Undergraduate Journal of Contemporary Issues & Media

Winter
2025

Load

The Undergraduate Journal of
Contemporary Issues & Media

Volume 7 | Winter 2025

Content Editorial Assistants

Andi Barton

Brant Campanian

Parker Dunn

Thabata Fay

Cole Hanck

Sequoia Hennigs-Cornell

Zabrina Le

Emmerson Martin

Cora Romero

Shay Wilson

Design Editorial Assistants

Andi Barton

Hillary Curd

Patrick Foley

Shinaya Harris

Lisa Kasabyan

Samantha Kivett

Zabrina Le

Emmerson Martin

Cora Romero

Jessica Sipes

Emma Zeigler

Cover design by Lisa Kasabyan

Table of Contents

Letter from the Editors

Rachel Bryson,
Co-Editor

Maureen Mathison,
Co-Editor

We are thrilled to present this year's edition of the *Undergraduate Journal of Contemporary Issues & Media* with its focus on artificial intelligence. As many readers are no doubt aware, Generative Artificial Intelligence (GenAI) has reshaped much of how we interact with creative processes in all forms of media. The articles in this volume, from undergraduate students across the globe, interrogate the personal, cultural, and academic implications of this emerging technology. While the various articles are diverse in their approaches and methods, they share in common a commitment to nuanced understanding of GenAI and its affordances and constraints. Several writers attend specifically to questions of ethics and emerging technology. Others explore GenAI's applications in academic writing and research processes. Still more examine GenAI's connections to literature and popular culture. In short, this collection offers a snapshot of how undergraduate students are wrestling with what GenAI means—in society, in their personal lives, in their educations, and more.

We would be remiss if we did not acknowledge the stellar work of our undergraduate editorial assistants, whose efforts shape and enable the publication of this journal. Our content editorial assistants worked with individual authors over the course of a full semester to prepare these articles for publication. Our design editorial assistants brought the journal's design to life, carefully attending to issues such as readability, accessibility, and visual design. The work of these undergraduate student editors and authors is the heart of this journal, and the care and collaboration represented in these pages should serve as a model for academic publishing. We hope you enjoy the results of their collective efforts.

ChatGPT & Student Writing

Catalina Barraza, Salt Lake Community College

Introduction

In November 2022, the software company Open AI launched ChatGPT. Since then, instructors and administrators across the nation have scrambled to find ways to prevent students from abusing ChatGPT's artificial intelligence to do their schoolwork for them (Anders, 2023; Winter, 2023). At the same time, students and others have found unique ways to use ChatGPT to do their work for them (i.e., brainstorming, drafts, structure, and editing) through its humanlike response (Anders, 2023). Currently, we do not understand if a student presenting ChatGPT's work as their own is plagiarism or if it is an advanced composition tool.

Students have been warned of the 'no plagiarism' rule since the beginning of their education in composition classrooms—and for good reason. Oxford Dictionary defines plagiarism as "the practice of taking someone else's work or ideas and passing them off as one's own" (NP). The unique problem with catching plagiarism from ChatGPT is that it's not directly plagiarizing one source. ChatGPT combines all the knowledge it was "trained on" to output human-like responses (Open AI FAQs). Oftentimes, older plagiarism checkers used to see if students are plagiarizing from the internet or a database don't catch plagiarism from AI because of ChatGPT using its own "thoughts," which is possible through Reinforcement Learning with Human Feedback (RLHF).

At this point, there are many unanswered questions about using this tool, and it is worth investigating if ChatGPT can write a college-level assignment. In this paper I will examine if ChatGPT can meet the requirements of length, citations, and use of sources like those expected of a college student, and if it cannot, how close can it get? Further, I'll be discussing the student in the classroom who turns to ChatGPT and encourage instructors to reflect on their own teaching styles and assignment sheets. Are students using ChatGPT because of their lack of engagement, or is it perhaps an issue that starts in the classroom?

Literature Review

ChatGPT is a Large Language Model that produces its own coherent, human-like dialogue. This artificial intelligence software was launched to the public on November 30, 2022, and was created by Open AI. According to Open AI, "ChatGPT was optimized for dialogue by using Reinforcement Learning with Human Feedback (RLHF)." RLHF is "a method that uses human demonstrations and preference comparisons to guide the model toward desired behavior" (Open AI, n.p.). ChatGPT gets its human-like dialogue because it was trained on "vast amounts of data from the internet written by humans" (Open AI, n.p.). Open AI discusses the credibility of ChatGPT or the possible lack thereof on their site, "ChatGPT is not connected to the internet, and it can occasionally produce incorrect answers. It has limited knowledge of the world and events after 2021 and may also occasionally produce harmful instructions or biased content." ChatGPT is not a consistently credible source of information, which raises concerns for students using the software as a tool to do their work for them. ChatGPT's credibility is only one of many concerns that instructors and scholars in academia are discussing.

