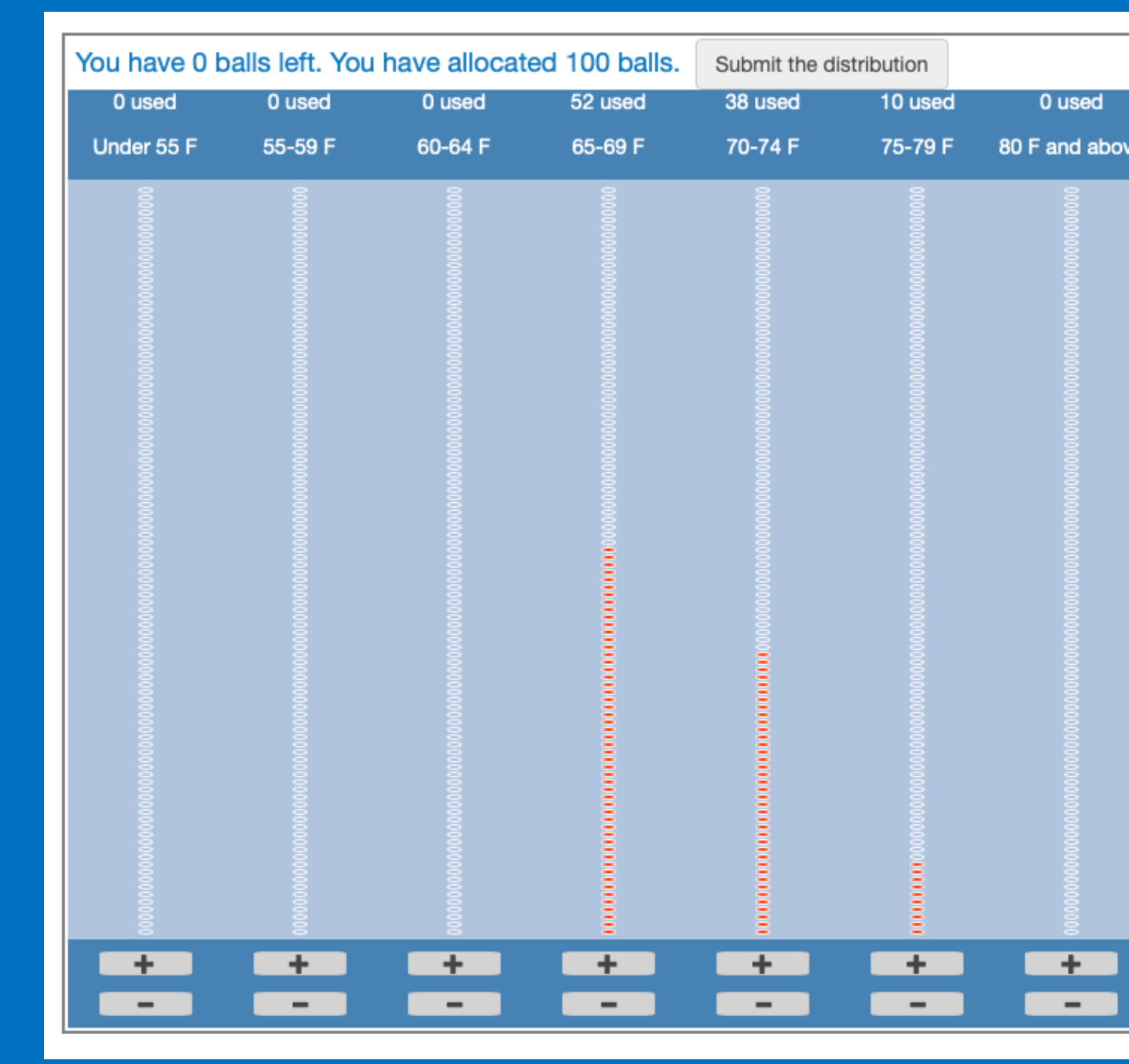
