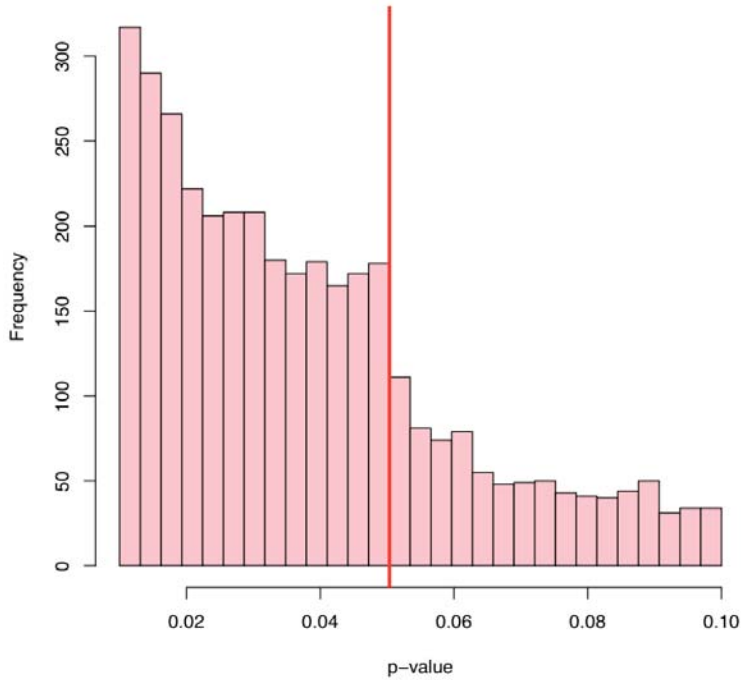
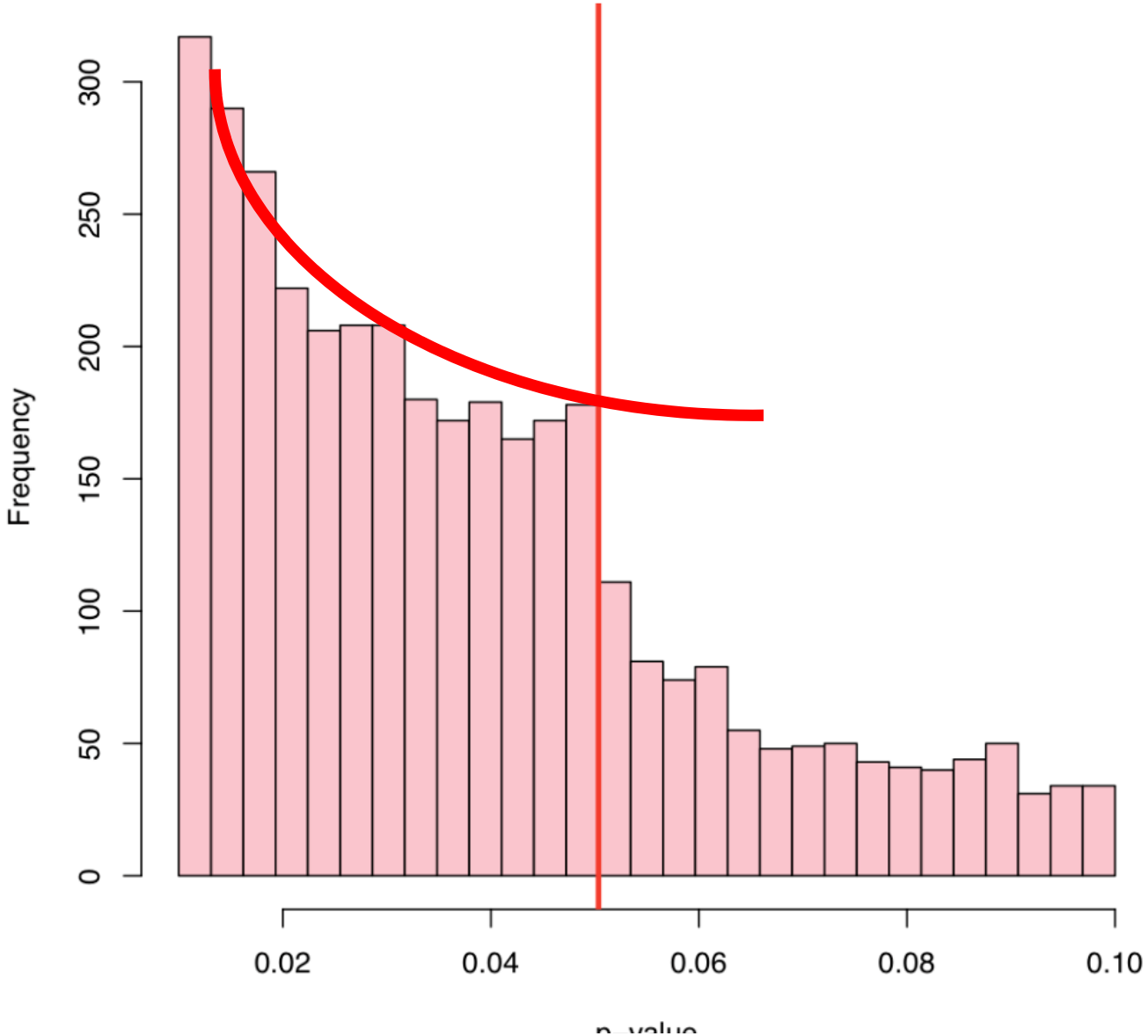
Publication Bias