Because ChatGPT is so new, there are few peer-reviewed sources about how it will affect student writing classrooms. Instead, what is available so far are speculations of ChatGPT's effect on students' critical thinking and writing skills and composition classrooms in general. Brent Anders, in "Is using ChatGPT cheating, plagiarism, both, neither, or forward thinking?" reminds us that AI will be a forever part of our world and specifically the workforce. This means that AI literacy will now become a workforce skill that students will need to know to succeed (Anders 2023). Xiaoming Zhai (2023) agrees with Anders in "ChatGPT user experience: Implications for education" by maintaining that AI is a part of everyone's future, including in education (p. 9). Many scholars have concluded that we need to include artificial intelligence in learning spaces (Anders, 2023; De Winter, 2023; Thorp, 2023; Zhai, 2023) to keep up with the fast-evolving world. This could potentially look like "using AI tools to conduct subject domain tasks" that way students can focus on "creativity and critical thinking rather than general skills" (Zhai, 2023, p. 10).

We must ask: when students use AI to generate answers for them, are they missing out on the satisfactory and necessary process of writing as described in Composition Studies? Donald Murray (1972), in "Teach Writing as Process Not Product," has hopes that

students will see writing as "the process of discovery through language," which "is the process of exploration of what we know and what we feel about what we know through language" (p.2). Nancy Sommers (1980) furthers that by talking about the importance of fulfillment when writing in "Revision Strategies of Student Writers and Experienced Adult Writers." Sommers argues that "it is a sense of writing as discovery" (p. 387). Will students find the 'discovery,' as Murray and Sommers describe, if they have software to find their sources and ideas for them?

Many scholars are concerned that students will graduate without having become proficient in the important learning processes that schools teach, specifically gaining critical thinking skills (Bishop, 2023; Bleumink, & Shikule, 2023; University of Oxford 2024). One definition of a student being able to think critically means the student will be able to be "good at figuring out what is true", and "exercising good judgment in deciding what to do with the truth once you find it" (Bishop, 2023, 16-17). As an experiment, I asked ChatGPT why it's important for students to learn critical thinking skills in school. Here is its response: "It equips them [students] with tools to navigate complex situations, evaluate information critically, and make informed choices, which are crucial skills for success in academics, careers, and life" (ChatGPT 4, Open AI). ChapGPT only knows this definition because of its ability to synthesize the information fed to it. Most people can agree that critical thinking skills will help students in all aspects of their lives. When we are given an answer from ChatGPT, are students really learning? Is this kind of information gathering different from plagiarism? It's unclear at present.

We might also ask how much a teacher and their assignments will determine whether a student will turn to ChatGPT. And, in the case that the student does, will the teacher be able to recognize the usage of AI in a student's paper? In "A Computer Wrote This Paper: What ChatGPT Means for Education," Lea Bishop (2023) maintains that a key way for teachers to find the difference between student writing and AI writing is to "Look for substance, not style. Ideas, not information. Look for critical thinking" (p. 17). Tracing a similar idea, Arend Groot Bleumink and Aaron Shikule (2023), co-founders of *AICheat Check* (a tool for teachers to check plagiarism in their students' papers), argue that teachers are key to preventing students from plagiarizing, "an important reason [for students] to plagiarize included a poor explanation by teachers" (p. 2). Instructors have the opportunity to not only give students the resources and help they need to not feel like plagiarizing is their only option, but to also make connections with their students to know when AI pla-

giarism is present (Bishop, 2023; Bleumink & Shikule, 2023).

Returning to early compositionists, Murray (1972) argues that we should teach "unfinished writing, and its glory in its unfinishedness" (p. 2). If instructors teach the idea that students are not expected to be professionals in their writing, then students would respect where their writing skills are and give their essays a shot for themselves. Examining how instructors coach and hopefully encourage students to embrace writing could be a big difference in whether a student turns to ChatGPT or not.