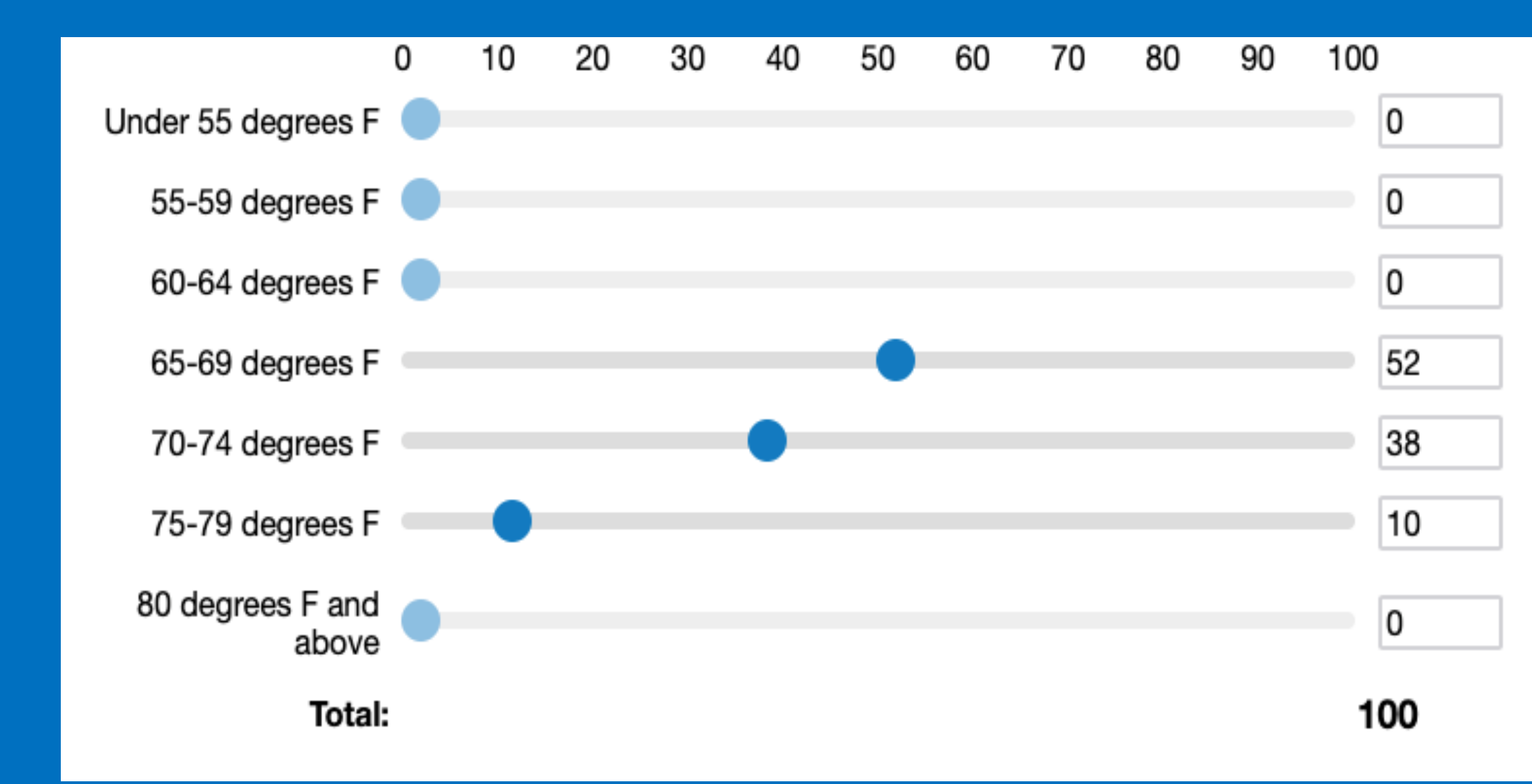


