Can people accurately draw statistical conclusions from dot plots?

Sara Jaramillo   Benjamin Margolin Rottman

University of Pittsburgh

Introduction

- Dot plots are increasingly being used to communicate scientific findings.

- People give weaker judgments for dot plots than plots that show a mean + 95 CI [1]. But this does not assess bias.

- We used Common Language Effect Size (CLES) and Bayes Factors (BFs) as a method of measuring bias.
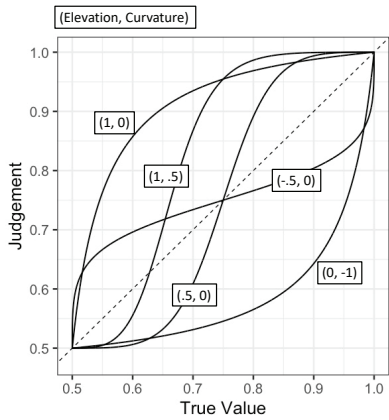
Design

- N = 227 for Ex1, 367 for Ex2

- 7 Effect Sizes (Cohen's d): 0.025, 0.1, 0.2, 0.4, 0.8, 1.6, 3.2

Effect Size Conversion

Cohen's d	0.025	0.1	0.2	0.4	0.8	1.6	3.2
CLES	.51	.53	.56	.61	.71	.87	.99

- 5 Sample Sizes: 16, 32, 64, 128, 256

- 4 Graph Styles

- DVs (between ss): CLES vs. DBF

Gonzalez & Wu's probability weighting function [2]
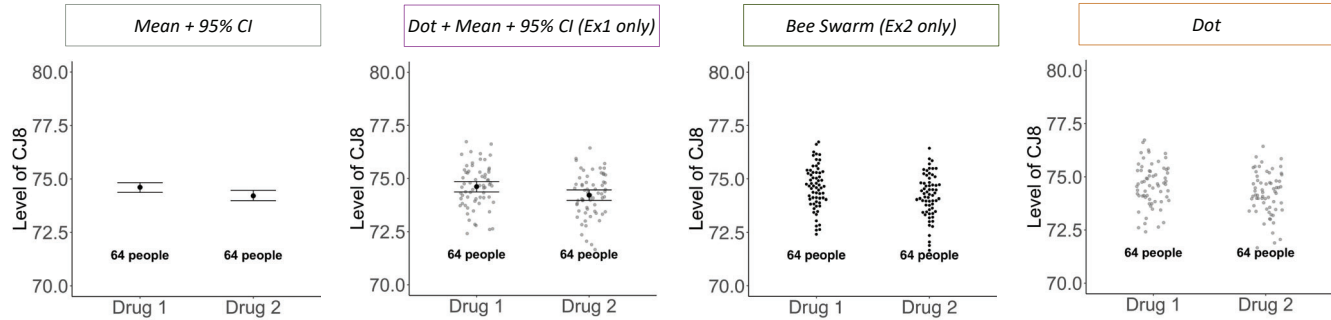

Elevation
  E > 0 judgments too strong
  E < 0 judgments too weak

Curvature
  C > 0 underestimate small effect sizes or BFs and overestimate large ones
  C < 0 overestimate small effect sizes or BFs and underestimate large ones

Discussion

- Effect sizes and BFs are underestimated for Dot plots and Bee swarm plots compared to Mean + CI.

- Higher variance for Mean + CI plots.

- People interpret error bars as deterministic, failing to incorporate uncertainty into judgements [3].

- We showed 95 CI's but there are also SD, SEs, and different thresholds for CI's (e.g., 80).