Teachers and administrators may need to adapt to the new world that students are living in outside of the classroom. I'm curious what it says about the writing assignment given to students if something like ChatGPT can write it. ChatGPT is a software that we know is incapable of critical thinking and text to self connections. "ChatGPT doesn't need to be thoughtful, reflective, or creative. It simply does what it is told to do" (Bishop, 2023, p. 16). Holden Thorp (2023), in "ChatGPT is Fun but Not an Author" remarks that ChatGPT "did well finding factual answers, but the scholarly writing still has a long way to go. If anything, the implications for education may push academics to rethink their courses in innovative ways and give assignments that aren't easily solved by AI" (p. 313). Is the only skill needed to pass a college level writing assignment syntax and grammar? If so, ChatGPT should be able to do that just fine.

Methods

Below I discuss the three methodologies I used in conducting the research and analysis for this project.

The Student's Writing

For this project, I am using an anonymous sample of student writing while using ChatGPT to write the same prompt in an attempt to see which piece of writing is more successful. The student example I am using to compare to ChatGPT is a 20 year old college student majoring in English. The student has an indie post-college rock band and works at a local record shop. This fact is important to know since the student draws on their own experiences with the music industry in their researched argument essay like in this excerpt, "I am someone who works in a dying—or perhaps long dead—industry. Three days a week, I brave a mixture of vintage-crazed teenagers, audiophiles and retired hippies/Vietnam war vets at [a local record shop]'' (Author 1). Attached to their essay were their assignment sheet and the score they received in their first-year composition course.

I took the writing assignment sheet from the college student and determined what would be valuable to put through ChatGPT to get a response similar to the student's work. Next, I compared the different choices that AI made versus the student writer I am using as an example.

The assignment sheet lets the student pick a topic of their choice and highlights a general outline of an argumentative essay and directs students "to review [their] field's literature in order to situate your argument in the conversation and show your reader that you understand the conversation that is going on and that you have something to add to it." The sheet gives format guidelines for a researched argument essay: claims, supported by good evidence, formal tone, logical connections between ideas, paragraphs structured around topics that move the reader smoothly and logically through the paper, a thesis in the first section of the paper, in-text parenthetical citations, a works cited page, titles and headings, 9-12 pages long, etc. (Instructor 1).

To set up my comparison, I asked the student to send me their essay, the writing prompt, the grade they received, and the course in which they were enrolled. When I found out the essay was a topic of the student's choice I made sure to read the students' essay first to craft an input for ChatGPT that would be comparable enough to the student's sample.

ChatGPT's Writing

The students' essay was titled "How has TikTok Changed the Economic and Cultural Influence of Popular Music?" (Author 1, 2022, 1). At first, I prompted ChatGPT to "Write me an essay about how has Tik-Tok changed the economic and cultural influence of popular music." Here is where I first ran into ChatGPT's five-paragraph format issue. The assignment sheet the student gave me required the paper to be 9-12 pages long. ChatGPT was roughly outputting about 500 words. The students' 9-page essay was 3235 words. Further, I added the word criteria to the prompt for ChatGPT, "Write me a 3200-word long essay about how has TikTok changed the economic and cultural influence of popular music." ChatGPT could not stay on topic past about 1500 words and the essay felt surface level.

In an attempt to get a more thoughtful and lengthy response from ChatGPT, I changed my wording input to: "Write me a college-level 3200 word researched argument essay about How TikTok has changed the economic and cultural influence of popular music." Once again, I was left unsatisfied with the answer. I questioned if I added too many criteria because

ChatGPT began to ignore it. My next move was to simplify. "Write me a researched argument essay about how has TikTok changed the economic and cultural influence of popular music." This time, I could tell ChatGPT was pulling information from sources. After asking ChatGPT to do MLA in-text citations, it could only output a Works Cited page.

Here is where I discovered that ChatGPT has a generating limitation. At this point, the software can only generate so many words before it stops. This limitation explains why the first prompts I was giving ChatGPT created short essays. This limitation would create a problem for students like the one I am studying since there is a 9-12 page requirement. I did a quick Google search to see if anyone has found a way to go around the limiting barrier. Unsurprisingly, this problem was well documented. In the article "Here's the ChatGPT word limit and how to get around It," Fionna Agomuoh (2023) mentions that ChatGPT's word limit is around 4000 characters or 500 words. A way around the limit is to simply say "Go on'' or in the beginning prompt write "Write me the first 500 words of a 3200 word essay about..." (Agomuoh, 2023). I tried both ways and was able to push Chat GPT to give me 1644 words before it completely abandoned the subject and just talked about TikTok with no connection of the app to music. To keep up with the word limit I set for ChatGPT, the software began to write about body positivity and mental health amongst its younger users. These points have no connection to music and ChatGPT doesn't try to make that any less apparent. Even with the hacks to get around the barrier of words, we still see the limitations of AI in producing long coherent essays like a student can.