Researchers have recently recognized the benefits of measuring people's entire belief distributions.

- This approach divides the entire space of possible outcomes into mutually exclusive and exhaustive intervals, to which participants allocate subjective probabilities that add up to 100%.
- Two commonly used variants are the **Distribution Builder** [1] and the slider scales (or the "SPIES" method) [2].



Distribution Builder



SPIES

Research Question:
Do distributions elicited through these two methods differ and, specifically, does one method lead to more accurate results?

Overview of Studies

We ran 6 pre-registered studies where we asked participants to recall distributions using either the Distribution Builder or the Sliders.

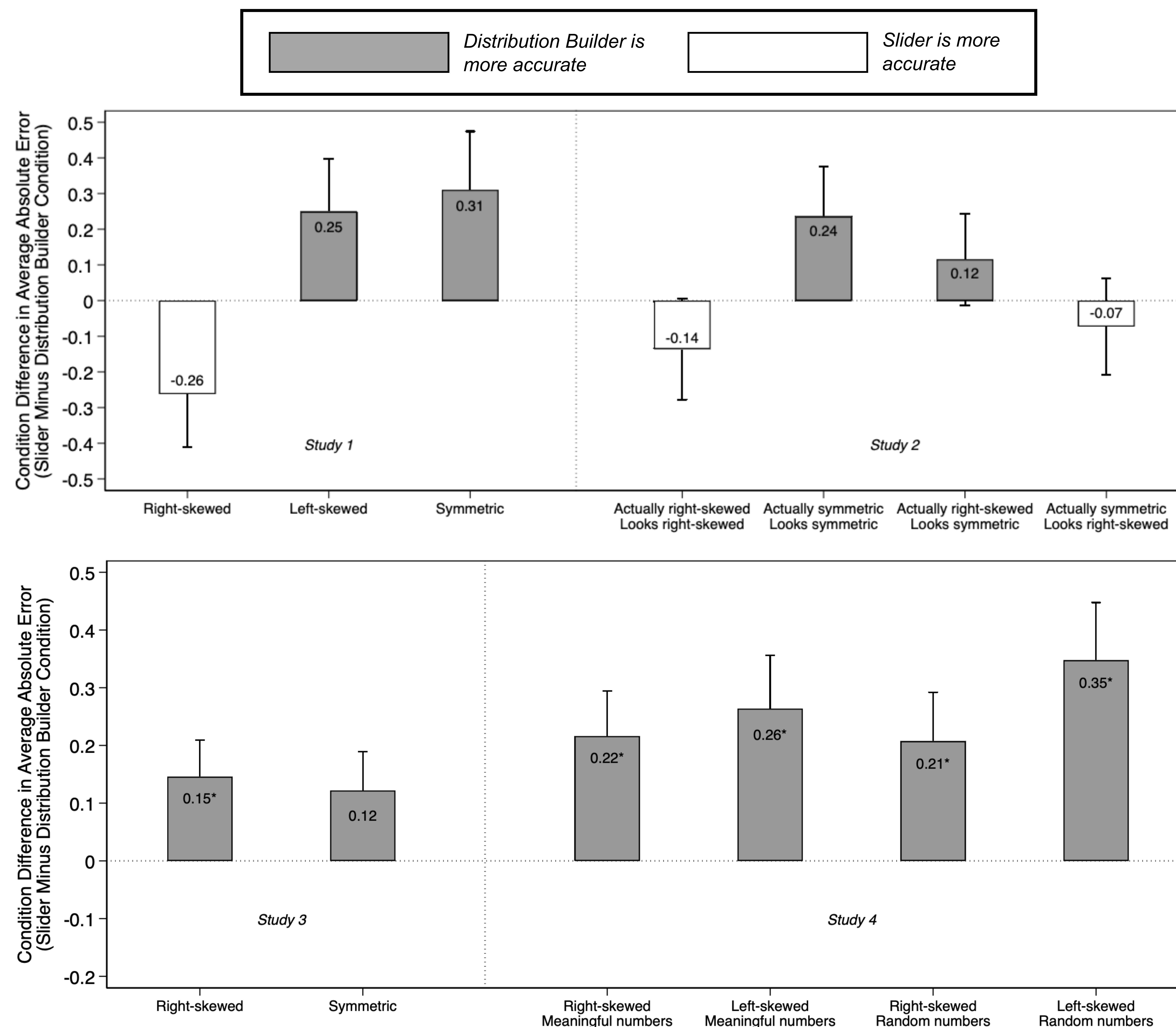
Task type:

- In 5 studies (Studies 1-4, 6), we asked them to recall the distribution of numbers we showed them previously (DV: average absolute error of individual responses).
- In Study 5, we asked them to provide estimates for general knowledge questions (DV: the percentage allocated to the category containing the true answer).

Study	Stimuli	Additional IVs manipulated			Sample
		Shape	Partition	Number perception	
1	50 cars' MPG values	Right-skewed; Left-skewed; Symmetric			1,276
2	50 cars' MPG values	Right-skewed; Symmetric	Even partition; Uneven partition		1,573
3	20 U.S. States' COVID testing rates	Right-skewed; Symmetric			1,536
4	50 heart rates; 50 mile times	Right-skewed; Left-skewed		Random numbers; Meaningful numbers	1,905
5	4 general knowledge questions	Right-skewed; Symmetric			1,576
6	50 song's danceability scores	Right-skewed; Left-skewed			1,787

Results

Studies 1-4: Condition Difference in Average Absolute Error
(Positive numbers = Distribution Builder is More Accurate Than Slider Condition)



The Distribution Builder elicited more accurate subjective distributions than the Sliders most of the time, except for a few occasions when the distribution should have a right-skewed shape.

Results are robust to:

- Different types of recall tasks (Study 3; Study 5)
- Different accuracy measures (Study 5)
- Different number perceptions (Study 4)

What could have led to different results produced by the two methods?

- We tracked the order by which each Distribution Builder/Slider category was first activated and found across all studies that **significantly more Slider users started from the first category than Distribution Builder users**.
 - Sliders led to more right-skewness** in that (1) more mass was allocated to the first bucket in the Slider condition and (2) more participants in the Slider condition allocated more to the first half than the second half.
- One reason for Sliders' disadvantage could be that Slider users tend to start with the lowest category in the interface and were more likely to put excessive mass in the lower categories.

Our research affords a practical recommendation to future research: All else equal, the distribution builder is the preferred method to elicit belief distributions.

The prior literature shows that discrete subjective estimates can be influenced by elicitation methods. Our research shows that the same is true of subjective belief distributions.



Selected References:

- [1] Distribution Builder: Goldstein, D. G., & Rothschild, D. (2014). Lay understanding of probability distributions. *Judgment & Decision Making*, 9(1)
- [2] SPIES: Haran, U., Moore, D. A., & Morewedge, C. K. (2010). A simple remedy for overprecision in judgment. *Judgment and Decision Making*, 5(7), 467.

Acknowledgments: We thank the Wharton Behavioral Lab, the Wharton Dean's Research Fund, and the Wharton Risk Center Russel Ackoff Doctoral Student Fellowship for support of this research.

Contact: Beidi Hu, beidihu@wharton.upenn.edu