

## AI-biases influence on human decision-making: AI presentation order. (#103020)

Created: 07/21/2022 10:14 AM (PT)

Public: 06/22/2023 04:51 AM (PT)

### Author(s)

Lucía Vicente (University of Deusto) - lucia.vicente@deusto.es

Helena Matute (University of Deusto) - matute@deusto.es

### 1) Have any data been collected for this study already?

No, no data have been collected for this study yet.

### 2) What's the main question being asked or hypothesis being tested in this study?

In this study, we want to test whether the different order in which we present the phase with the AI-biased recommendations has an influence on participants' performance on a health-related classification task. We predict that the group of participants assisted by a biased recommendation system during the first part of the classification task will make more mistakes during this AI-assisted phase and during a second unassisted phase in comparison to the group that first completed the task without assistance. We also predict that the bias in the first AI-assisted group will generalize to ambiguous stimuli during the unassisted phase of the task so that these ambiguous stimuli become classified more often in the direction of the errors made by the AI during the assisted phase. Thus, the results of the group that completed first the AI-assisted phase should replicate the results of our previous experiment. The results of the group that completes first the unassisted phase of the task should extend those results. We expect this group to make fewer mistakes during the AI-assisted phase than the other group. We believe that prior experience classifying without assistance will protect this group against the AI-biased recommendations.

### 3) Describe the key dependent variable(s) specifying how they will be measured.

During the AI-assisted phase, the main dependent variable is the number of mistakes that participants make in all the trials of the task and particularly in the trials where the AI recommendation is erroneous, that is, where the recommendation contradicts the objective information (i.e. the proportion of dark/light cells) presented in a fictitious tissue sample (the bias of the AI consists on a systematic wrong recommendation on trials where the proportion of dark and light cells in the tissue sample is 40/60). During the unassisted phase of the experiment, the main dependent variables are also the number of errors in all trials of the classification task and particularly the number of errors in the classification of the tissue samples with the proportion 40/60. During this phase, we will also assess the number of trials in which the ambiguous (50/50) stimuli are classified in the same direction as the AI bias.

### 4) How many and which conditions will participants be assigned to?

The experiment comprises two groups as a function of the order in which they are exposed to the different phases of the experimental task. Half of the participants will be randomly assigned to the With AI – Without AI group, while the other half will be randomly assigned to the Without AI – With AI group.

### 5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.

Analyses to replicate previous findings and test the effect of order:

1. Between groups t-test on the number of errors in the 40/60 samples and on the total number of errors of the AI-assisted phase.
2. Between groups t-test on the number of errors in the 40/60 samples and on the total number of errors of the unassisted phase.
3. Between groups t-test on the number of biased classifications of the ambiguous stimuli 50/50 of the unassisted phase.

Also, we will test the changes in participants' behaviour across all the 40/60 trials of the task for each group with a mixed ANOVA with blocks of trials as a within-subject measure and group as a between-subjects measure.

### 6) Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.

Data from participants that do not get at least 5 correct responses (out of 6) on the second repetition of the practice phase will be excluded from the analysis. Also, data from participants that do not perform better than chance (more than 30 hits out of 60) in the classification task will be excluded from the analysis.

### 7) How many observations will be collected or what will determine sample size? No need to justify decision, but be precise about exactly how the number will be determined.

We planned a sample of 200 participants (100 per group). This sample size should allow us to detect a size effect of  $d = 0.35$  on the Student's t-test for the difference between two independent means with 80% power.

### 8) Anything else you would like to pre-register? (e.g., secondary analyses, variables collected for exploratory purposes, unusual analyses planned?)

We included some post-experimental questions to know if participants have seen and followed the recommendation of the AI and to explore the confidence participants have felt in their own ability to perform the task, and the confidence they place in general in artificial intelligence algorithms in health. This information is collected for exploratory purposes and thus we do not pre-register any prediction.