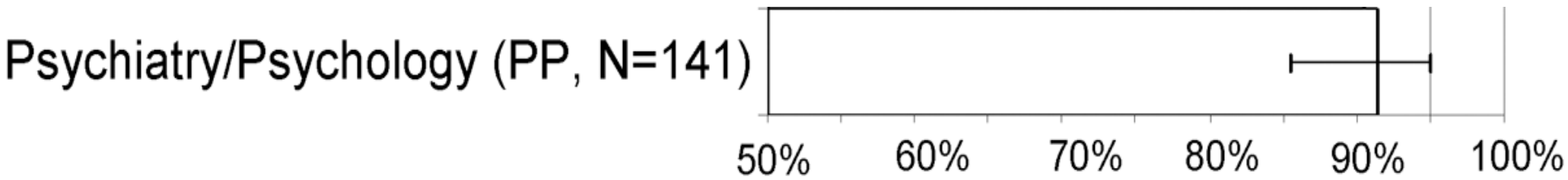


One of the bigger problems of *p*-values is their use as a threshold to publish.


Frequency
0
50
100
150
200
250
300
0.02
0.04
0.06
0.08
0.10

And this is where we put the non-significant results.

Most published findings confirm the hypothesis (Fanelli, 2010)

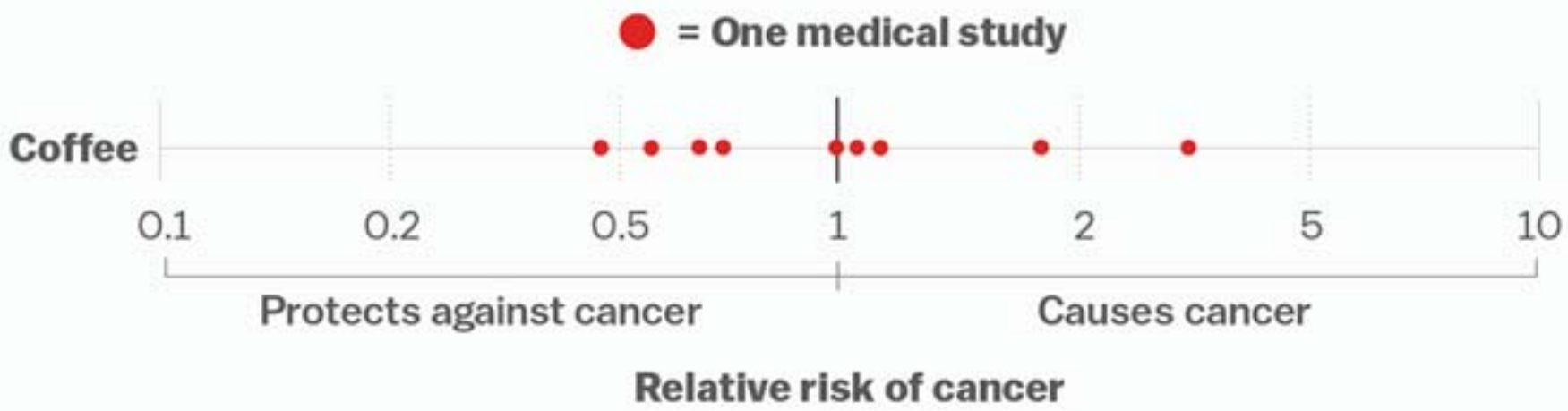


Null-results are difficult to interpret.

