

Who Wrote This?

The Use of Artificial
Intelligence in the Academy

Rahul Kumar, PhD (rkumar@brocku.ca)

Michael Mindzak, PhD (mmindzak@brocku.ca)

Rachel Racz, MEd (rracz@brocku.ca)

Land acknowledgement

Brock University, where this study was conducted, sits on the traditional territory of Anishinaabeg and Haudenosaunee peoples, many of whom continue to live and work here today. The territory is covered by the Upper Canada Treaties and is within the land protected by the Dish with One Spoon Wampum Agreement. Today these lands are the home to many First Nations, Métis, and Inuit peoples and acknowledging them reminds us that our great standard of living is directly related to the resources and friendship of Indigenous people. Recognition and respect are essential elements of establishing healthy, reciprocal relations. These relationships are key to reconciliation.

Agenda

Introduction

Literature Review

Research Question

Methodology

Findings & Limitations

Discussion & Conclusion

Introduction

- If the current AI technology is sophisticated enough, it might be difficult to distinguish small compositions whether they are human generated or computer generated
- A simplified and modified Imitation Game (Turning Test) was constructed to investigate if university students and faculty can accurately identify who wrote a small composition
- This presentation reports these results and discusses implications for PSE instructors and administrators



Literature Review

Academic Integrity, Plagiarism, & Contract Cheating

- The educational sphere is increasingly surrounded by the use, advancement, and integration of artificial intelligence (Selwyn, 2019; Kerr, 2020)
- Increased sophistication and undetectability of AI writing software LLMs like GPT-3 is posing a threat to academic integrity (McKnight, 2021)
- Detection of AI generated text from their human counterparts is increasingly difficult (Luitse & Denaka, 2021; Dehouche, 2021) and insufficient (Abd-Elaal et al, 2022)
- Our traditional understanding of plagiarism, academic integrity, and authorship is likewise being rapidly reshaped by technological developments (Mindzak & Eaton, 2021)

Research Questions

- How sophisticated is a readily accessible AI tool for text generation?
 - Can participants distinguish human generated text from AI generated one?
 - Who is better at determining the origin of the text?
- What are the implications for assessment and evaluation?

