Artificial Intelligence and Human Creativity – Hypotheses and Analysis

Hypotheses:

H1: Overall scores for the observed matchstick designs are higher in the treatment condition than in the baseline.

H2: AI assistance will more strongly improve the overall scores for the observed matchstick designs among participants with low self-rated creativity than among participants with high self-rated creativity.

H3a: Among participants in the treatment condition, those who choose not to use AI assistance will create designs that are more original than those created by participants who use AI assistance with pre-written prompts.

H3b: Among participants using AI assistance in the treatment condition, those who generate their own prompts will create designs that are more original than those created by participants who select from pre-written prompts.

H4: Participants who decline AI assistance in the treatment condition will receive on average a higher overall score for building their designs than participants in the baseline condition.

H5: Participants using AI assistance in the treatment condition and who generate their own prompts will receive on average a higher overall score for building their designs than participants in the baseline condition.

Note: Manipulation checks indicated that our initial experimental design failed to create enough differentiation in the task performance and to design AI assistance conditions that potentially affect human creative outcomes. To address these issues, we redesigned the study with a more sophisticated task and distinct AI assistance conditions. The updated design enables us to test revised versions of Hypotheses 3a and 3b, as well as the newly formulated Hypothesis 5, which was not testable under the previous experimental framework.

For publication purposes, the language, abbreviations, and sequence of hypotheses may undergo minor adjustments to enhance clarity.

We use the overall score (consisting of the simple sum of the single scores of the dimensions originality and elaboration) assigned by the judgment panel to participants' completed designs as main variable of interest. For answering Hypotheses H3a and H3b we use the single score of the dimension of originality.

To answer our hypotheses, we will use non-parametric tests, such as Mann-Whitney U-Tests, Chi^2 Tests etc. In addition we will use linear regression analysis such as OLS regressions to corroborate the findings of the non-parametric results.