After I was able to input instructions the program understood, I got a 1644 word essay from ChatGPT. I put the ChatGPT essay into a Google doc and set it in MLA format with a Works Cited page with the ten sources the program generated.

Comparing the Student to ChatGPT

Next, I annotated both essays, noting their differences and similarities in structure, layout, and contents. After I did my own initial analysis of the two essays side by side, I turned to the instructor's assignment sheet and created a rubric to follow and grade the essays on my own. Once I created the rubric, I graded both essays out of 50 points because that is how the instructor graded the original student's essay. This process makes each of the 9 categories roughly about 5 points each, with an extra 5 points lingering in limbo up to the instructor to decide. The student example I used got 47/50 on his essay from the instructor, so I have attempted recreating that score in my rubric

while also grading ChatGPT's essay the same as the student.

Here is the grading rubric based on the instructor for the student and ChatGPT:

Criteria	Student	ChatGPT
Claims supported by good evidence	5	1
Formal tone	4	5
Logical connections between ideas	4	5
Paragraphs structured around topics that move the reader smoothly and logically through the paper	5.6	5
A thesis in the first section of the paper	5.6	5
In text parenthetical citations	5.6	0
Works Cited	5.6	5
Titles and headings	5.6	5
9-12 pages	5.6	1
Totals:	**47/50**	**32/50**

Table 1: Table showing the grading rubric used to analyze the essays from the student and ChatGPT

Discussion

We see the student scored 15 points higher than ChatGPT based on the instructor's grading rubric. ChatGPT could hit the requirements of formality, thesis, and most format requirements but was missing the deep connections in the literature review and was incapable of doing in-text citations despite what was entered as the prompt for the AI to output. ChatGPT is incapable of making thoughtful text that connects to the self (Bishop, 2023, p. 16). I told ChatGPT to write me an essay about how TikTok affects the music industry and it did just and only that.

ChatGPT is successful in the mechanics of writing, it "consistently writes with nearly perfect grammar, syntax, spelling, punctuation, vocabulary, and sentence and paragraph structure" (Bishop, 2023, p. 7). ChatGPT's essay is in formal third person writing. While the student's essay uses "I" statements, the assignment sheet does say "formal tone" as one of the criteria and technically formal tone is in third-person writing. But does ChatGPT make "claims supported by

good evidence," as the instructor desired, if it cannot make text that connects to something personal? This is the problem we run into with this assignment sheet and grading rubric. What are "good evidence," "formal tone," and "logical connections," and what exactly makes a transition go "smoothly and logically" (Instructor 1)? These types of words and descriptions can be confusing to students because they are subjective. It blurs the line between what constitutes a good grade from a bad one.

In the student's essay, we see that they recall their own life experience to relate to the argument they are trying to portray. ChatGPT cannot make these types of choices in writing due to it not being a human. This output constitutes a generic essay that anyone can write, even software. ChatGPT also explains well-known concepts and the "wisdom of the crowd" (Bishop, 2023, p. 7) in its writing. What ChatGPT lacks is critical thinking. Bishop explains that critical thinking includes being aware of your assumptions and biases. The student author makes a connection to their job at a record shop and that being the old-fashioned way of hearing new music but now we see TikTok becoming the new wave for artists to find fame (Author 1). The student author working in the physical media music industry creates a bias towards a more internet approach to music. ChatGPT could never make that connection unless it was specifically entered into the input.

Here is what ChatGPT has to say on behalf of record stores and discovering music:

> Record stores have long been a staple of the music industry, serving as a hub for music enthusiasts to discover new artists and albums. However, with the rise of digital music platforms and streaming services, record stores have faced significant challenges in recent years. Despite this, many record stores have continued to thrive, thanks in part to the resurgence of vinyl records and the growing interest in physical media among music enthusiasts. For some music fans, record stores remain the best way to discover new music, offering a unique and personal experience that cannot be replicated online. While TikTok has certainly changed the way that many people discover music, it is important to remember that there are still many other avenues for music discovery, including the traditional record store (ChatGPT).

This is what the student wrote about on the same topic:

The consumption of popular music is ever chang-

ing. For most of the twentieth century, physical media was king. People found music they liked from Record Stores, the radio, or a bootleg cassette tape someone gave them, ripped directly from their own collection of Vinyl and CD's. People had stereos, turntables, cassette decks, walls of records and CD's. Nick Horsby in his book *High Fidelity* calls record collecting "(not) like collecting stamps, or beermats, or antique thimbles. There's a whole world in here, a nicer, dirtier, more violent, more peaceful, more colorful, sleazier, more dangerous, more loving world than the world I live in." (Hornsby 38). Relationships were formed by lending or purchasing a physical piece of music. Artists could expand on their music by adding whatever they wanted onto the covers (Author 1).

