Practical Exam

Advanced Mathematics and Statistics for Data Science and AI

MSc Data Science and AI
University of the Arts London

3 December 2025

Introduction

You are given a dataset collected in the context of a research study. The goal of the study was to understand the relationship between the use of generative AI and human creativity, and in particular on how the use of language models affects creative output through the expression of short fiction.

The study was designed in two phases: (A) the writing and (B) the evaluation phase. In phase A, we recruited a group of "writers" who are asked to write a short eight-sentence story. Based on the randomly assigned study condition, writers would either write the story entirely from scratch on their own, could use one or up to five different AI-generated ideas for inspiration. In phase B, the stories composed by the writers were evaluated by a separate group of "evaluators". Evaluators read randomly selected stories without being informed whether writers used AI or not.

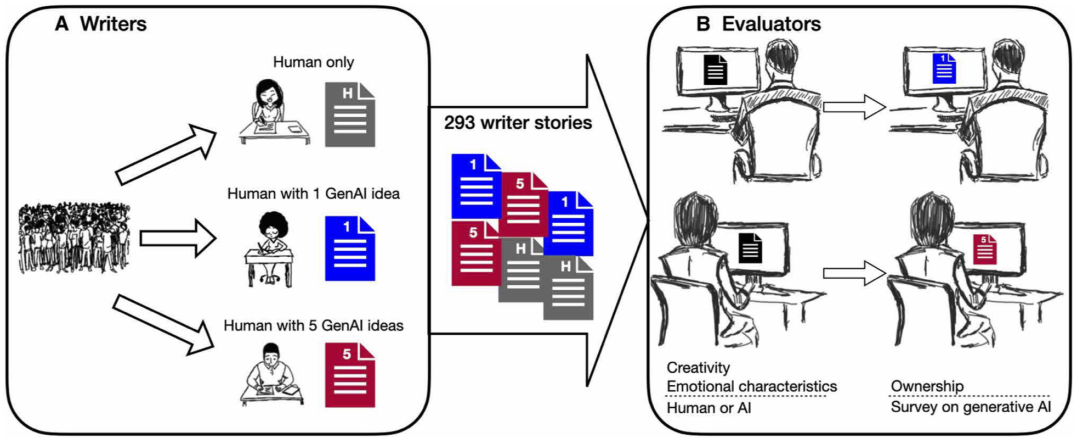

Figure 1: Study phases A and B

All stories were evaluated by multiple evaluators on novelty, usefulness, and several emotional characteristics.

Instructions

Perform a data analysis of the given dataset to answer an original research question. You are welcome to address additional research questions.

Deliverables

1. Marimo notebook containing all necessary code for the the data analysis and a written report. The notebook needs to run locally without errors and compute all results within 1 hour. If you require more complex calculations (e.g. vector embeddings of the short stories), please provide intermediate results as separate data files as well as the scripts used to create them.
2. Any intermediate data files, scripts or additional data used in your process. It is not necessary to include the original dataset you were given for the exam.

Report

In the report, please cover the following sections:

1. Abstract (max 200 words)
    - A brief summary of the whole project.
    - As a guide, include one sentence each for: project description, main research question, primary methods used, main results and overall conclusions.
2. Introduction
    - A shot description of the project where you have the chance to mention any specific personal interests or motivation for the project.
    - Explain what research question(s) you want to answer.
    - Where applicable, your alternative hypotheses and the corresponding null hypotheses.
3. Methodology
    - Explain which methods you used and why you chose them to answer specific research questions.
4. Results

    - Describe the results of your analyses.
5. Conclusions
    - Summarise your findings. Are your results conclusive or do any open questions remain?
    - Discuss what conclusions you draw from the results about your research questions.

Your submission will be marked on the three assessment criteria: Process, Communication and Realisation.

Study

Phase A: Writers

For the writer study (phase A), we recruited 293 participants. Writers were not selected based on prior writing skills or their creativity. We first asked each participant to complete the divergent association task (DAT) in which they are asked to write down ten maximally different English nouns. We use these words to compute the participants' DAT scores and determine their baseline creativity. Each participant was then provided with instructions to complete a story writing task. Participants were randomised into writing about one of the following three topics:

1. an adventure on the open seas
2. an adventure in the jungle
3. an adventure on a different planet

As an example, participants (using the "open seas" writing topic as an example) received the following instructions: "We would like you to write a story about an adventure on the open seas. You can write about anything you like. The story must be exactly eight sentences long and it needs to be written in English and appropriate for a teenage and young adult audience (approximately 15 to 24 years of age)."