Contact: saj107@pitt.edu

[1] Witt, J. K. (2019). Introducing hat graphs. *Cognitive research: principles and implications*, *4*(1), 1-17.
[2] Gonzalez, R., & Wu, G. (1999). On the shape of the probability weighting function. *Cognitive Psychology, 38*(1), 129–166.
[3] Belia, Sarah & Fidler, Fiona & Williams, Jennifer & Cumming, Geoff. (2005). Researchers Misunderstand Confidence Intervals and Standard Error Bars. *Psychological methods. 10*. 389-96.
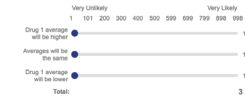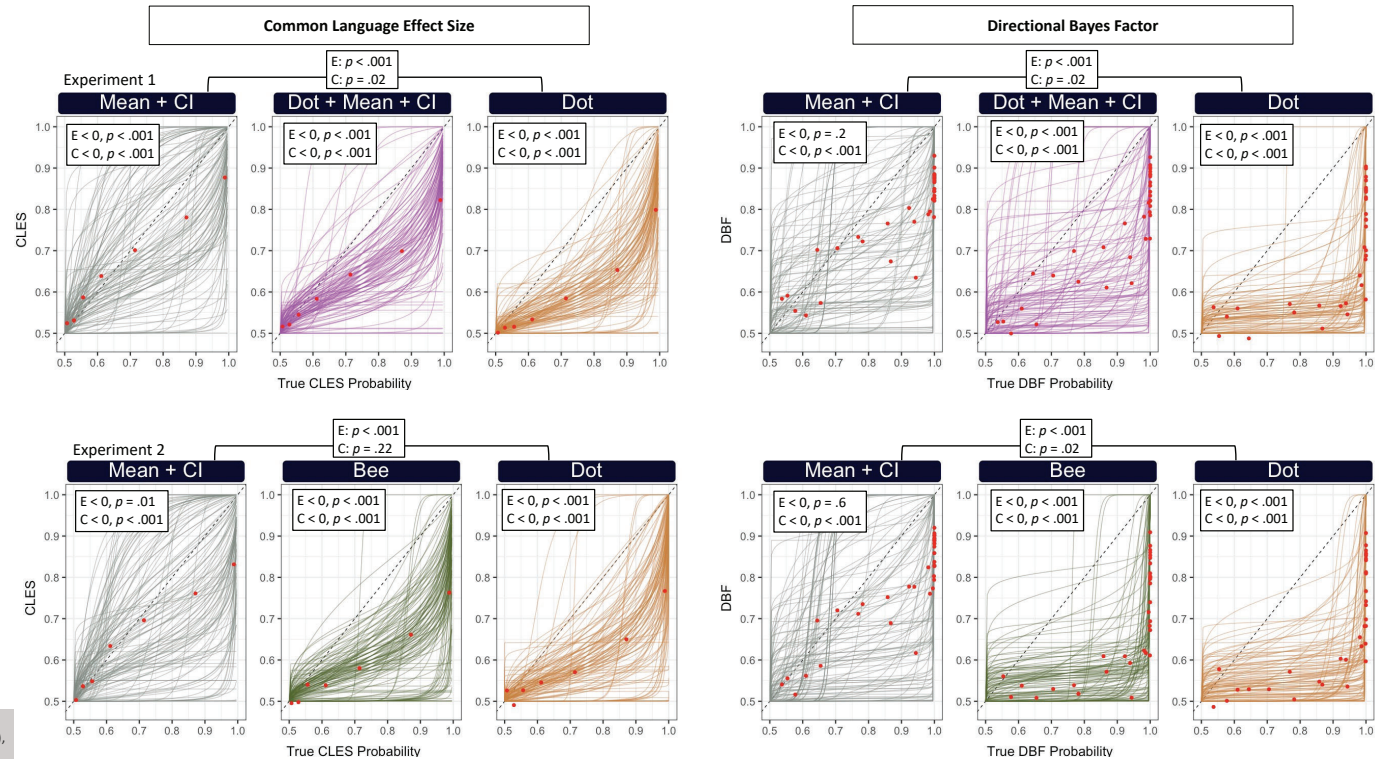
Stimuli

Common Language Effect Size

"What is the probability that the person from the Drug 1 group will have a higher level of CJ8 than the person from the Drug 2 group from this much larger study?"

Directional Bayes Factor

"What are the chances that the average levels of CJ8 of the Drug 1 group will be higher, the same, or lower than the average of the Drug 2 group in this much larger study?"

Results