-There is no effect

-The study wasn’t good

Everything we eat both causes and prevents cancer



SOURCE: Schoenfeld and Ioannidis, *American Journal of Clinical Nutrition*

Vox

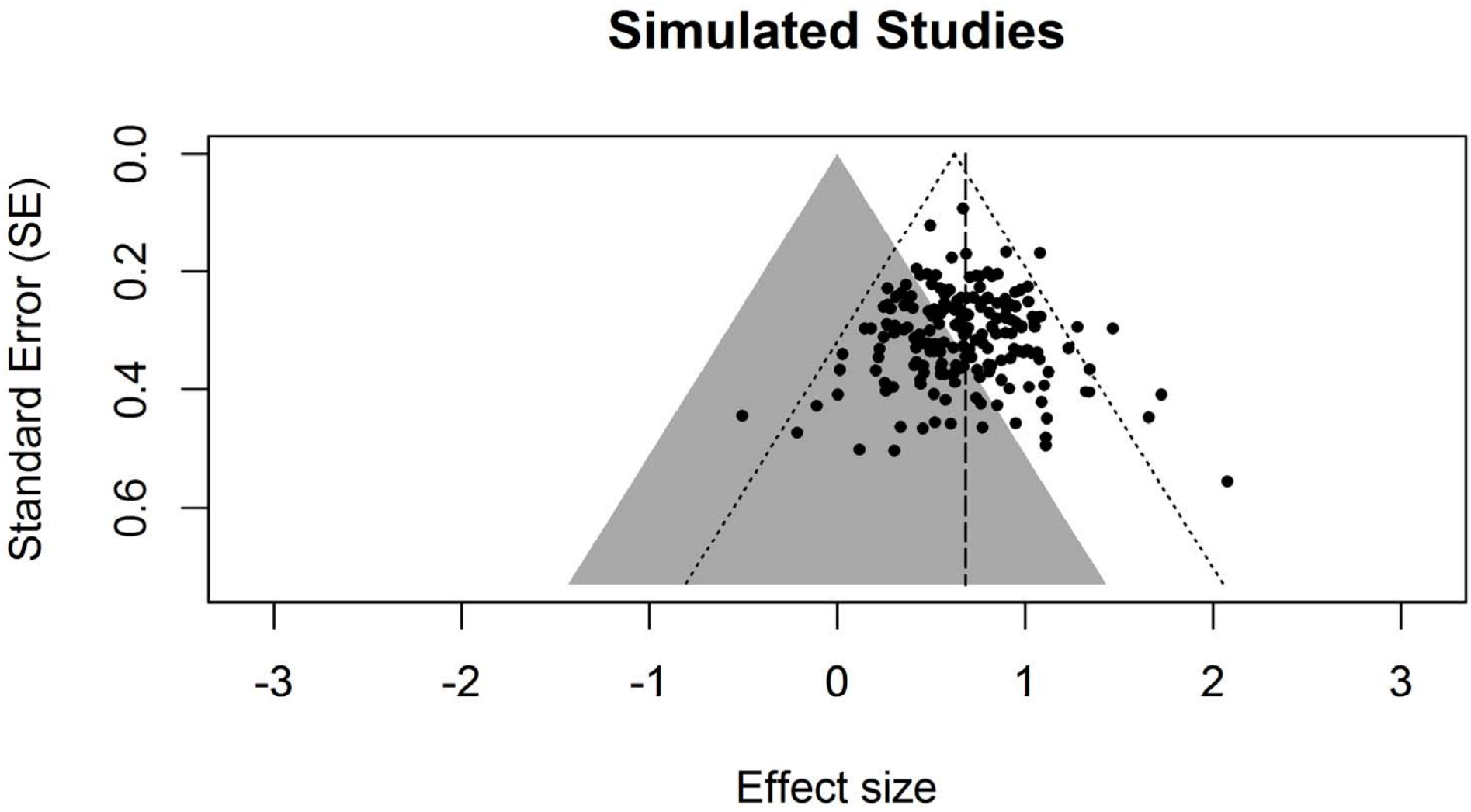
As long as a research area doesn’t share all results, it’s not a quantitative science.