Participants were randomly assigned to one of three conditions:

1. Human-only (human): 97 participants
2. Human with one generated idea (human_1AI): 100 participants
3. Human with five generated ideas (human_5AI): 98 participants

In the Human-only baseline condition, writers were assigned the task with no mention of or access to generative AI. In the Human with one generated idea condition and the Human with five generated ideas conditions, the participant had the option to

receive a three-sentence idea for a story from a language model. When a participant clicked on "Generate Story Idea...," we passed the following prompt to OpenAI's GPT API (again, using the open seas topic as an example): "Write a three-sentence summary of a story about an adventure on the open seas." The response from the API was passed to the participant. At the time of the study, we used the API from OpenAI's latest model, GPT-4. Those in the Human with one generated idea condition could only receive one story idea, while those in the Human with five generated ideas condition could receive up to five story ideas, each of which was visible to the participant. Participants were not able to copy and paste the generative AI idea text.

After completing their story, writers were asked several follow-up questions, and among others to evaluate the creativity of their story based on novelty and usefulness. We further asked them how much they agreed with six stylistic statements, including whether they enjoyed writing it, how well written it was, how boring it was, how funny it was, to what extent there was a surprise twist, and whether it changed their expectations of future stories (questions were asked in a random order across participants). We then asked participants about their view of story profits they should receive (as a percentage) and whether the story reflected their own ideas, as well as the novelty and usefulness of the story (on a nine-point scale). We also asked the Human-only condition whether they used AI to help them complete the task.

Phase B: Evaluators

The stories written by participants in phase A were evaluated by a different group of participants, the evaluators. Participants were not selected on the basis of prior experience in the publishing industry, but represent "regular" readers. Each evaluator was shown six stories (two stories from each topic). The order in which the stories were presented for review was randomised across evaluators. Evaluators were presented with one story at a time and asked to provide their feedback on the stylistic characteristics, novelty, and usefulness of the story. We presented the evaluator the same stories a second time and asked for an assessment of whether the story was written by a human or AI (as a percentage). We then disclosed whether the writer was offered the opportunity to generate an AI idea and, if so, whether the writer made use of it. If the author did use AI, we provide the evaluator with the text of the idea. Following that disclosure, we asked about the extent to which the story reflects the author's ideas and the extent to which the author has an ownership claim over the story. If the author used AI, we also asked the share of the profit the author should receive. After all story evaluations, we asked participants to assess six statements about the use of AI in writing stories.

Dataset

The dataset contains of three different parts of information in two files:

1. Information about the writers and the evaluations of their own stories.
2. The written short stories (all text).
3. Information about the evaluators and their evaluations of the stories.

Most dependent variables were assessed on a nine-point Likert scale from 1 (not at all) to 9 (extremely) to capture how much a participant disagrees or agrees with a statement or a question.

writers.csv

As an example, a few column names with an overview of the data they contain.

- Writer ID (`articipant.code`)
- Topic of the short story, randomly assigned to the writer (`participant.topic`)
- The writer's study condition, randomly assigned (`participant.condition`)
- Short story written by the writer (`ai_story_gen.1.player.story`)
- Short story id (`ai_story_gen.1.player.story_id`). Used to match stories and evaluations.
- Up to five AI generated story ideas (`ai_story_gen.1.player.ai_idea0` to `ai_story_gen.1.player.ai_idea4`). Generated by GPT-4 upon request. Writers in conditions `human_1AI` could request only one idea and up to five in condition `human_5AI`. Some columns may be empty, not all writers necessarily requested all possible ideas even if they could.
- DAT words to evaluate the writer's baseline creativity (`ai_story_gen.1.player.word1` to `ai_story_gen.1.player.word10`).
- Writer demographics, education and occupation
  - `payment.1.player.age`
  - `payment.1.player.gender`
  - `payment.1.player.education`
  - `payment.1.player.employment`
  - `payment.1.player.job_title`
  - `payment.1.player.income`

Divergent Association Task (DAT)

The writers were recruited without any limitations to their creative writing experience. Different people will thus have different baseline abilities. To estimate the baseline creativity of the writers, the authors asked them to perform the Divergent Association Task (DAT).

There are additional steps needed to calculate the DAT. Please refer to the README for instructions.

Documents

See the document 'Study questions.pdf' for a summary of writer and evaluator questions.

See the document 'Study screenshots.pdf' for screenshots of the online questionnaires.

Different columns in the dataset correspond to different questions posed to writers and evaluators. If you are having difficulties finding the data you need for your research question, please reach out to me (s.berns@arts.ac.uk).

Submission

All deliverables need to be submitted to the correct element in the Submission Area on Moodle.

The regular deadline for submissions is Wednesday, 10 December, 2 pm UK time (GMT).