

DISSERTATION DEFENSE

Subjective Probability Interval Estimates: A Simple and Effective Way to Reduce Overprecision in Judgment

Uriel Haran

Tuesday, November 22, 2011

1:30 pm

324 GSIA (West Wing)

Overprecision in judgment is the most robust type of overconfidence, and the one least susceptible to debiasing. It refers to people's excessive certainty in the accuracy of their estimates, predictions or beliefs. Research on overprecision finds that confidence intervals, estimated ranges that judges are confident will include the correct answer, tend to include the correct answer significantly less often than what their assigned confidence level would suggest. For example, 90% confidence intervals typically include the correct answer about 50% of the time (Klayman, Soll, González-Vallejo, & Barlas, 1999). By this standard, confidence intervals appear too narrow, or overprecise. This dissertation focuses on effectively reducing this bias.

In this dissertation, I present a novel elicitation method which can reduce overprecision, sometimes eliminating the bias. This method, called Subjective Probability Interval Estimates, or, in short, SPIES, presents the judge with the entire range of possible values, divided into intervals. The judge estimates, for each interval, the probability that it includes the correct answer. Since these intervals include the entire range of possible values, the sum of these subjective probabilities is constrained to equal exactly 100%.

This work presents six experiments, organized in two parts. Part I focuses on the use of SPIES for eliciting quantitative estimates, and tests it against other elicitation methods in three experiments. Experiment 1 included a within-subject comparison of SPIES and two other elicitation methods, namely 90% confidence intervals and 5th and 95th fractile estimates, and found that SPIES produce interval estimates with significantly higher hit-rates than the other two methods. Experiment 2 varied the range which the SPIES task spanned and the number of intervals included in it, and found that

SPIES outperformed the confidence interval method across all configurations. Experiment 3 tested the robustness of this effect to different value scales, and to variations in the extremity of true values on the range. SPIES again produced consistently more inclusive and better calibrated estimates than confidence intervals.

In Part II, I tested whether SPIES can improve estimates in other elicitation formats. Participants made multiple estimates, using SPIES for some and confidence intervals for others. Participants in Experiment 4 produced confidence intervals with better calibration with their assigned confidence level after having used SPIES in a prior estimate than before having practiced with SPIES. This effect held even when the two estimates had no shared content, suggesting that SPIES influence the estimation process, rather than merely increase the amount of relevant information already present in memory when making the second estimate.

Experiment 5 tested the effect of SPIES on subsequent confidence intervals in two types of estimates. When participants could retrieve a relatively homogeneous set of values but were asked to estimate likelihoods of values across a wide range of possible outcomes, they responded by improving the inclusiveness and calibration of their subsequent confidence intervals. However, when the value set in the first estimate was diverse, such that retrieving evidence for the entire range of the SPIES was easy, no effect was observed in subsequent estimates. This suggests that judges do not simply generalize the SPIES process to subsequent confidence intervals. Rather, they might react to the conflict between their knowledge and the estimates they had to make. This conflict may increase doubt, leading to an adjustment in subsequent estimates to account for this uncertainty.

Experiment 6 manipulated the existence of a conflict between participants' knowledge of the distribution of possible values and the structure of the SPIES task, by varying the value set's exposure time. When exposure time was very long, participants could assign each interval in the SPIES task its according likelihood without conflict, and when it was very short, participants' knowledge of the value set came mainly from the SPIES task itself. SPIES did not improve subsequent confidence intervals in either of these two conditions. Rather, only when exposure time was moderate, as it was in Experiment 5, did SPIES result in improved calibration of subsequent confidence intervals.

Together, results of all experiments show that SPIES is an effective method for reducing overprecision in judgment. It allows for the elicitation of more inclusive and better calibrated estimates than those produced by the confidence interval method for a wide variety of estimate types. In addition, it can enact changes in judges' estimation process, such that their subsequent estimates, elicited by traditional methods, display better accuracy. These features make SPIES an effective tool to reduce one of Judgment and Decision Making's most robust biases.