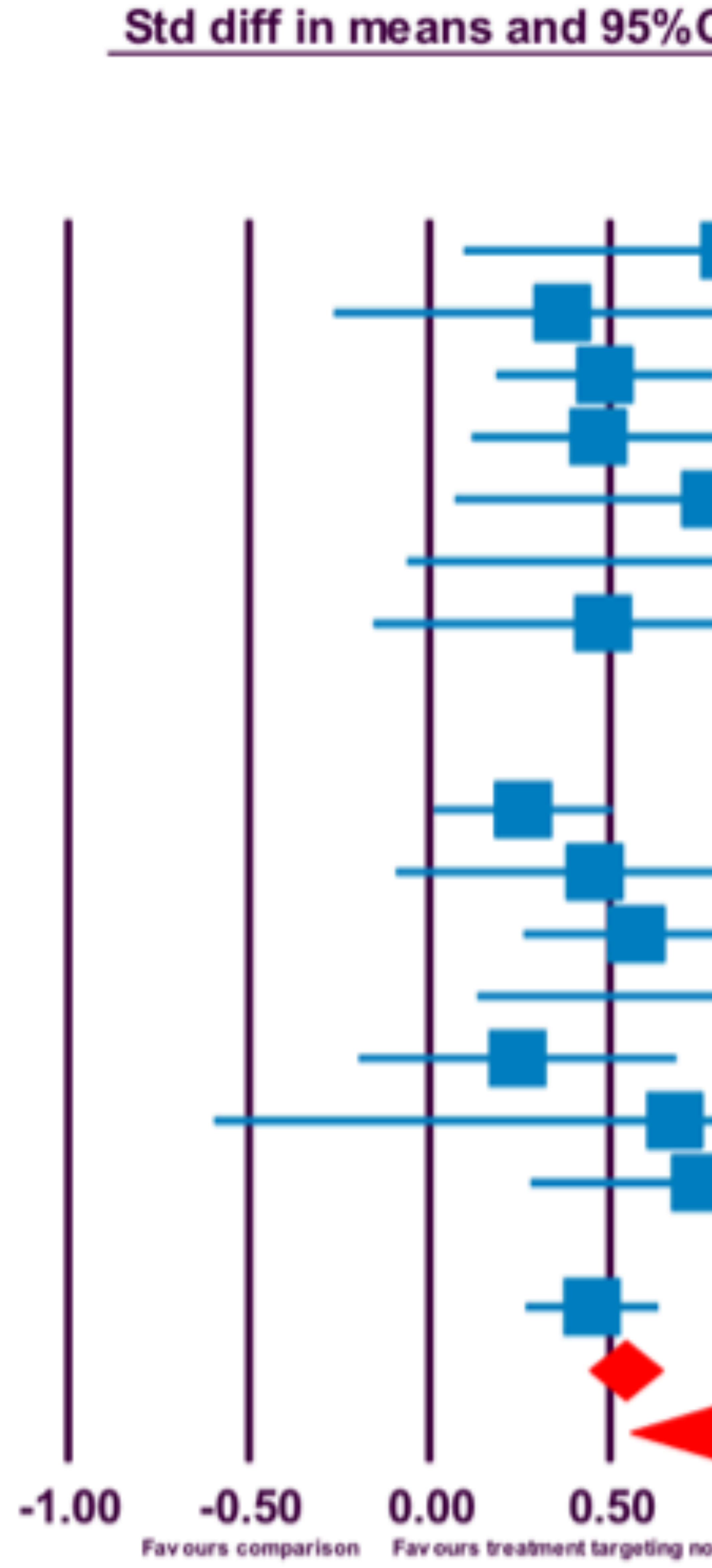
Cohen's d

Standardized Mean Difference



Intuitions for Cohen's d effect size

Model	Study name	Std diff in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
	Barnhofer et al. (2009)	0.829	0.374	0.140	0.095	1.563	2.213	0.027
	Clarke et al. (2014)	0.368	0.323	0.104	-0.265	1.001	1.140	0.254
	Eisendrath et al. (2016)	0.486	0.154	0.024	0.184	0.789	3.151	0.002
	Fonagy et al. (2015)	0.467	0.179	0.032	0.117	0.817	2.612	0.009
	Gloster et al. (2015)	0.777	0.361	0.130	0.069	1.484	2.152	0.031
	Harley et al. (2008)	0.881	0.481	0.232	-0.062	1.824	1.831	0.067
	Hartmann Souza et al. (2016)	0.480	0.324	0.105	-0.155	1.116	1.481	0.139
	Hinton et al. (2005)	2.226	0.402	0.162	1.438	3.015	5.532	0.000
	Isasi et al. (2010)	5.233	0.665	0.442	3.929	6.536	7.868	0.000
	Kocsis et al. (2009)	0.259	0.126	0.016	0.012	0.506	2.055	0.040
