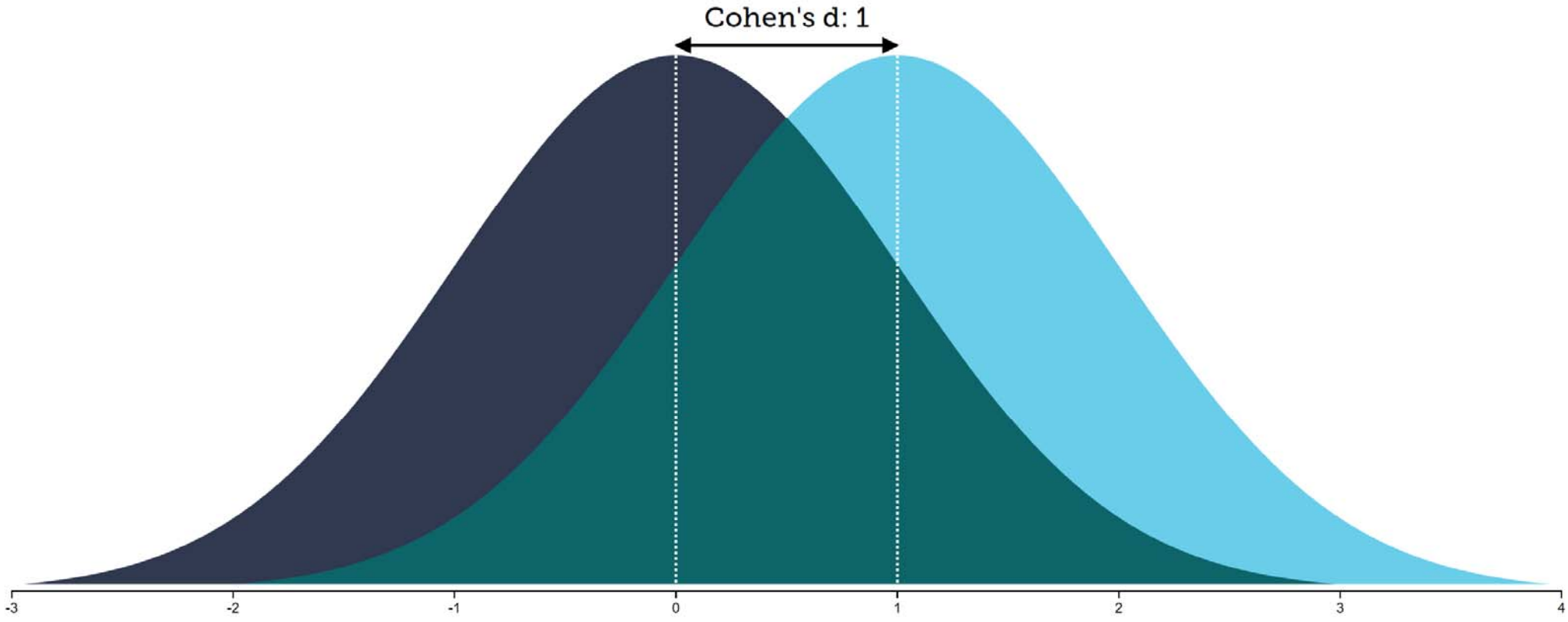
