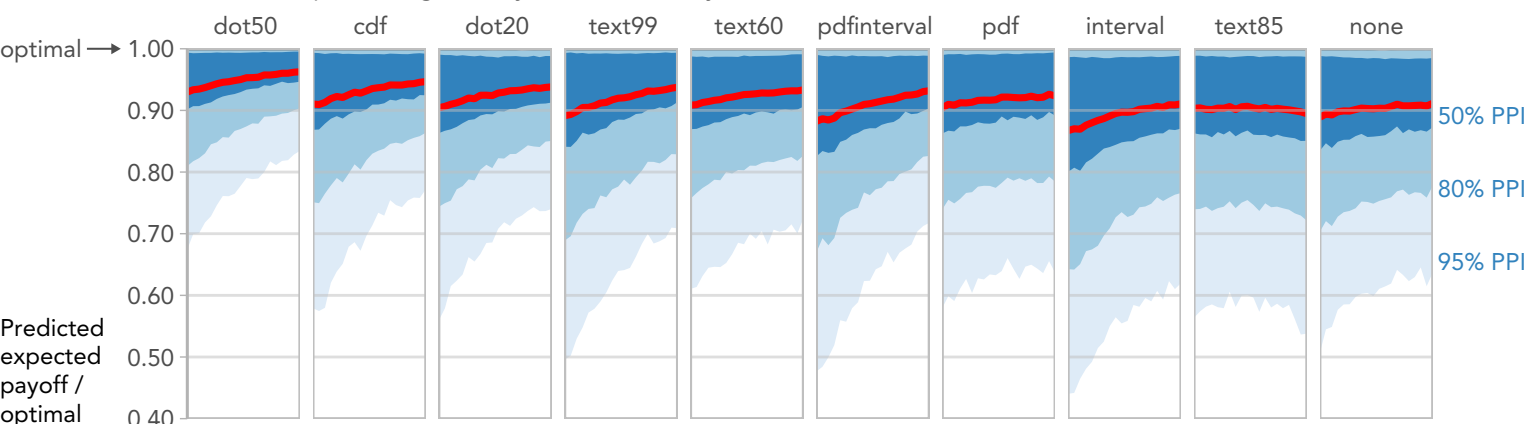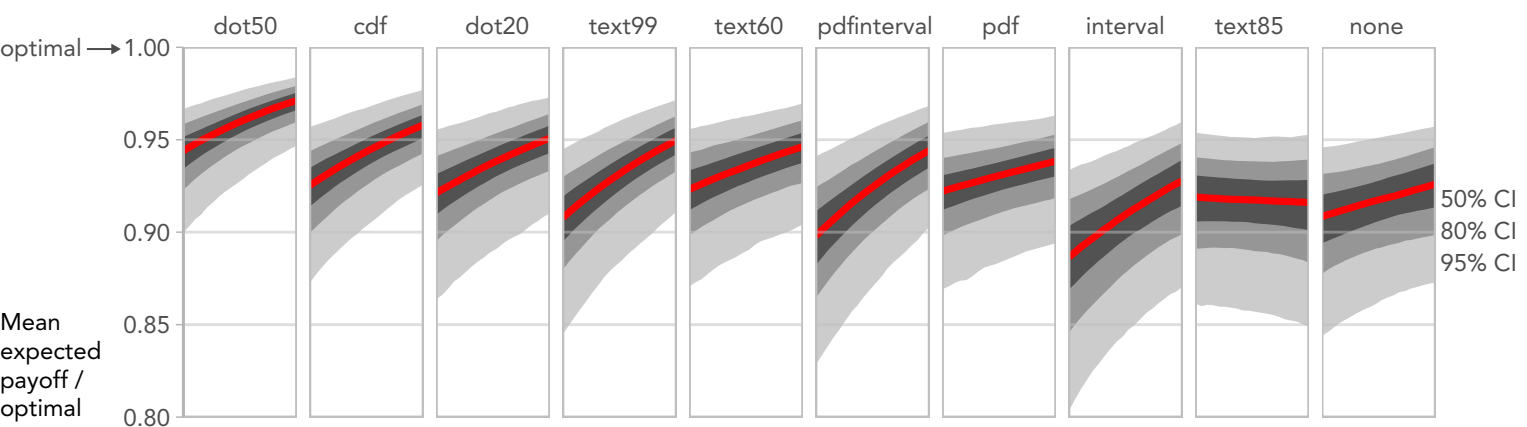1. **Posterior predictive intervals** and predicted **mean** for performance in each condition. These intervals are what we would predict 50%, 80%, or 95% of new observations of performance to fall into. Performance improves with additional trials, especially for dot50, cdf, and dot20. Meanwhile, performance in text depends on the risk threshold, with text85 performing similarly to no uncertainty.

2. Quantile **credible intervals** (Bayesian analog to confidence intervals) and posterior median of the **mean** performance in each condition. These intervals show the uncertainty in the location of the red line in chart #1, above.

3. Quantile **credible intervals** and posterior median of the **standard deviation** of performance in each condition. Not only does mean performance improve, but variance in performance also improves over time: people get better on average and more consistent; dot50, the best-performing condition, achieves an SD in performance likely less than 4 percentage points by the final trial.