Moore et al 2015, Overprecision

Kevin Dorst

Overprecision and overconfidence

Last time: point-estimates. How likely is it that *you got question i right*? Miscalibrated = of all the guesses you were x%-confident in, $y \neq x$ were true.

Or: How likely is it that you did better than a randomly-chosen person?

'Overconfidence' ≈ over-extremity.

Standard finding: hard/easy effect.

Interval estimates: What's your 90%-confidence interval for *the population of the UK*?

Miscalibrated = proportion of CIs containing the true value is $y \neq 90\%$.

Overprecision: hit rate < 90%
Underprecision: hit rate > 90%

Standard finding: overprecision.

90%-CIs contain the true value only 30–60% of the time.

The Puzzle

Interval estimation and point estimation are *inter-translatable*.
Usually[1], 90%-CI is [60,100] iff $P(X \geq 60) = 0.95$ & $P(X \leq 100) = 0.95$

[1] Given unimodality and centrality

But the two methods have very different empirical features:

- Interval estimates are more overprecise than intervals elicited with the *two-point method*[2].
- Interval widths are insensitive to confidence level: people give roughly the same intervals when asked for 90%- vs. 50%-CIs.
- People think fewer than 90% of *their own 90%-CIs* contain the truth.
- Overprecision in interval estimates is much more robust—no analogue of 'easy' side of hard-easy effect.

[2] 'What's the highest number that you're 95%-sure is below the true population of the UK? The lowest number that you're 95%-sure is above it?'

Give 10 CIs; asked what proportion contain the true value; avg answer: *6*

Even when elicited in different ways, like SPIES

Why?

What might explain this?