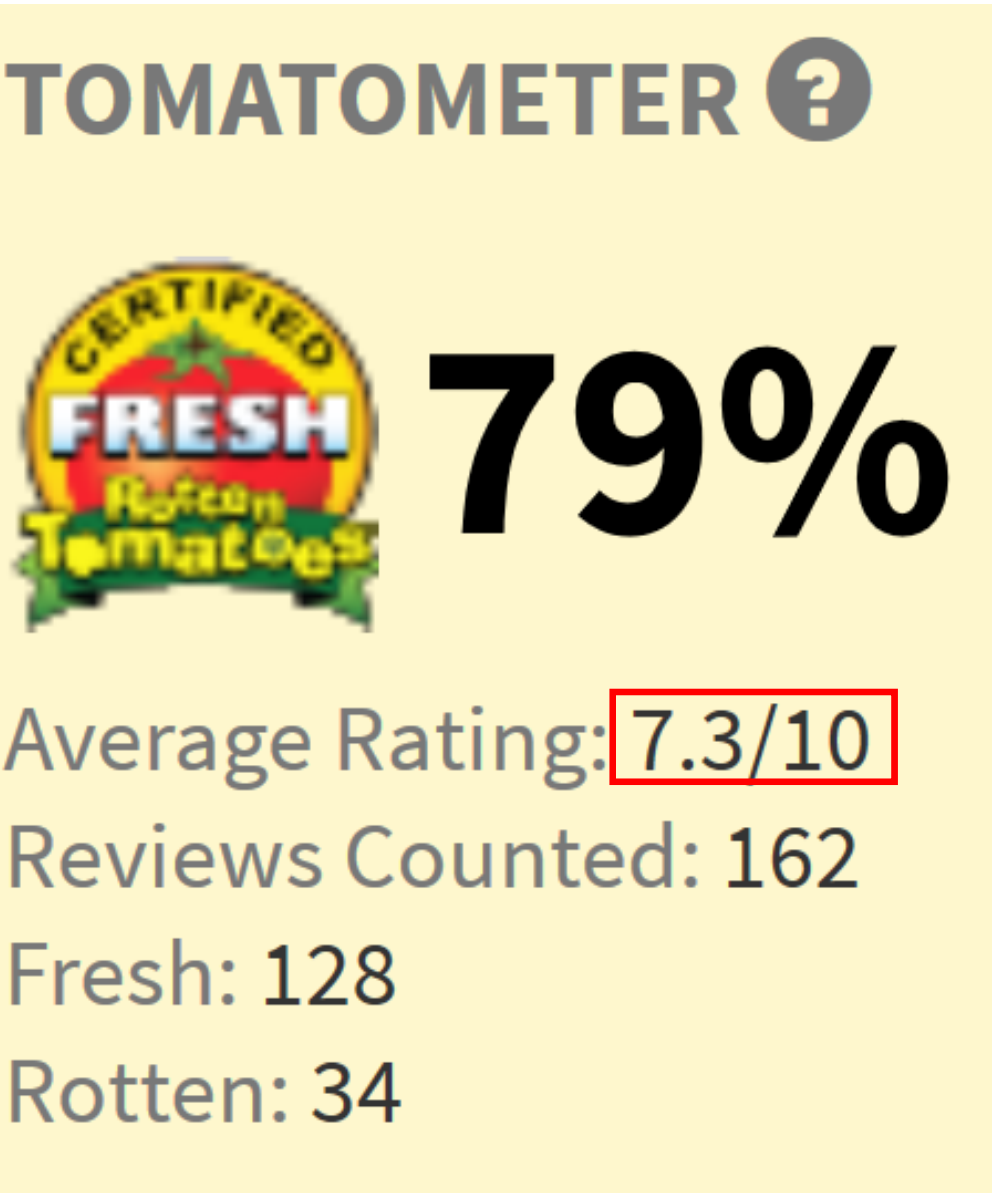
Cohen’s d