	Ludman et al. (2007)	0.458	0.281	0.079	-0.093	1.008	1.629	0.103
	Mantani et al. (2017)	0.572	0.159	0.025	0.260	0.885	3.592	0.000
	Moore & Blackburn (1997)	1.653	0.776	0.602	0.132	3.173	2.130	0.033
	Nakagawa et al. (2017)	0.243	0.224	0.050	-0.197	0.683	1.084	0.278
	Otto et al. (2003)	0.679	0.650	0.423	-0.596	1.954	1.044	0.297
	Town et al. (2017)	0.748	0.239	0.057	0.280	1.216	3.131	0.002
	Watanabe et al. (2011)	1.811	0.391	0.153	1.044	2.578	4.627	0.000
	Wiles et al. (2013)	0.450	0.094	0.009	0.266	0.633	4.807	0.000
Fixed		0.540	0.050	0.003	0.442	0.638	10.765	0.000
Random		0.818	0.134	0.018	0.556	1.081	6.106	0.000



Cohen's d

aut =

Effectiveness at post-treatment on symptom reduction

Intuitions for Cohen's d effect size

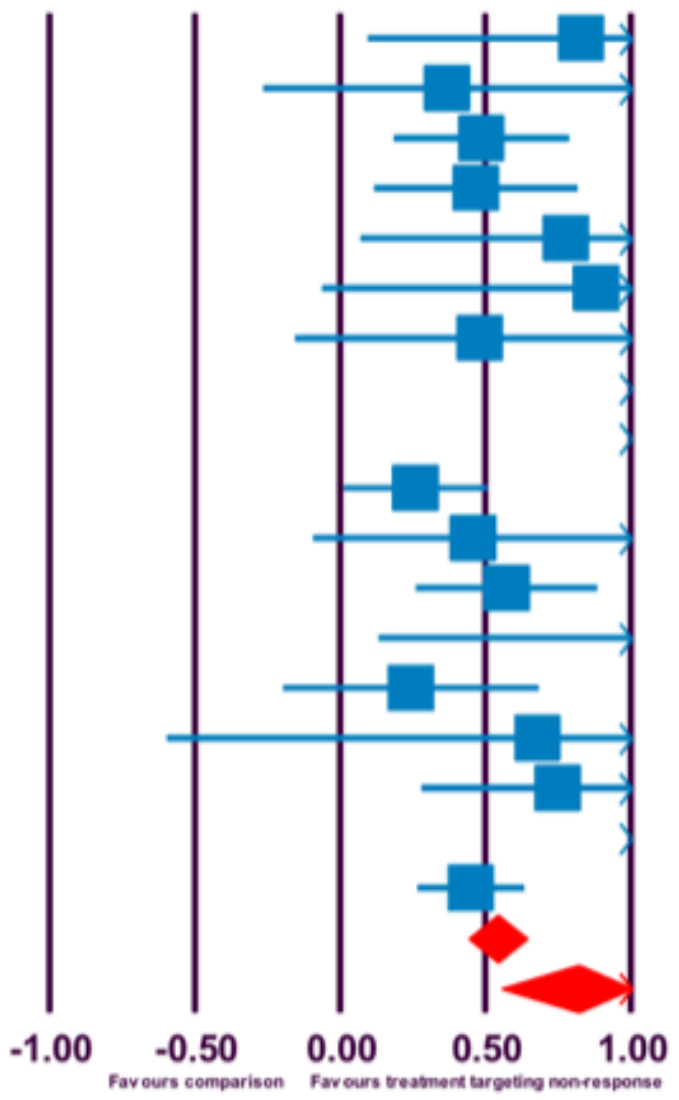
Gloster et al. (2020) Treating treatment non-responders

