Thinking About ChatGPT in Education: College Study, Mar 2023 (#125155)

Author(s)
Tal Waltzer (University of California, San Diego) - twaltzer@ucsc.edu
Gail Heyman (University of California, San Diego) - gheyman@ucsd.edu

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?
The main research questions are how good instructors and students are at differentiating between ChatGPT-generated essays and ones generated by college students, and what factors are predictive of this skill. We will look at factors associated with participants who try to make these judgments, with a focus on experience with ChatGPT, confidence, (and for instructors only, expertise in the topic). We will also look at effects of the quality of the essays written by students (as assessed by the average grade assigned by two graders).

This is an exploratory study, and we do not have specific hypotheses, but we do expect that both instructors and students will have trouble differentiating between these essays, scoring below 75% on average. We also expect that, compared to student-generated essays that are rated as less good, better student-generated essays will be harder to distinguish from ChatGPT-generated ones.

3) Describe the key dependent variable(s) specifying how they will be measured.
Students and instructors will see 6 pairs of short essays in response to one of two possible essay prompts. In each pair, one will be ChatGPT-generated and the other will be written by a college student. The specific prompts will be randomly selected from larger pools of responses (25 for each question for both ChatGPT-generated and student-generated ones). Participants will be told: "Please read the following pair of texts. One was written by a college student and one was written by an artificial intelligence chatbot (ChatGPT). Your job is to figure out which one was written by the chatbot." The primary DV will be percent correct.

4) How many and which conditions will participants be assigned to?
There are no different conditions, although instructors will be asked some additional questions about their background (e.g., "Have you ever taught or helped teach an undergraduate Psychology class?"). We will be comparing the results of students and instructors.

5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.
We will test our hypotheses by looking at correlations between possible predictors (e.g., confidence ratings and participant experience with ChatGPT), and we will compare the overall accuracy of students and instructors with a t-test.

6) Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.
We will exclude participants who do not indicate whether they are a student or an instructor. We will exclude participants who don’t answer any of the key questions in the survey (discrimination task, confidence in ability to discriminate, experience with subject area).

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 aim to have a minimum of 100 students and 100 instructors.

8) Anything else you would like to pre-register? (e.g., secondary analyses, variables collected for exploratory purposes, unusual analyses planned?)
In addition to our primary research questions, we will examine attitudes about ChatGPT. For example, all participants will evaluate 5 uses of ChatGPT (e.g., "A student uses Chat GPT to write an essay for them and submits the direct generated answer."). We will also ask participants to rate the extent to which they are concerned about ChatGPT in education (0 to 100), how optimistic they are about it (0 to 100), and how much they expect students to use ChatGPT to cheat (0% to 100% of students). We will compare responses of instructors and students, and we expect that students will have more favorable attitudes about ChatGPT than instructors.