Another of the limitations I noticed of ChatGPT was its incapability of in-text citations, such as the one the student used above from Nick Horsby (Author 1). ChatGPT can synthesize information but cannot show directly where it got that information from using in-text citations.

Additionally, a technical struggle I ran into was copying and pasting ChatGPT's response onto a Google document. The background is highlighted in black no matter my attempts to fix it. After a Google search, I found a Chrome extension called *Copy for ChatGPT* launched by web and mobile developer Sethu Senthil. The extension will copy the contents of ChatGPT and paste them into a document like normal. Another thing *Copy for ChatGPT* does is when I copied contents a little box would appear and mention whether the content I am copying could set off 'AI detectors' like Bluemink and Shikule created (AICheatCheck). When it did have the possibility to set off the detectors it would say "This text was flagged as AI generated by plagiarism detectors, make some changes before submitting" (*Copy for ChatGPT*). What's interesting here is that using ChatGPT for coursework may require additional types of digital literacy.

Similarly, ChatGPT touched on the same subject as the student about small, independent artists. Here is what ChatGPT said about smaller artists: "The platform has become a powerful promotional tool for new music, enabling independent and emerging artists to gain exposure and build a following" (p. 5). Now here is what the student said about independent artists: "People [artists] create followings with likes, instead of fans at their shows'' (Author 1), and the student further explains how the song *Supalonely* by Benee took off because of an influencer who created a dance on Tik Tok to the song. "Her video has since amassed more than 45 million views, turning her into a Tiktok celebrity

and helping to make Benee a global sensation" (Author 1). In the instructor's assignment sheet, they mention "Claims supported by good evidence" (Instructor 1). The student used specific examples that I would say are "good" while ChatGPT just made a claim and has no evidence. The specificity of the human writing makes it better.

Students often learn about synthesizing sources in first-year writing courses. It is interesting to examine how ChatGPT does this work. The AI bot takes all its knowledge from a large database, 570 GB to be exact, ranging from sources such as books, articles, Wikipedia, and other pieces of writing (Hughes, 2023). Still, in this essay, ChatGPT mentions "Lil Nas X's 'Old Town Road,' which initially gained popularity on TikTok before becoming a global hit" (Chat GPT, 2023, 1). We see that this information is not just known knowledge and the AI pulled this from a direct source. The problem was from where is ChatGPT pulling this information, and is it peer-reviewed? Factual or trustworthy? These are the problems we run into with essays written with AI. The sources that form the synthesis are not clear and there is no discernment in ChatGPT selecting them. We cannot interrogate its sources if we don't know what they are. Unfortunately, the student may be either out of time or doesn't know how or what to look for; we cannot assume the student would take that extra precaution to check its sources. So, I did. I went to the Works Cited page from ChatGPT where it created ten references for me. Here are the stats on those references:

- Only 1 of the 10 was a scholarly source
- Only 2 of the URLs pulled up articles or videos that existed
- 8 of the references were "404: Page not found"

With these results, the integrity of ChatGPT appears to be wildly inaccurate. Did ChatGPT just create these articles and names and pretend they came from mass media sources such as the *New York Times* or *The Guardian*?

Imagining a classroom where a ChatGPT essay passes when compared to the student's example is concerning. It reminds me of what my English professor said in regard to this project: ChatGPT will be the factor that forces teachers to get rid of "bad" writing assignments. The elimination of bad assignments is because ChatGPT is incapable of the critical thinking that a "good" college writing assignment requires. We must wonder what it means for the assignment if ChatGPT, which is incapable of creativity, thoughtfulness, reflectiveness, critical thinking, and credibility is able to pass (Bishop, 2023). Teachers need to have a

connection well enough to students that 1) the teacher can determine something off in the student writing that might connect to an AI generator like ChatGPT, and 2) students are more likely to turn to ChatGPT when they don't understand something in class and don't have time to cram for an assessment (Bishop, 2023; Bleumink & Shikule, 2023).

Conclusion

The effects ChatGPT has on classrooms start with plagiarism. Students copying others' work or, in this case, a software's work, prevents students from the satisfaction of writing and the process of discovery through writing. Students who plagiarize will not learn critical thinking skills. Critical thinking is a skill that carries into the rest of their lives and is a great tool for students to harness in school. Despite it being easy to blame students alone for plagiarism from ChatGPT, it beckons a call to instructors everywhere on whether they are adequately equipping their students to feel confident enough to do their own work. Students are more likely to turn to ChatGPT when they don't understand a subject in school. In the case a student does turn to ChatGPT, does the instructor know their students well enough to notice? Does the instructor have an assignment the program can complete? Can ChatGPT even pass as a functioning student brain to an engaged instructor?