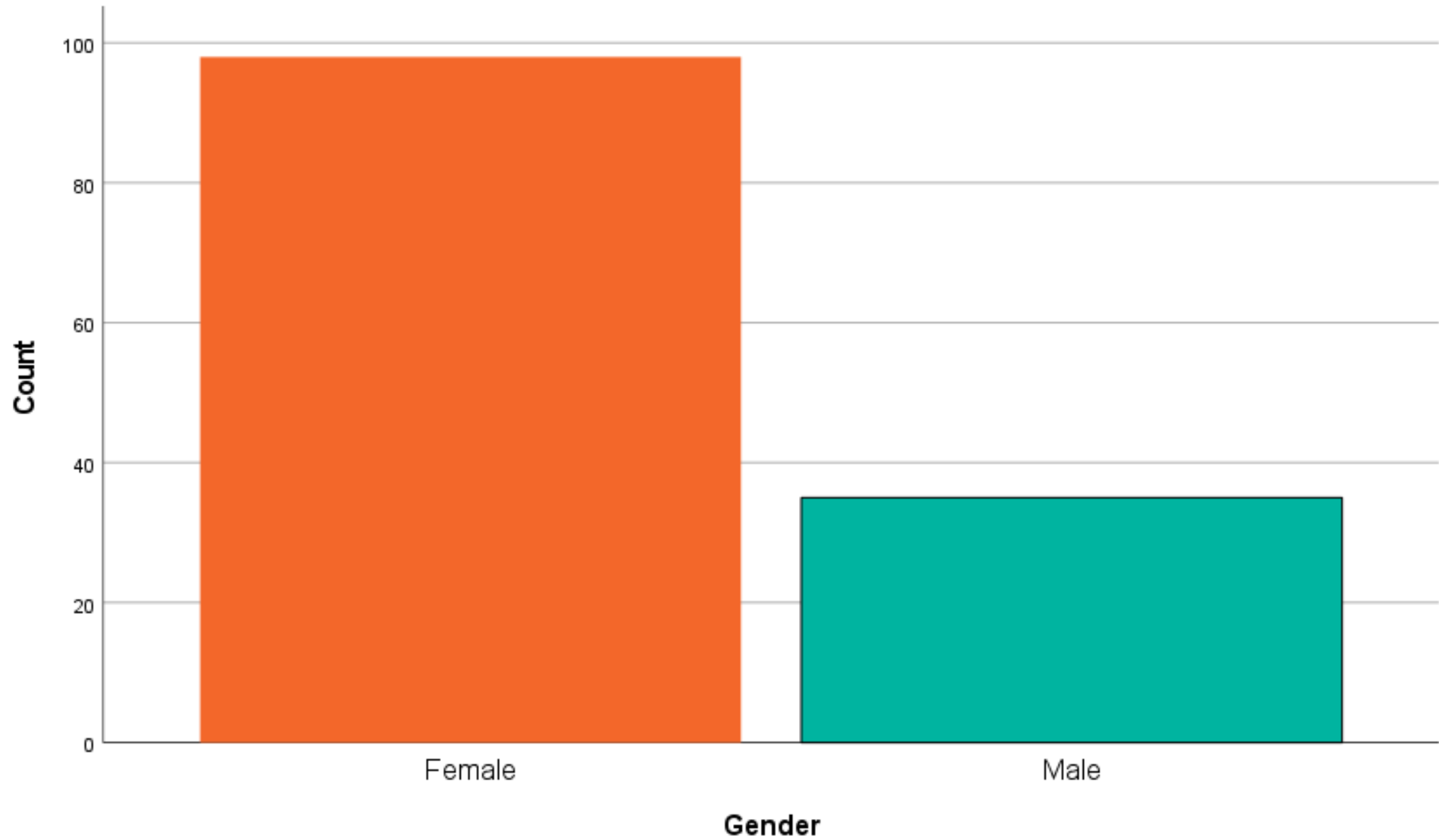
Methodology

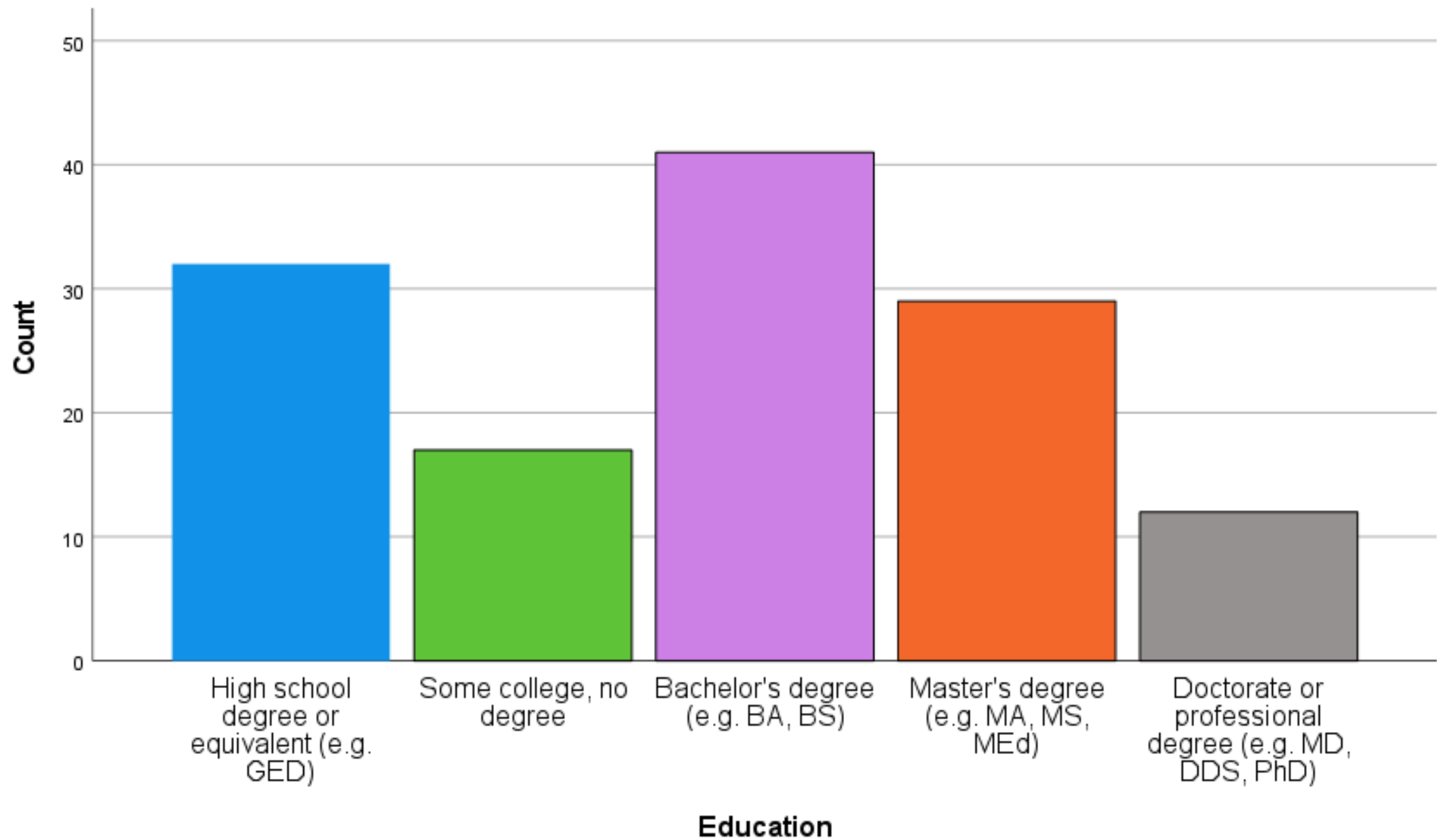
- Research Design
 - Participant demographic characteristics were collected
 - Three pools of writing samples:
 1. Human generated (By our research team)
 2. AI generated (Inferkit, using GPT-3)
 3. Found from the Internet
 - Participants shown two random passages from these pools and asked to identify the author and origin of the passages presented
- Analysis
 - Descriptive
 - Correlation
 - Logistic Regression

Findings

- $N = 135$ (including 17 partial responders)
 - Two (of 6) writing evaluation passages were randomly assigned to participants. Thus, regarding analyses, predicting Author (A)1-6, or Origin (O)1-6, the sample sizes are smaller:

Author 1	$n = 40$	Origin 1	$n = 40$
Author 2	$n = 37$	Origin 2	$n = 37$
Author 3	$n = 40$	Origin 3	$n = 40$
Author 4	$n = 35$	Origin 4	$n = 26$
Author 5	$n = 37$	Origin 5	$n = 37$
Author 6	$n = 39$	Origin 6	$n = 39$





Correlations

		Age_winsoriz ed	Education	Gender
Age_winsorized	Pearson Correlation	1	.556 ^{**}	.202 [*]
	Sig. (2-tailed)		.000	.020
	N	134	131	133
Education	Pearson Correlation	.556 ^{**}	1	.037
	Sig. (2-tailed)	.000		.678
	N	131	131	130
Gender	Pearson Correlation	.202 [*]	.037	1
	Sig. (2-tailed)	.020	.678	
	N	133	130	133

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).



Logistic Regressions (predictors on their own)

- **Gender** was **not** a significant predictor of accuracy of authorship or the origin
- **Age** was not a significant predictor of accuracy of authorship or the origin
- **Education level** was **marginally** significant predictor
 - **More educated** got the **lower probability** of getting the answer correct (Author for the first passage - A1) [$b = -.382$, $p = .072$, OR = .682]
 - **More educated** got a **higher probability** of getting the answer correct for the Origin for the fourth passage (O4) [$b = .422$, $p = .098$, OR = 1.52]
 - Not significant for other passages



Logistic Regressions (predictors entered together)

- Controlling for education and gender, **older** (vs. younger) **participants** had a **lower probability** of getting the answer correct for A5 ($b = -.866$, $p = .082$, OR = .421) and A6 ($b = -.715$, $p = .076$, OR = .489)



Demographics and Marks Assigned to Passages

Correlations:

- Participants who were older (vs. younger) assigned lower marks to the fifth passage ($r = -.42, p = .009$).
- Participants who were more (vs. less) educated assigned lower marks to the first ($r = -.32, p = .051$), third ($r = -.34, p = .033$), and fifth ($r = -.44, p = .007$) passages.
Note: no gender differences emerged.