Do movie ratings differ between websites? And if so, how much?

FIGHT CLUB


IMDb
8.9/10
1,294,602


TOMATOMETER
CERTIFIED
FRESH
Rotten Tomatoes
79%
Average Rating: 7.3/10
Reviews Counted: 162
Fresh: 128
Rotten: 34

$IMDB - TomatoMeter$

$8.9 - 7.3$

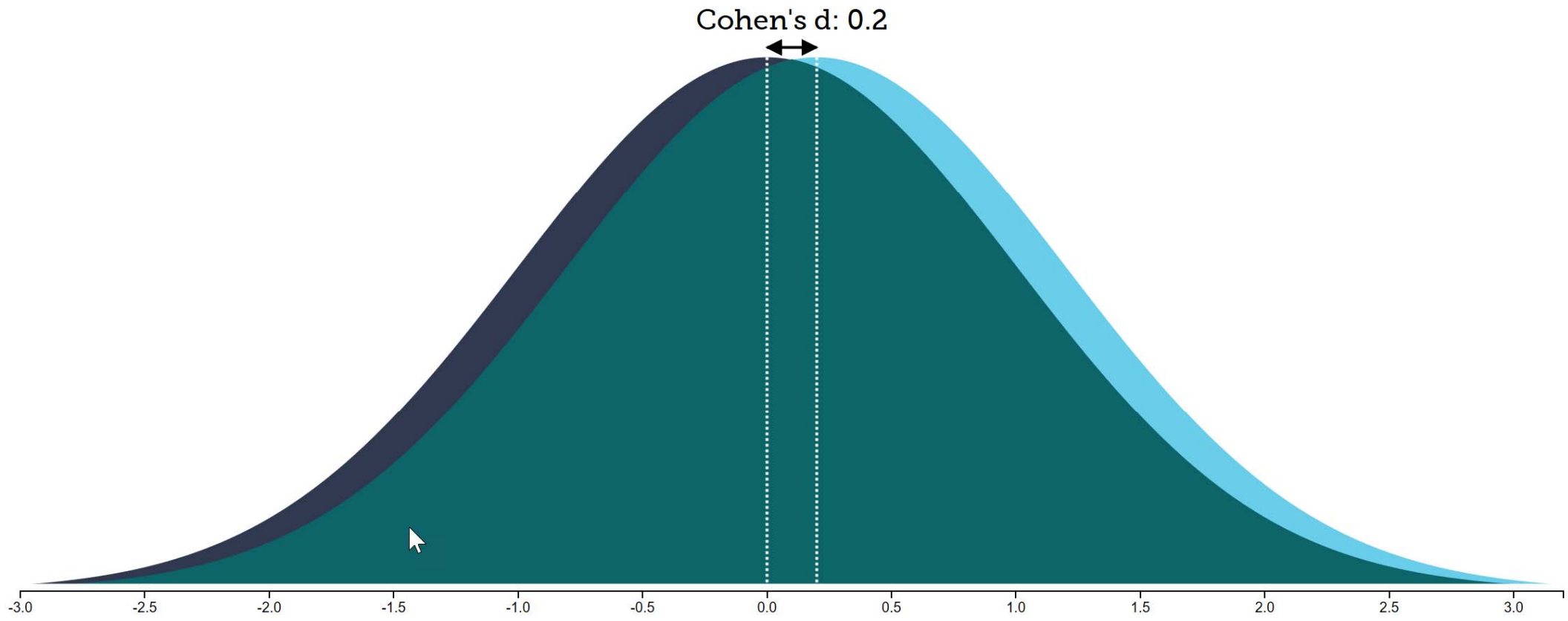
$$\frac{IMDB - TomatoMeter}{Standard\ Deviation}$$