My tiny research project compares one student's essay to one ChatGPT wrote. From this small sample, we might infer that ChatGPT cannot adequately create a long enough essay with consistent, reliable, and coherent thought. The software caps out at about 500 words, and if you tell the AI bot to 'go on', it can, but it begins to include filler that does not contribute to the essay's main ideas. Further, at this time, ChatGPT cannot complete accurate in-text citations, have solid evidence, or back up its claims with specific evidence. ChatGPT is incapable of making text-to-self connections. Text-to-self connections are a key factor to show that a student is learning and applying their learning to their own life. One final finding from using ChatGPT is its inability to be trusted. As found, ChatGPT will create sources that do not exist. Instead of finding a credible source, ChatGPT will string a bunch of information together, creating something that does not exist.

When comparing ChatGPT's essay to the student's example, the student rose above in categories of good evidence to back up claims, drawing on personal insight, citations, MLA format, length requirements, text to text connections, text to self connections, and maintaining academic integrity while harnessing their critical thinking skills. What ChatGPT was able to do,

but at a less well-executed level, was the mechanics of writing such as spelling, grammar, and punctuation.

In today's composition classrooms, teachers are trying to find ways to prevent ChatGPT from being used, and in the case it is used, how to recognize it. I believe the answer is simple: address the elephant in the room. Explain to students what ChatGPT is. More importantly, explain why ChatGPT prohibits the benefits of learning, critical thinking, and blocks students from the discovery of writing (Sommers 1972; Murray 1972). And, without being insensitive to teacher burnout, teachers should attempt to make themselves available to their students for questions about lessons or deeper explanations to the students who need it. This engagement will be the key to making students not feel so alone in assignments and turn to plagiarism. ChatGPT is a new tool, and it's too soon to tell if it will cause good, bad, or in the most probable outcome, mixed effects in the education world. What we do know so far is that ChatGPT causes instructors, admins, and board of education members to question course programs. If AI can write an essay for a course and pass, what does that say about the learning that is supposed to take place in the classroom?

Catalina Barraza

Catalina Barraza is a sophomore majoring in Writing Studies at Salt Lake Community College. Born and raised in West Valley City as a second-generation Mexican-American, Catalina was a teen peer educator for Planned Parenthood and continues to be passionate and outspoken about social justice and the issues that impact the communities around her.

Catalina is the editorial assistant for the National Council of Teachers of English (NCTE)–*Teaching English in the Two-Year College Journal*. In addition, she is a Promoter Representative at Kilby Court in Downtown Salt Lake City. Both of these jobs continue her education in writing and publication studies, fostering community and leadership skills.

In the past year, Catalina has been a keynote speaker at Utah's Coalition Against Sexual Assault's (UCASA) statewide conference and was a keynote speaker for YCC's Family Crisis Center's Teen Healthy Relationships Summit.

Works Cited

Agomuoh, F. (2023, March 17). How to easily bypass the ChatGPT word limit. *Digital Trends*. www.digitaltrends.com/computing/easily-bypass-the-chatgpt-word-limit/#:~:text=To%20remedy%20this%20issue%2C%20you.

Anders, B. A. (2023). Is using ChatGPT cheating, plagiarism, both, neither, or forward thinking? *Patterns, 4*(3).

Author 1 (2022). How has TikTok changed the economic & cultural influence of popular music.

Bishop, L. (2023). A computer wrote this paper: What ChatGPT means for education, research, and writing. *SSRN Electronic Journal* https://doi.org/10.2139/ssrn.4338981.

Bleumink, G. & Shikhule, A. (2023). Keeping AI Honest in Education: Identifying GPT-Generated Text. *Educato AI Research*.

ChatGPT (2023). How TikTok Changed the Economic and Cultural Influence on Popular Music.

De Winter, J. (2023). Can ChatGPT pass high school exams on English language comprehension?" *International Journal of Artificial Ingelligence in Education, 34*, 915-930.

Hughes, A. (2023, Jan. 4). ChatGPT: Everything you need to know about OpenAI's GPT-3 tool. *BBC Science Focus Magazine*, www.sciencefocus.com/future-technology/gpt-3/.

Murray, D. (1972). Teach writing as a process not product. *The Leaflet, 71*(3), 11-14.

