

Visual Simulation and Subjective Probability Estimation: When Seeing Is Believing

Participant-visible method context. Results, discussion, captions, and the target figure are excluded.

Participants The target sample size is 100 per treatment group to ensure sufficient statistical power ($>.90$) to detect a medium-to-large-sized effect. Therefore, the planned sample size is 200 participants. In total, I recruited 200 American Mechanical Turk³ participants (age: $M = 40.5$, $SD = 11.8$) with an incentive of \$1, where 82 participants identified as Female. **Procedure** Before starting the experiment, participants read an attention-check question.⁴ Participants are then introduced to a ball drop machine known as a quincunx. In this experiment, the quincunx involves a ball bouncing down 10 rows of pegs, where it is equally likely to fall to the left or right at each juncture, eventually landing in one of 11 buckets (labeled A to K). The quincunx machine was also used in Studies 2–5 of Moore et al. (2015) and is presented in Figure 2.

The participants for Studies 1 and 2 were screened to have a human intelligence task approval rate greater than 98% and to have completed at least 1,000 approved human intelligence tasks.

The question asks “approximately how many cups of water do you drink each day” and is followed by a brief description about drinking water with the instructions “when answering the question below, you must select other and enter the phrase soda.”

VISUAL SIMULATION AND PROBABILITY ESTIMATION

Figure 2 The Image of the Quincunx Machine Presented to Participants

Note. See the online article for the color version of this figure.

Probability Measures I compared the tail and middle probabilities between the participant groups to test whether the visual group was more precise than the control group. For each experiment, participants estimate probabilities across $B + 1$ bins. Let P_i be the standardized probability of bin $i \in \{0, B\}$. (Note that the bins are relabeled to be numbered from 0 to B .) To tail probabilities, I considered $T_n = P \text{ Pmeasure } n-1 \text{ } \delta \text{ } i=0 \text{ } P_i + B \text{ } i=B-n+1 \text{ } P_i \text{ } B=2n$, where I varied n depending on the number of bins B . Similarly, to measure probabilities near the distribution's $PB=2+n-1 \text{ } P_i = \delta 1 + 2\delta n - \text{center}$, I used $M_n = i=B=2-n+1 \text{ } 1 \text{ } P_i$. Specifically, in Study 1, I compared $T_2 =$

$4 \text{ } \delta p_0 + p_1 + p_9 + p_{10} \text{ } B$ and $M_3 = 5 \text{ } \delta p_3 + p_4 + p_5 + p_6 + p_7 \text{ } B$. Generally, greater overprecision should decrease the probability of T and increase the probability of M .