

Notice

IRB STUDY #1806940835 INDIANA UNIVERSITY STUDY INFORMATION SHEET FOR

Defending Against Authorship Attribution Attacks

You are invited to participate in a research study of the distinctiveness of individuals' writing. You were selected as a possible subject because you are able to write English at a college level or higher and are at least 18 years old. We ask that you read this form and ask any questions you may have before agreeing to be in the study.

The study is being conducted by Allen Riddell of Indiana University. It is funded by the National Science Foundation.

STUDY PURPOSE

The purpose of this study is to investigate to what extent individuals' writing style is identifiable. People writing documents (memos, emails, books) tend to use distinctive vocabulary and syntax. The study investigates how much pre-existing prose is needed to identify an individual's writing style. The study also explores ways individuals can mask or obscure their distinctive writing style.

PROCEDURES FOR THE STUDY:

If you agree to be in the study, you will do the following things:

Provide us with at least 6,500 words of pre-existing formal prose you have

written (essays, memos, etc). You will then be asked to write a short essay of about 500 words. You will be asked to describe a familiar environment to someone unfamiliar with it. The entire exercise should take no longer than 60 minutes. You will also be asked to provide basic demographic information about yourself including approximate age and approximate region of residence.

The prose samples you provide and the short essay you will write during this study will both be entered into a public collection of writing samples and may be used for future research. No identifiable information will be shared.

RISKS AND BENEFITS

The risks of participating in this research are minimal. Writing approximately 500 words in English is not an activity associated with known risks.

There is also a risk of loss of confidentiality. This risk is minimized if you do not include any personally identifying information in your submitted prose documents.

You are not expected to benefit from participating in this research.

CONFIDENTIALITY

Efforts will be made to keep your personal information confidential. We cannot guarantee absolute confidentiality.

No personally identifying information will be collected during this research.

Organizations that may inspect and/or copy your research records for quality assurance and data analysis include groups such as the study investigator and his/her research associates, the Indiana University Institutional Review

Board or its designees, the study sponsor, the National Science Foundation, and (as allowed by law) state or federal agencies, specifically the Office for Human Research Protections (OHRP).

PAYMENT

You will receive payment for taking part in this study. You will be paid \$7 if you complete the research exercise. If you, in the judgment of the researchers, respond to the questions carefully, you will be paid a bonus of \$3.

CONTACTS FOR QUESTIONS OR PROBLEMS

For questions about the study, contact the researcher Allen Riddell at 812-856-4730.

For questions about your rights as a research participant or to discuss problems, complaints or concerns about a research study, or to obtain information, or offer input, contact the IU Human Subjects Office at (800) 696-2949. You may also email irb@iu.edu.

VOLUNTARY NATURE OF STUDY

Taking part in this study is voluntary. You may choose not to take part or may leave the study at any time. Leaving the study will not result in any penalty or loss of benefits to which you are entitled. Your decision whether or not to participate in this study will not affect your current or future relations with Indiana University.

Thank you for agreeing to participate in our research. Before you begin, please note that the data you provide may be collected and used by Amazon

as per its privacy agreement. Additionally, this research is for residents of the United States over the age of 18*; if you are not a resident of the United States and/or under the age of 18, please do not complete this survey.

This research is intended for individual 18 years of age or older. If you are under age 18, do not complete the survey.

This research is for residents of the United States. If you are not a U.S. resident, do not complete the survey.

Demographic Questions

What is your sex?

- ☐ Male
- ☐ Female

What is your age?

- ☐ 18-34 years old
- ☐ 35-49 years old
- ☐ 50-64 years old
- ☐ 65 or more years old

Upload writing samples

Submit at least 6500 words total from multiple documents of your own writing that was done for a formal purpose (school essays, grant proposals, etc). Formatting is not important. If possible please try to avoid writings on

very specific topics with lots of highly specialized technical terms and acronyms. Ideally the passages you submit should be reasonably suited for a general audience. All submissions must be your own work. **Please ensure you remove any and all personal information from the documents you submit.**

Guidelines to Follow

- The submitted pre-existing writing to be “scholarly” in nature (i.e: a persuasive piece, opinion paper, research paper, journals, etc.).
- Anything that is not the writing content of the work should be removed (i.e: citations, urls, section headings, editing notes, etc.).
- The papers/samples should have a minimal amount of dialogue/quotations.
- Please refrain from submitting samples of less than 500 words, laboratory and other overly scientific reports, Q&A style samples such as exams, and anything written in another person’s style.

You may submit up to eight documents using the forms below:

Upload one document here.

Upload one document here.

Upload one document here.

Upload one document here.

Upload one document here.

Upload one document here.

Upload one document here.

Upload one document here.

Before you submit, please check the following box to acknowledge that you understand that all information you submit through these forms will be added to a public collection of writing samples that may be used for research purposes:

☐ I understand that my submission will be included in a public, open source research repository.

Experiment

Write approximately 500 words (+/-25) on the following topic.

TOPIC: You are asked as part of a college application to describe your neighborhood to someone who has never been there before. Discuss the houses, people, stores, parks. Anything you think is relevant.

Write approximately 500 words (+/-25) on the following topic but in a way in which you are trying to obscure your own writing style. Your goal is to fool the computer into thinking that your passage was NOT written by you. You may use whatever means you wish so long as the writing would not raise any eyebrows when a human reads over it (no scrambled words, mixed up semantics, etc) and the point is still clearly conveyed.

TOPIC: You are asked as part of a college application to describe your neighborhood to someone who has never been there before. Discuss the houses, people, stores, parks. Include anything you think is relevant.

Read the sample passage from *The Road* by Cormac McCarthy:

<http://homes.sice.indiana.edu/riddella/theRoadSample.txt> (link opens in a new tab/window)

Write approximately 500 words (+/-25) on the following topic but in a way in which you are trying to mimic the writing style of Cormac McCarthy. Your goal is to fool the computer into thinking that your passage was written by Cormac McCarthy. You may use anything you wish from the example passage as clues as to what signifies his writing style, but again remember that your writing should not raise any eyebrows when read by a human and the point should still be clearly conveyed.

TOPIC: You are asked as part of a college application to describe your neighborhood to someone who has never been there before. Discuss the houses, people, stores, parks. Include anything you think is relevant.

You will be asked to write using Special English, a simplified version of English. There are two basic rules for writing in Special English:

1. Write using words in the Special English word list, available here: <http://www.manythings.org/voa/words.htm> (link opens in a new tab/window). (Using an occasional proper noun which is not on the list is acceptable.)
2. Write using short, simple sentences. Try to use the simplest sentences that make sense.

Write approximately 500 words (+/-25) on the following topic but do so using Special English to hide your writing style. Your writing should not raise any eyebrows when a human reads over it and the point should still be clearly conveyed.

TOPIC: You are asked as part of a college application to describe your neighborhood to someone who has never been there before. Discuss the houses, people, stores, parks. Include anything you think is relevant.

Before you submit, please check the following box to acknowledge that you understand that all information you submit through these forms will be added to a public collection of writing samples that may be used for research purposes:

☐ I understand that my submission will be included in a public, open source research repository.

Comments

Comments (optional)

Feel free to enter any comments you have about this survey.

Powered by Qualtrics