Simultaneous Regressions:

- No significant predictors of marks assigned when all 3 predictors were entered together, except:
 - Controlling for age and gender, **participants** who were **more** (vs. less) **educated** assigned **lower marks** to the passages

Discussion

- Because participants were inclined to ascribe the writing sample composed by AI to humans, we conclude that it is difficult for participants to differentiate writing samples written by humans vs those generated by a machine
- Ongoing developments in the field of artificial intelligence and large-language models continue to improve—likely meaning that “machine writing” will become further indistinguishable from human writing
- How these digital technologies might be utilized in teaching and learning will also likely continue to evolve—requiring new ways of thinking about matters such as plagiarism, academic integrity and the assessment/evaluation of writing.

Implications

- Implications for Students
 - Plagiarism & Academic Integrity
 - AI Writing Tools, Products & Services
- Implications of Educators
 - Assessment and Evaluation of Writing
 - AI Writing and Digital Pedagogies
- Implications for Researchers
 - Authorship, Publication & Academic Ethics
 - Finding relevant variables/characteristics to detect Writing

Conclusion

- The present study provides a glimpse into the future intersections of AI and writing in higher education
- While the study has several limitations, further research on the topic can continue to investigate how educators are responding pedagogically to the evolution of these technologies
- Nevertheless, AI writing technologies will in all likelihood significantly reshape the ways in which we thinking about writing in higher education in the near future

References

- Abd-Elaal, E., Gamage, S. H.P.W., & Mills, J. E. (2022). Assisting academics to identify computer generated writing. *European Journal of Engineering Education*.
<https://doi.org/10.1080/03043797.2022.2046709>
- Dehouche, M. (2021). Plagiarism in the age of massive generative pre-trained transformers (GPT-3). *Ethics in Science and Environmental Politics*, 21, 17-23.
<https://doi.org/10.3354/esep00195>
- Gehrmann, S., Strobel, H., & Rush, A. M. (2019). GLTR: Statistical detection and visualization of generated text. arXiv.
<https://doi.org/10.48550/arXiv.1906.04043>
- Kerr, K. (2020). Ethical considerations when using artificial intelligence-based assistive technologies in education. In B. Brown, V. Roberts, M. Jacobsen and C. Hurrell (Ed.). *Ethical Use of Technology in Digital Learning Environments: Graduate Student Perspectives*. (pp. 9-14). Pressbooks.
<https://openeducationalberta.ca/educationaltechnologyethics/>
- Luitse, D., Denkena, W. (2021). The great transformer: Examining the role of large language models in the political economy of AI. *Big Data & Society*. 8(2)
<https://doi.org/10.1177/205395172111047734>
- McKnight, L. (2021). Electric sheep? Humans, robots, artificial intelligence, and the future of writing. *Studies in Culture and Education*. 28(4), 442-455.
- Mindzak, M., & Eaton, S. A. (2021a, November 9). *Paging the plagiarism police—AI is getting better at writing students' essays*. The Next Web.
<https://thenextweb.com/news/plagiarism-improving-writing-students-essays-syndication>
- Mindzak, M., & Eaton, S. A. (2021b, November 4). *Artificial intelligence is getting better at writing, and universities should worry about plagiarism*. The Conversation. <https://theconversation.com/artificial-intelligence-is-getting-better-at-writing-and-universities-should-worry-about-plagiarism-160481>
- Selwyn, N. (2019). *Should robots replace teachers? AI and the future of education*. Polity Press.