Cohen's d

Standardized Mean Difference

Model	Study name	Std diff in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
	Barnhofer et al. (2009)	0.829	0.374	0.140	0.095	1.563	2.213	0.027
	Clarke et al. (2014)	0.368	0.323	0.104	-0.265	1.001	1.140	0.254
	Eisendrath et al. (2016)	0.486	0.154	0.024	0.184	0.789	3.151	0.002
	Fonagy et al. (2015)	0.467	0.179	0.032	0.117	0.817	2.612	0.009
	Gloster et al. (2015)	0.777	0.361	0.130	0.069	1.484	2.152	0.031
	Harley et al. (2008)	0.881	0.481	0.232	-0.062	1.824	1.831	0.067
	Hartmann Souza et al. (2016)	0.480	0.324	0.105	-0.155	1.116	1.481	0.139
	Hinton et al. (2005)	2.226	0.402	0.162	1.438	3.015	5.532	0.000
	Isasi et al. (2010)	5.233	0.665	0.442	3.929	6.536	7.868	0.000
	Kocsis et al. (2009)	0.259	0.126	0.016	0.012	0.506	2.055	0.040
	Ludman et al. (2007)	0.458	0.281	0.079	-0.093	1.008	1.629	0.103
	Mantani et al. (2017)	0.572	0.159	0.025	0.260	0.885	3.592	0.000
	Moore & Blackburn (1997)	1.653	0.776	0.602	0.132	3.173	2.130	0.033
	Nakagawa et al. (2017)	0.243	0.224	0.050	-0.197	0.683	1.084	0.278
	Otto et al. (2003)	0.679	0.650	0.423	-0.596	1.954	1.044	0.297
	Town et al. (2017)	0.748	0.239	0.057	0.280	1.216	3.131	0.002
	Watanabe et al. (2011)	1.811	0.391	0.153	1.044	2.578	4.627	0.000
	Wiles et al. (2013)	0.450	0.094	0.009	0.266	0.633	4.807	0.000
Fixed		0.540	0.050	0.003	0.442	0.638	10.765	0.000
Random		0.818	0.134	0.018	0.556	1.081	6.106	0.000

Std diff in means and 95% CI



Effectiveness at post-treatment on symptom reduction

Intuitions for Cohen's d effect size

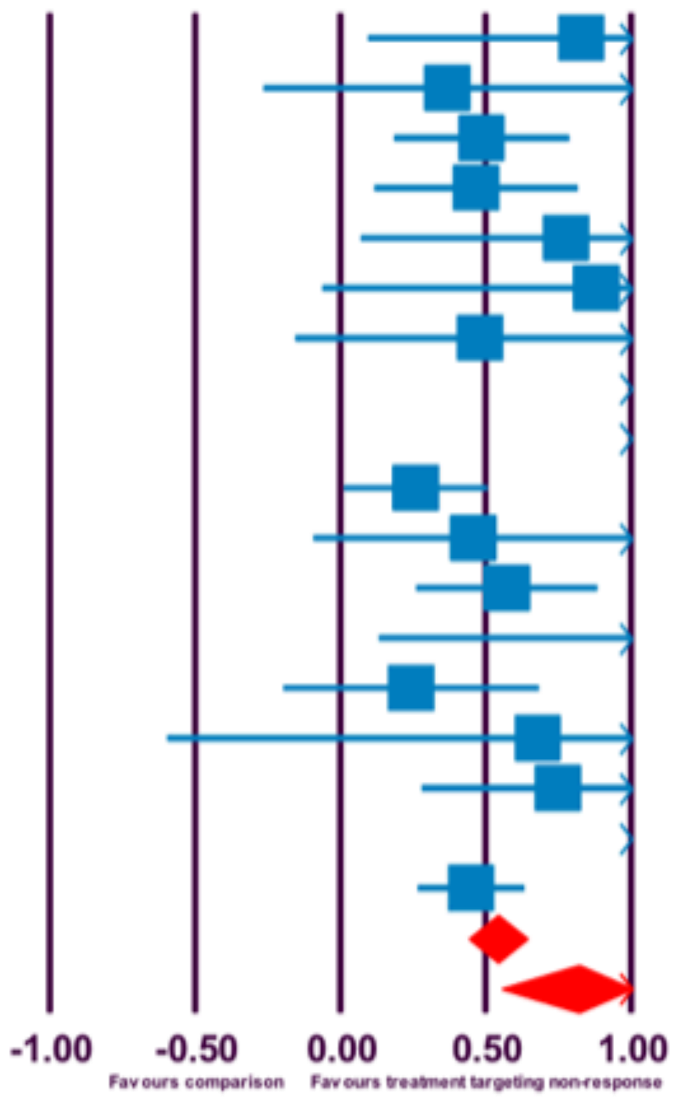
Gloster et al. (2020) Treating treatment non-responders

