

A Sufficient Statistic for Designing and Evaluating Human-AI Collaboration (Amendment to Pre-Analysis Plan)

Nikhil Agarwal, Alex Moehring, Alexander Wolitzky

August 20, 2024

As specified in the original pre-analysis plan, we pre-register this amendment that specifies the policies we will test in the second round of the experiment. In the second round, we will ask 2000 participants to assess claims under varying information environments. This design will include both a within and across comparison. This will allow us to take advantage of the within comparisons to control for heterogeneity in participant skill while preserving the ability to do pure across participant comparisons. Specifically, each participant will assess 8 claims in each of the following information environments¹

1. No disclosure: participants observe only the prior mean $E[\omega]$
2. Full disclosure: participants observe the AI assessment θ
3. Automation + no disclosure: we automate cases where $\theta < 0.4$ or $\theta > 0.67$ and disclose $E[\omega|\theta \in [0.4, 0.67]]$ among cases not automated
4. Optimal automation and disclosure: participants observe θ when $\theta \in [0.33, 0.69]$ and automate the cases where $\theta < 0.33$ and $\theta > 0.69$
5. Stoplight policy: The policy that maximizes performance where at most 3 values are revealed to participants. Participants observe $E[\omega|\theta \in [0, 0.39]]$ if $\theta \in [0, 0.39]$, $E[\omega|\theta \in [0.4, 0.67]]$ if $\theta \in [0.4, 0.67]$, and $E[\omega|\theta \in [0.68, 1]]$ if $\theta \in [0.68, 1]$

¹We originally planned to run policies 1-4 below and have participants assess 10 statements per arm. After round 1, we decided to include policy 5 and have participants assess 8 claims per arm. This decision was motivated to provide a better test of the sufficient statistic framework and given we expect the performance of this restricted policy to approximate full disclosure. Specifically, this policy allows us to test if the predicted performance equals actual performance for a more complicated disclosure policy compared to the optimal policy.