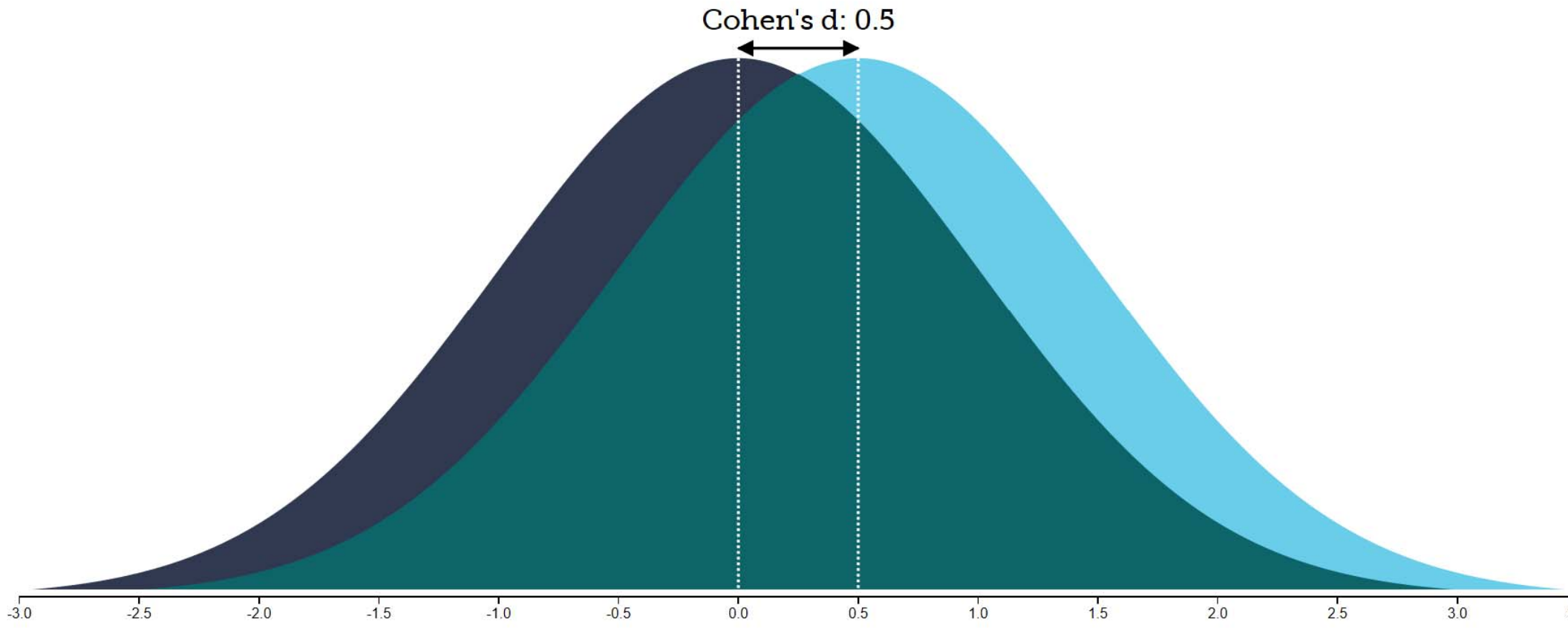
$$\frac{8.9 - 7.3}{1.4}$$

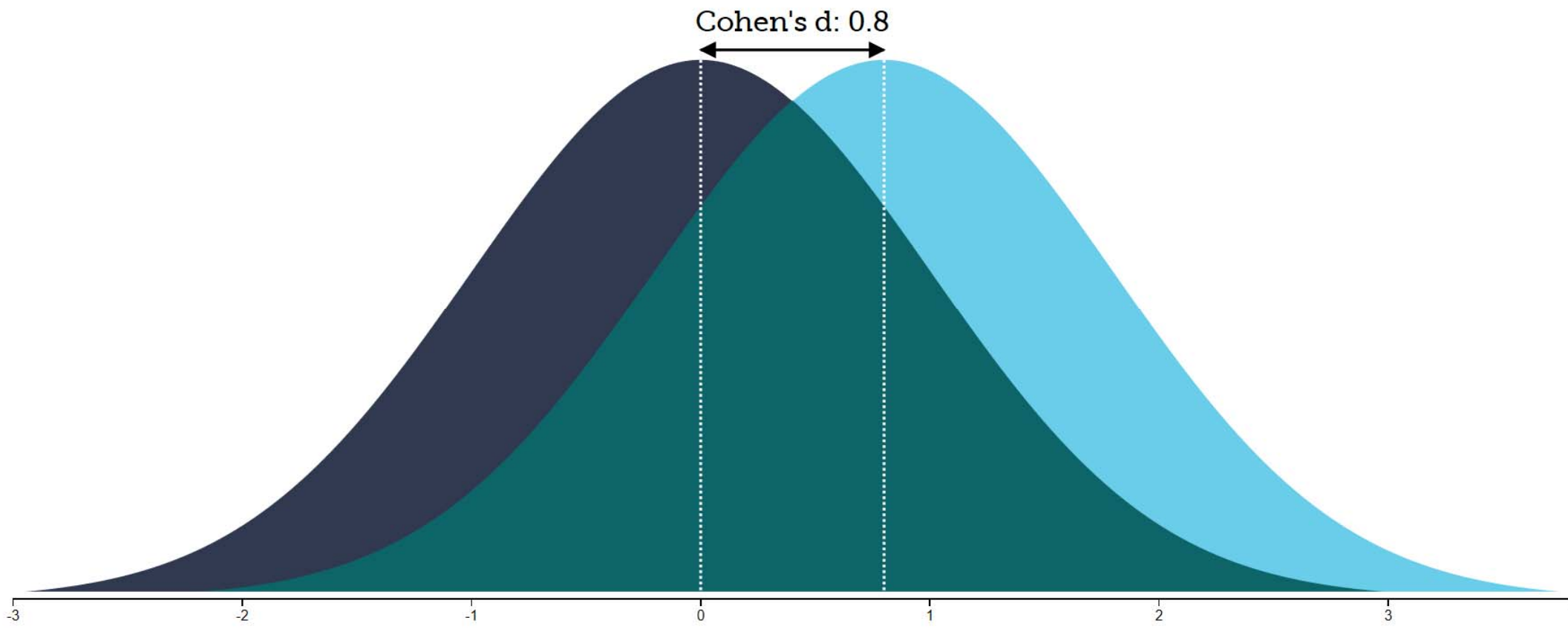
$$\frac{8.9 - 7.3}{1.4}$$

Cohen's $d = 1.14$

Cohen's *d* ranges from 0 (no effect) to ∞


Cohen's d: 0.2
-3.0
-2.5
-2.0
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
2.0
2.5
3.0


Cohen's d: 0.5
-3.0
-2.5
-2.0
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
2.0
2.5
3.0