Cohen's d

Standardized Mean Difference

Model	Study name	Std diff in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
	Barnhofer et al. (2009)	0.829	0.374	0.140	0.095	1.563	2.213	0.027
	Clarke et al. (2014)	0.368	0.323	0.104	-0.265	1.001	1.140	0.254
	Eisendrath et al. (2016)	0.486	0.154	0.024	0.184	0.789	3.151	0.002
	Fonagy et al. (2015)	0.467	0.179	0.032	0.117	0.817	2.612	0.009
	Gloster et al. (2015)	0.777	0.361	0.130	0.069	1.484	2.152	0.031
	Harley et al. (2008)	0.881	0.481	0.232	-0.062	1.824	1.831	0.067
	Hartmann Souza et al. (2016)	0.480	0.324	0.105	-0.155	1.116	1.481	0.139
	Hinton et al. (2005)	2.226	0.402	0.162	1.438	3.015	5.532	0.000
	Isasi et al. (2010)	5.233	0.665	0.442	3.929	6.536	7.868	0.000
	Kocsis et al. (2009)	0.259	0.126	0.016	0.012	0.506	2.055	0.040
	Ludman et al. (2007)	0.458	0.281	0.079	-0.093	1.008	1.629	0.103
	Mantani et al. (2017)	0.572	0.159	0.025	0.260	0.885	3.592	0.000
	Moore & Blackburn (1997)	1.653	0.776	0.602	0.132	3.173	2.130	0.033
	Nakagawa et al. (2017)	0.243	0.224	0.050	-0.197	0.683	1.084	0.278
	Otto et al. (2003)	0.679	0.650	0.423	-0.596	1.954	1.044	0.297
	Town et al. (2017)	0.748	0.239	0.057	0.280	1.216	3.131	0.002
	Watanabe et al. (2011)	1.811	0.391	0.153	1.044	2.578	4.627	0.000
	Wiles et al. (2013)	0.450	0.094	0.009	0.266	0.633	4.807	0.000
Fixed		0.540	0.050	0.003	0.442	0.638	10.765	0.000
Random		0.818	0.134	0.018	0.556	1.081	6.106	0.000