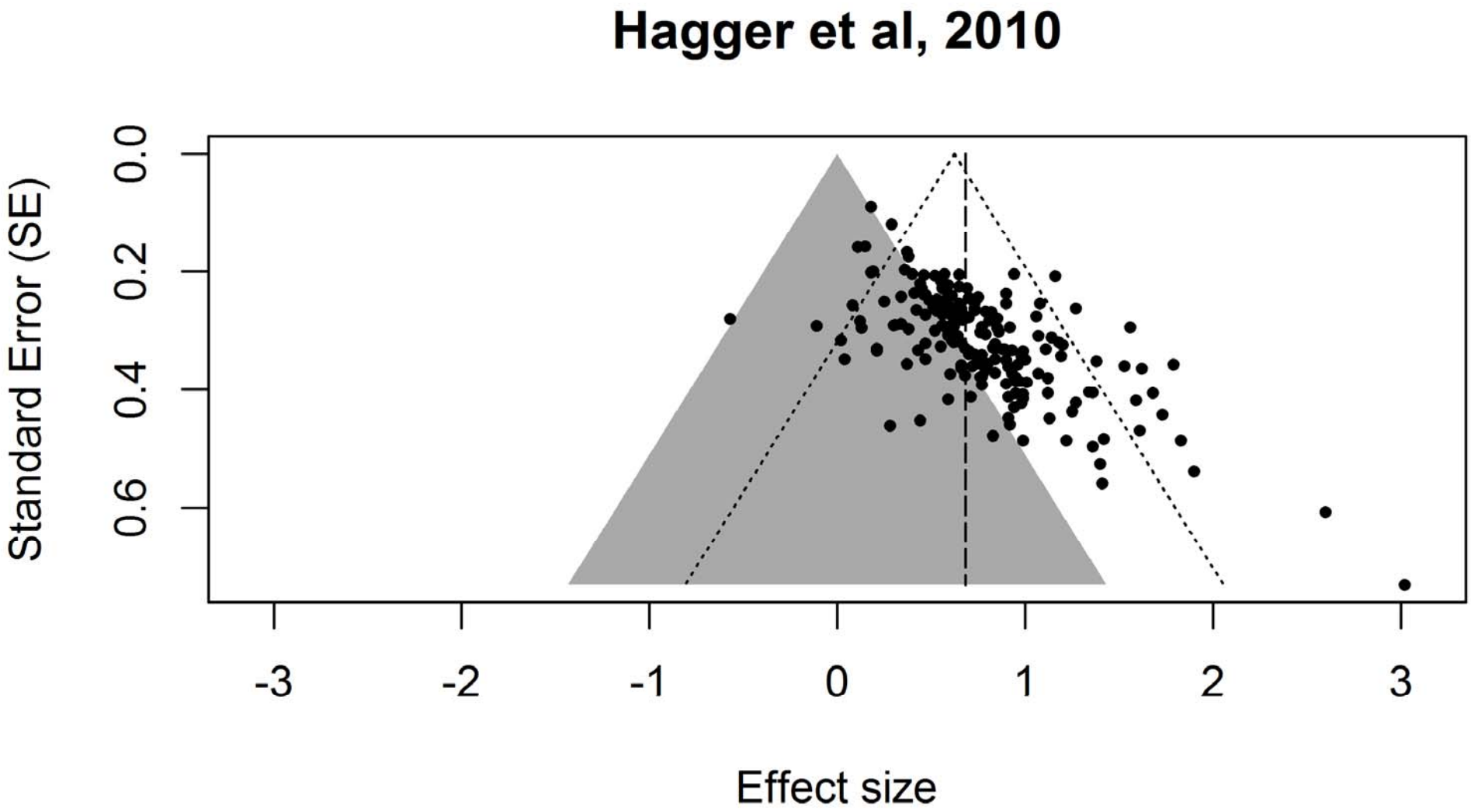
Promoting excellence in
parapsychological research and education