University of Oxford (n.d.). Plagiarism. *www.ox.ac.uk*, www.ox.ac.uk/students/academic/guidance/skills/plagiarism#:~:text=Passing%20off%20another%27s%20work%20as.

Senthil, S. (2023, Apr. 20). Sethu Senthil. *Sethusenthil.com*, sethusenthil.com/#aboutMe.

Sommers, N. (1980). Revision strategies of student writers and experienced adult writers. *College Composition and Communication, 31*(4), 378-388.

Thorp, H. H. (2023). ChatGPT is fun, but not an author. *Science, 379*(6630), 313. https://doi.org/10.1126/science.adg7879.

Zhai, X. (2022) "ChatGPT User Experience: Implications for Education." *Social Science Research Network*, RELX Group (Netherlands). https://doi.org/10.2139/ssrn.4312418.

Society's New Oracle: An Ethical Analysis of ChatGPT and its influence

Leena Bath, Utah State University

Introduction

In this era of rapid technological advancement, ChatGPT has become a global sensation, emerging as a transformative tool that has already begun to shape the dynamics of our ever-evolving world. Developed by OpenAI, ChatGPT stands at the forefront of generative artificial intelligence (AI), using sophisticated algorithms to simulate human-like conversation. These algorithms, known as "generative pre-trained transformers," enable ChatGPT to generate text that closely resembles human speech (What is ChatGPT?, 2023). With its wide range of applications, ChatGPT has found its way into various aspects of our daily lives, revolutionizing how we interact with technology. However, beyond its capabilities lies a complex web of interactions and moral quandaries that shape our engagement with this tool, society, and our world. This growing complexity necessitates a deeper exploration of the ethics surrounding ChatGPT. Using Foucauldian ethics to analyze this issue offers a unique perspective, enabling a comprehensive analysis of ethical nuances, societal implications, and embedded structures within this seemingly omnipresent AI phenomenon.

Technical Foundation of ChatGPT

ChatGPT is based on natural language processing (NLP), a branch of AI that amalgamates linguistic understanding with computer programming. At its core, NLP strives to endow machines with the ability to comprehend, interpret, and generate human language, essentially bridging the gap between the intricacies of linguistic expression and the computational power of machines. This fusion of linguistics and computer science serves as the backbone for ChatGPT's remarkable language generation capabilities. Within the realm of NLP, ChatGPT is distinguished by its reliance on what is called a neural language model (NLM). This model uses neural networks as a way to perform predictive language modeling and give it a basic "understanding" of language structures (Language Models Explained, 2024), and it represents an innovative approach in machine learning, particularly in the domain of language processing. Through the NLM, ChatGPT is able to gain an understanding of the contextual nuances of language, allowing it to gener-

ate responses that closely mimic human conversation (What is ChatGPT?).

This mastery of language is largely due to the pre-training phase of ChatGPT. In this phase, the AI model is exposed to extensive corpora of textual data, spanning various linguistic expressions and styles (Language Models Explained, 2023). The model's language proficiency is then shaped by the dataset, giving it the ability to understand and respond to an array of prompts. The pre-training process not only imparts linguistic knowledge, but also instills a sense of contextual awareness, enabling a conversation that, upon interaction, provides an experience akin to engaging with a human. This proficiency in simulating natural conversation has positioned it as a transformative tool across various domains.

Current Uses of ChatGPT

Despite its relatively recent introduction into society in late 2022, ChatGPT has rapidly become a versatile tool, finding application in various aspects of daily life. Personally, I frequently use ChatGPT to simplify and enhance various aspects of my life. The range of applications is diverse: from crafting satirical romance stories to debugging code, revising papers, simplifying complex topics, initiating brainstorming sessions for projects, generating gift ideas for friends and family, and creating recipes from random fridge contents. I even use it as a tutor in the writing center, where I guide students in using it as a valuable tool for revising their own work. Teaching students to leverage ChatGPT as a writing revision tool is particularly valuable, as it not only saves them time and alleviates stress, but also imparts skills in engaging with a technology that is likely to remain prominent in the foreseeable future.

Beyond these uses, ChatGPT can also play games, reword content, create articles, find keywords, help you prepare for your epic trivia night showdown, and even compose a love poem that is sure to make your true love swoon and accept your proposal. It can help draft emails, do research, have conversations, answer questions, streamline presentations, and organize information (What is ChatGPT?, 2023). It is even used

by many as a quasi-search engine, with users typing in questions to receive immediate answers.