Std diff in means and 95%CI



Effectiveness at post-treatment on symptom reduction

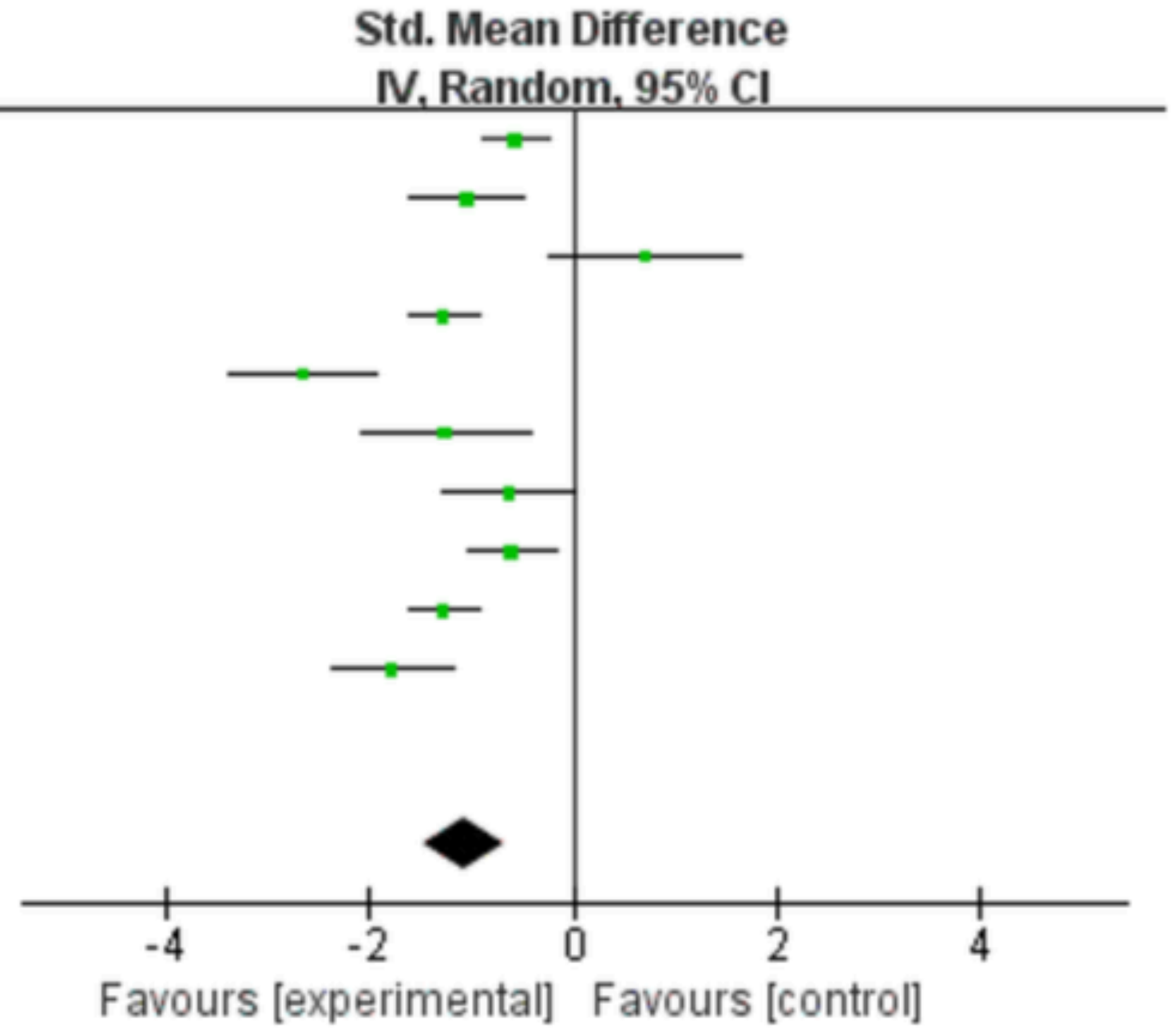
Intuitions for Cohen’s d effect size

Zhao et al. (2023) Effect of acceptance and commitment therapy for depressive disorders

Cohen’s d
Standardized Mean Difference

	Experimental			Control				Std. Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI
Wendy T.M.Pots 2016	14.68	8.05	71	19.34	8.55	78	11.7%	-0.56 [-0.89, -0.23]
Situ Yuyi 2022	12.51	3.26	30	15.52	2.35	30	10.3%	-1.05 [-1.59, -0.50]
Shima Tamannaei Far 2017	28.2	16.28	10	18.54	7.65	9	7.4%	0.71 [-0.22, 1.65]
Ren Zhihong 2012	13.61	9.48	92	26.05	10.06	76	11.7%	-1.27 [-1.60, -0.94]
Mehdi Zemestani 2020	20.7	3.4	26	32.57	5.15	30	8.8%	-2.64 [-3.37, -1.91]
Louise Hayes 2011	66.05	3.24	19	70.68	4.2	11	8.2%	-1.25 [-2.06, -0.43]
Lappalainen 2015	13.34	6.75	18	17.85	7.34	20	9.4%	-0.62 [-1.28, 0.03]
Ernst T.Bohlmeijer 2011	15.94	10.37	39	22.07	9.99	42	11.0%	-0.60 [-1.04, -0.15]
Chunxiao Zhao 2022	13.61	9.48	92	26.05	10.06	76	11.7%	-1.27 [-1.60, -0.94]
Chen Juan 2021	49.36	2.18	30	54.36	3.29	30	9.8%	-1.77 [-2.37, -1.17]
A-Tjak 2018	11.08	1.36	33	7.91	1.51	25	0.0%	2.19 [1.53, 2.86]
Total (95% CI)			**427**			**402**	**100.0%**	**-1.05 [-1.44, -0.66]**