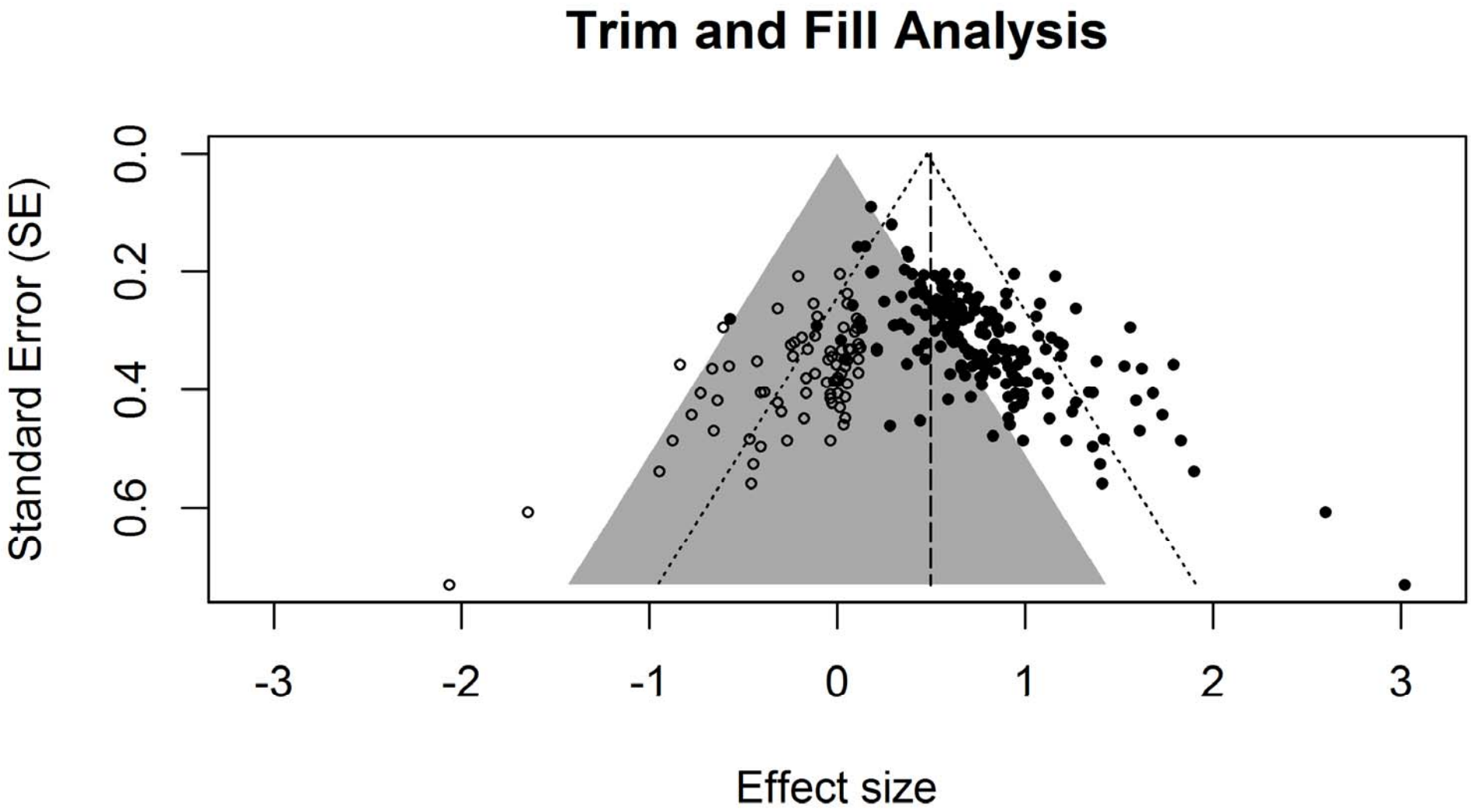
Koestler Parapsychology Unit

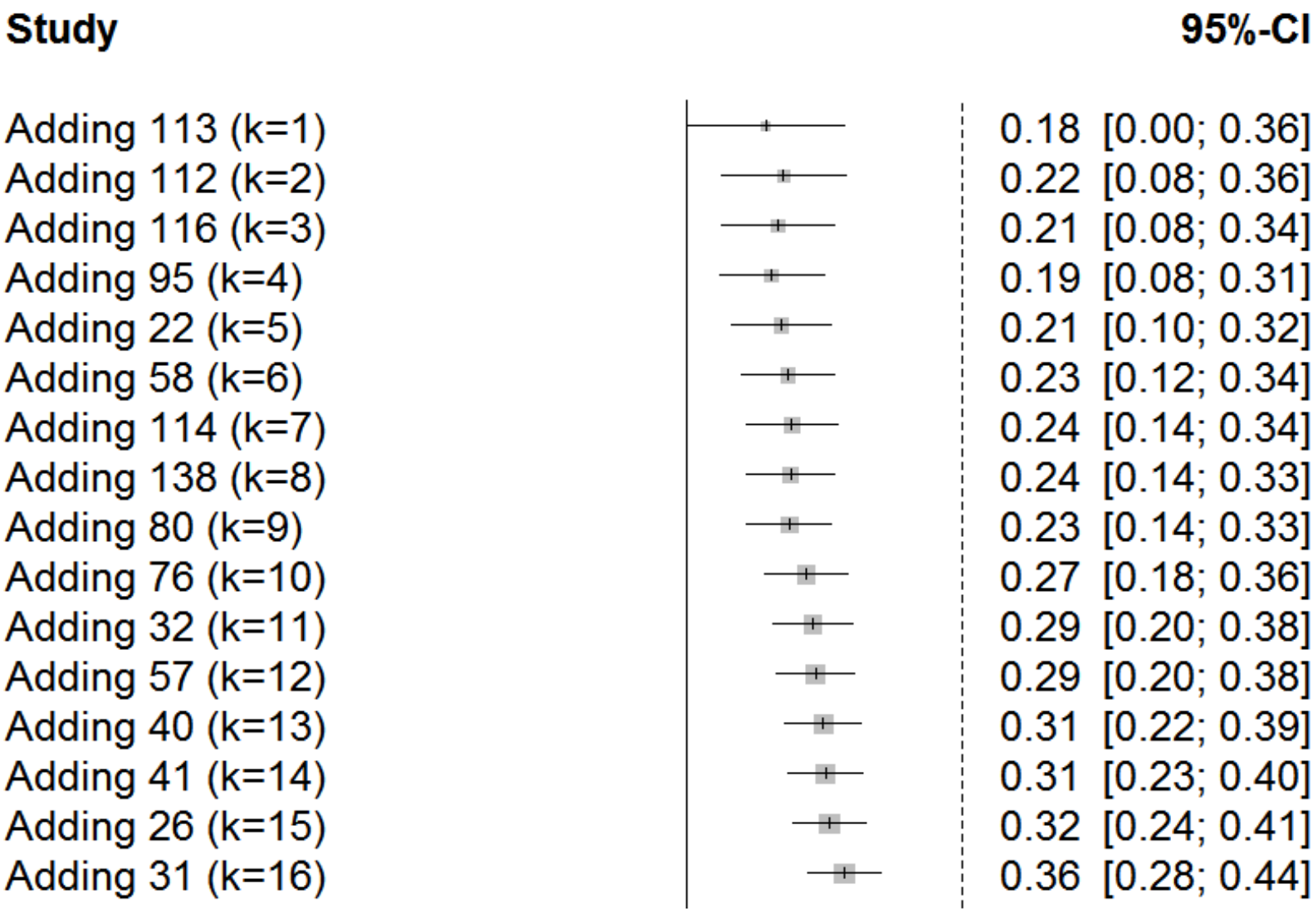
There can be 200 published studies with $p < 0.05$, but no true effect.

Publication bias can not be **corrected**, but it can be **detected**.


Simulated Studies
Standard Error (SE)
0.0
0.2
0.4
0.6
-3
-2
-1
0
1
2
3
Effect size


Hagger et al, 2010
Standard Error (SE)
0.0
0.2
0.4
0.6
-3
-2
-1
0
1
2
3
Effect size


Trim and Fill Analysis
Standard Error (SE)
0.0
0.2
0.4
0.6
-3
-2
-1
0
1
2
3
Effect size


Study
95%-CI
Adding 113 (k=1) 0.18 [0.00; 0.36]
Adding 112 (k=2) 0.22 [0.08; 0.36]
Adding 116 (k=3) 0.21 [0.08; 0.34]
Adding 95 (k=4) 0.19 [0.08; 0.31]
Adding 22 (k=5) 0.21 [0.10; 0.32]
Adding 58 (k=6) 0.23 [0.12; 0.34]
Adding 114 (k=7) 0.24 [0.14; 0.34]
Adding 138 (k=8) 0.24 [0.14; 0.33]
Adding 80 (k=9) 0.23 [0.14; 0.33]
Adding 76 (k=10) 0.27 [0.18; 0.36]
Adding 32 (k=11) 0.29 [0.20; 0.38]
Adding 57 (k=12) 0.29 [0.20; 0.38]
Adding 40 (k=13) 0.31 [0.22; 0.39]
Adding 41 (k=14) 0.31 [0.23; 0.40]
Adding 26 (k=15) 0.32 [0.24; 0.41]
Adding 31 (k=16) 0.36 [0.28; 0.44]

Failsafe N – how many studies reduce an effect to zero? **Don’t use it**.

Meta-regression techniques might be useful (e.g., Egger's regression)