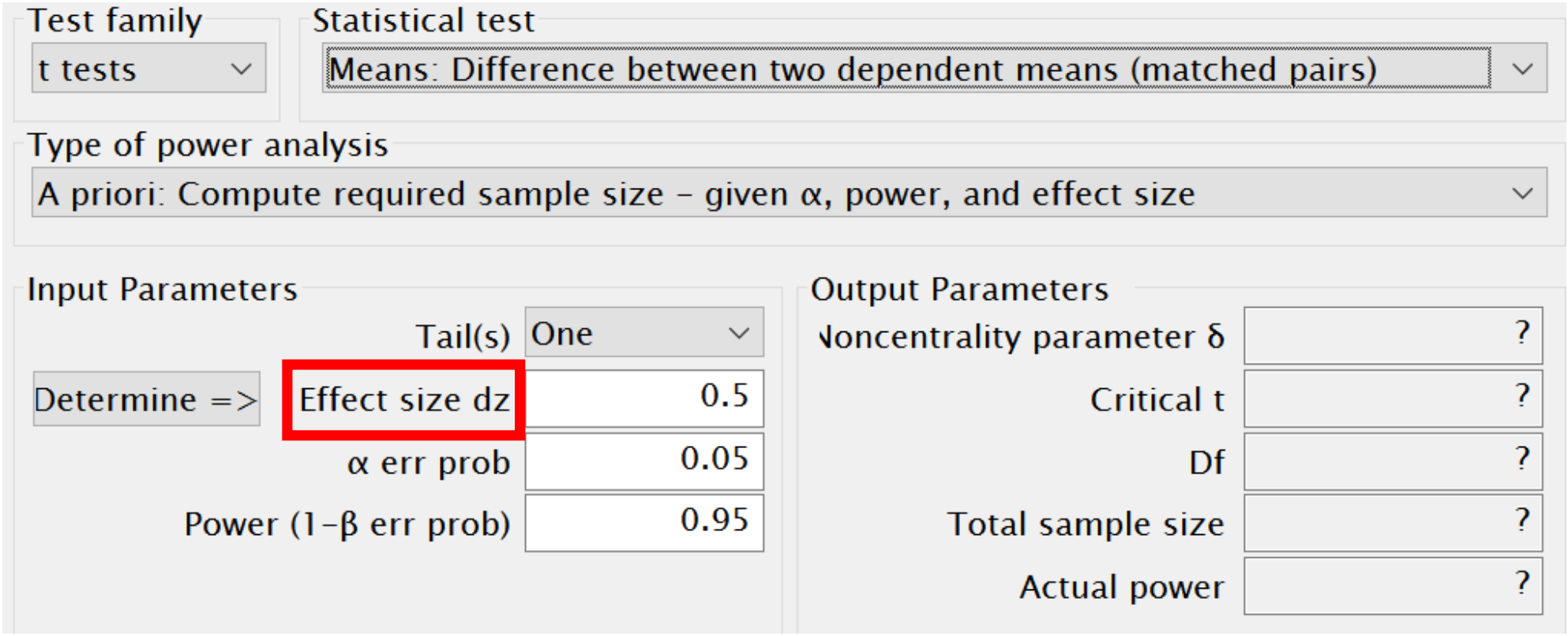

Cohen's d: 0.8
-3
-2
-1
0
1
2
3

Use benchmarks ‘large’, ‘medium’, or ‘small’ only as a last resort.

Cohen's *d* is the **difference** divided by the **standard deviation**

But slightly different versions exist: d_z (within) d_s (sample) d_{av} (corrected within)

Lakens, 2013


Test family
t tests
Statistical test
Means: Difference between two dependent means (matched pairs)
Type of power analysis
A priori: Compute required sample size – given α, power, and effect size
Input Parameters
Tail(s)
One
Determine =>
Effect size dz
0.5
α err prob
0.05
Power (1–β err prob)
0.95
Output Parameters
Noncentrality parameter δ
?
Critical t
?
Df
?
Total sample size
?
Actual power
?

Hedges' g is an unbiased version of Cohen's d.

$$g = d \times \left(1 - \frac{3}{4(n_1+n_2)-9}\right)$$

d can be calculated from *t* and *n*

$$d_s = t \times \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

Cohen's d for a within and between design (same *M* and *SD*) differ by **$\sqrt{2(1 - r)}$**

When $r > 0.5$,
$\sqrt{2}(1 - r) > 1$,
d is higher,
power is higher.

When needed, **Cohen's *d*** provides a standardized mean difference effect size.