Heterogeneity: $Tau^2 = 0.31$; $Chi^2 = 55.18$, df = 9 (P < 0.00001); $I^2 = 84\%$
Test for overall effect: Z = 5.27 (P < 0.00001)



Intuitions for Pearson’s r

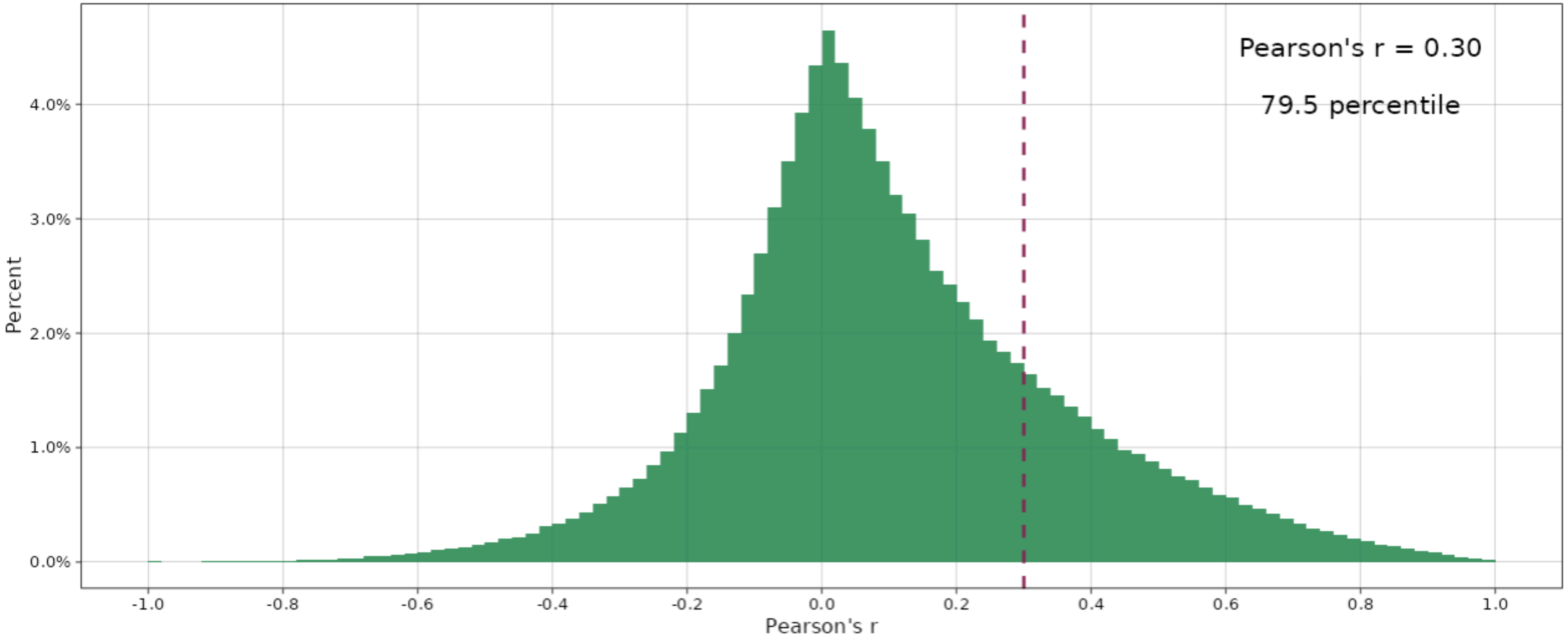
Pearsons’ r

Correlation coefficient

$$r = \frac{Cov(X,Y)}{SD_X \times SD_Y}$$

Intuitions for Pearson's r

https://errors.shinyapps.io/correlations_percentiles/



Pearsons' r

Correlation coefficient

Intuitions for Cronbach’s alpha

https://errors.shinyapps.io/cronbachs_alpha_percentiles/

Cronbach’s alpha

Internal consistency