Characteristic	Original protocol	ML2	ML5	Testable implication	Conclusion
Dimensions of scale measuring perceived likelihood	Likelihood scale contained 11 options, ranging from 0 to 10.	Likelihood scale contained 10 options, ranging from 1 to 10.	Likelihood scale contained 11 options, ranging from 0 to 10.	Questionnaire design effects may produce differences in results between the otherwise comparable U.S. Mturk samples.	Unlikely; no meaningful differences in Mturk samples.
Missing data	Unknown, but likely allowed subjects to skip questions due to the pencil-and-paper format.	The online questionnaire allowed subjects to skip questions and experiments.	The online questionnaire required subjects to answer all questions.	Questionnaire design effects may produce differences in results between the otherwise comparable U.S. Mturk samples.	Unlikely; no meaningful differences in Mturk results and proportion of missing data <1% in ML2.
Presence of unrelated experiments	Experiment was administered alone.	Experiment was administered as part of a 30-minute block of experiments, with order randomized.	Experiment was administered alone, or only after other experiments pre-approved as unlikely to influence results.	If other tasks interfered in ML2, results may differ by the order of presentation.	Unlikely; order effects appeared minimal and non-systematic in ML2.
Sampling frame for primary analyses	Undergraduates at Cornell University	Undergraduates	Undergraduates at "similar" sites	Comparing all ML2 sites to all ML5 sites may result in better agreement, since these sampling frames are more directly comparable.	Unlikely; results still discrepant.
Statistical analysis	One-way ANOVA (assuming homoskedasticity)	Independent-samples t-test combining sites' data (not assuming homoskedasticity)	Linear mixed model combining sites' data and using model-based SEs (assuming homoskedasticity)	Aggregating individual subject level data from ML2 and ML5 using a single analysis method may result in better agreement.	Unlikely; results still discrepant using LMM, GEE, and RMD analysis approaches.
Outlying sites	Outlying sites	None evident	A single site with a small sample size estimated a large, positive main effect.	N/A	Unlikely; the only possible outlier would have made results less discrepant.