Distribution Builder Vs. Slider Scales: What Is the Best Way to Elicit Belief Distributions?

Beidi Hu, University of Pennsylvania
Joseph Simmons, University of Pennsylvania

We compare two popular methods used to measure consumers’ belief distributions: Distribution Builder vs. SPIES (using slider scales). We find that Distribution Builder usually elicits more accurate belief distributions, at least partially because SPIES users tend to complete the task “in order,” causing them to exhibit predictable response biases.

[to cite]:

[url]:
http://www.acrwebsite.org/volumes/3000718/volumes/v49/NA-49

[copyright notice]:
This work is copyrighted by The Association for Consumer Research. For permission to copy or use this work in whole or in part, please contact the Copyright Clearance Center at http://www.copyright.com/.
EXTENDED ABSTRACT

How should researchers assess consumers’ beliefs? By far, the most common approach is to simply ask consumers to make a simple judgment or choice (e.g., “What do you think Company X’s stock price is going to be one year from today?”; “Which product is most likely to provide the best value?”). But such elicitation methods omit a great deal of information. Most notably, they fail to capture consumers’ beliefs about all possible options or outcomes (e.g., “How likely is Company X’s stock to go up 0-2%, 2-4%, 4-6%, etc.?“). To remedy this, researchers have recently embraced the practice of eliciting consumers’ entire subjective belief distributions. Our research compares two common ways of eliciting such belief distributions and indicates that one of those ways is probably better.

To elicit consumers’ subjective belief distributions, researchers typically provide a graphical interface that (1) divides the entire range of options into several mutually exclusive and collectively exhaustive categories, and (2) asks participants to estimate the frequency of each category (if the task asks them to do their best to reproduce an existing distribution) or the probability that the true outcome will fall into each category (if the task asks them how likely each outcome is to materialize).

Two variants of this approach have been most frequently used: the Subjective Probability Interval Estimates (SPIES) method proposed by Haran, Moore, and Morewedge (2010) and the Distribution Builder proposed by Goldstein and Rothschild (2014; see also Sharpe, Goldstein & Blythe, 2000; Goldstein, Johnson & Sharpe, 2008). Both methods allow participants to create a visual histogram that best represents their subjective distribution. The key difference between the two methods is the interface. The SPIES interface (Haran et al., 2010) contains an ordered array of horizontal slider scales, each representing one category. Participants assign probabilities to each category by sliding the bars from left to right. The Distribution Builder (Goldstein & Rothschild, 2014; Andre, 2016) contains a graphical interface that allows participants to allocate a fixed number of balls into vertically displayed bins, each representing one category.

Both methods have been adopted to study a wide range of consumer research topics, such as financial decision making (Sharpe et al., 2000; Goldstein et al., 2008; Long, Fernbach, & De Lange, 2018; Camilleri, Cam, & Hoffmann, 2019; Reinholdt, Fernbach, & De Lange, 2021), consumers’ perceptions of income distributions (Page & Goldstein, 2016), statistical intuitions (Hofman, Goldstein, & Hullman, 2020; Andre, Reinholdt & De Lange, 2021), and overconfidence (Ren & Croson, 2013; Moore, Carter, & Yang, 2015; Moore et al., 2017; Prims & Moore, 2017; Soll et al., 2019). There is currently no research investigating how these two methods might induce different response patterns or whether one of them might be superior.

In this paper, we report the results of seven pre-registered experiments (N = 10,815) that investigate whether SPIES vs. Distribution Builders yield different response patterns and, specifically, whether one method leads to more accurate distributions. We examined accuracy using three different paradigms. In the first paradigm, participants were instructed first to observe a set of numbers as frequencies and then to do their best to reproduce the distribution of the numbers they saw. We recorded the absolute deviation of individual responses from the normative answer across all categories to derive a measure of individual-level accuracy. In the second paradigm, we moved away from memory tasks and instead assessed the accuracy of participants’ subjective probability distributions for general knowledge questions. We operationalized accuracy as the probability allocated to the category containing the correct answer. Finally, we used an incentive-compatible design to examine which elicitation method more accurately reflects people’s true beliefs. Participants made predictions by providing their entire belief distribution. They also indicated confidence in their predictions using a self-report measure and decided how much to wager on their predictions. We derived participants’ confidence in their prediction implied by the belief distribution and computed the absolute difference between that and self-reported confidence.

In every study, we manipulated whether participants used SPIES or Distribution Builder to indicate their belief distributions, and in most studies, we also manipulated the shape of the correct distribution (i.e., right-skewed, symmetric, left-skewed).

Our findings reveal that Distribution Builder elicits more accurate belief distributions than SPIES whenever the true distribution is not right-skewed. This was directionally true in 10 of 10 head-to-head comparisons, and significantly so in 5 of them. The average effect size for those 10 comparisons was d = 0.16. When the true distribution was right-skewed, the findings are more mixed. SPIES directionally, but non-significantly, outperformed Distribution Builder in 4 of 8 comparisons (average effect size: d = 0.07). In the other 4 right-skewed comparisons, Distribution Builder yielded directionally superior results, of which 2 were significant (average effect size: d = 0.10).

These somewhat complex results can be at least partially explained by a simple mechanism. Compared to participants who use Distribution Builder, participants who use horizontal slider scales (i.e., SPIES) tend to (1) start with the first and lowest categories of outcomes, and (2) subsequently assign higher frequencies and probabilities to those low categories. When the true distributions are right-skewed, containing more mass in the lowest categories, this form of biased responding can sometimes yield more accurate distributions. But when the true distributions are not right-skewed, this form of biased responding proves harmful, and accuracy suffers relative to Distribution Builder. On balance, these results suggest that, all else equal, it is probably best to assess subjective belief distributions using Distribution Builders rather than horizontal slider scales.

Given the rising popularity of both the Distribution Builder and the SPIES, our research affords a practical recommendation for future consumer research. All else equal, the Distribution Builder is the preferred method to elicit belief distributions. Furthermore, whereas a large prior literature shows that discrete subjective estimates can be influenced by elicitation methods (Juslin, Wennerholm, & Olsson, 1999; Kluyver et al., 1999; Teigen & Jorgensen, 2005; Funke, 2016; Thomas & Kyung, 2019), our research shows that the same is true of subjective belief distributions.

REFERENCES

Retrieved from https://quentinandre.github.io/DistributionBuilder/


Funke, F. (2016). A web experiment showing negative effects of slider scales compared to visual analogue scales and radio button scales. *Social Science Computer Review, 34*(2), 244-254.