Recently, I have been noticing a growing trend in conversations and classes where the phrase "let's ask ChatGPT" is used to answer a hard or confusing question. This seems to depict the AI model as an all-knowing entity, similar to the way society views search engines like Google. This reflects a shift in perception, with people turning to ChatGPT like they are consulting the oracle in The Matrix, (Wachowski, et al., 1999) highlighting its emerging role as a go-to source for information and insights in various spheres of life. This trend underscores ChatGPT's growing influence in our daily interactions.

ChatGPT and the Rise of the Internet

The parallels between ChatGPT and the rise of the internet are striking. Just as the internet revolutionized communication, ChatGPT introduces a new dimension to interactions. Initially emerging as a tool for government researchers in the 1960s, the internet's integration into public life expanded gradually over decades (A Short History). Similarly, the development of ChatGPT began in 2015, undergoing iterative improvements before its widespread release in 2022 (ScriptByAI, 2023).

Both technologies exhibit a pattern of gradual integration into society. Just as early internet models were not widely adopted until later stages, ChatGPT's initial versions saw limited use before becoming integral to daily conversations. This evolution suggests a trajectory wherein generative AI tools, like ChatGPT, could become ubiquitous, transforming academic, social, and professional spheres.

Despite their advantages in facilitating communication and access to information, both ChatGPT and the internet face challenges, such as credibility issues and ethical dilemmas. The evolving nature of these technologies complicates our understanding and ethical evaluation, emphasizing the need for continual reassessment.

In examining these parallels, we can anticipate the future integration of ChatGPT into society, while acknowledging the complexities and ethical considerations that accompany its development and use.

Foucauldian Ethical Framework

In this context, the insights of Michel Foucault can function as guiding principles for navigating the ethical complexities of AI use and integration. Foucault's framework allows us to engage in self-reflection in a way that allows us to think about our interaction with technology and how it affects us. In her article "Ethical Discourse and Foucault's Conception of Ethics," Mary Candace Moore gives a great summary of Foucauldian ethics when she says, "Ethics does not refer to a body of rules or prescriptions for right conducts, but rather it refers to the self-forming activity whereby one seeks to develop insight by practicing the inquiry that makes possible philosophical ethics" (Moore, 1987, p. 82). Essentially, she is saying that ethics don't have to mean a definitive answer of what is right and wrong but can be the process of analyzing yourself and your activity, changing your actions accordingly. Too often, society wants us to draw clear cut lines between what is ethical and what is not, but in this case, we cannot reasonably make many of those judgements because of the nature of new and adapting technologies.

Brenton Faber elaborates on Foucault's ethical framework, highlighting its emphasis on self-awareness and critical engagement. According to Faber, Foucault posits that ethics transcend mere intuition; ethics involves a profound understanding of underlying values and ideologies that shape our intuitive responses. This ethical orientation not only informs the way we interact with and understand organizations and institutions in our lives, but also influences the relationship everyone has with themselves. This is because Foucault's ethics are based on the idea of identifying power structures that are attempting to normalize your decisions, and deciding which structures you will or will not allow to affect you. He argues that one must first understand how their individual beliefs and values have been created and affected in order for them to act ethically, and that to do otherwise, "Without sufficient knowledge of the naturalized values that might be motivating oneself, or another individual, is unethical" (Faber, 2009, p. 195).

Technology will continue to change, and as such, so will its ethical and unethical uses. Using Foucault as an ethical framework allows us to analyze these issues not as fixed categories of 'ethical' and 'unethical,' but as products of ongoing power dynamics and social forces. This approach enables us to look at elements of the technology in a way that we can analyze and reevaluate as time goes on, assessing what is truly considered ethical in relation to it.

This analysis will focus on a few key elements of Foucault's concept of ethics. First, organizational power and power structures; second, normalization; and third, surveillance and privacy.

Foucauldian Analysis

Organizational Power in Foucault

In Foucault's view, power functions in a transfomative way because "It does not link forces together in order to reduce them; it seeks to bind them together in such a way as to multiply and use them" (Foucault, 1999, p. 97). Foucault saw organizational power as a threat to our individual autonomy and, as such, said that "Only after understanding the ways in which competing forces are attempting to naturalize one's behavior, thought, and judgement can one say one is prepared to act ethically" (Faber, 2009, p. 196). His ethical guidelines require that we look at how we are affected by organizational power structures to evaluate how such structures are affecting individual autonomy. Power structures in Foucault are interesting because instead of being defined as we would initially think—a well-established organization, like governments and schools, that are often seen as oppressive—power structures can be anything and everything that affects and changes us. Because of this, Foucault does not speak of power structures as oppressive forces, but rather like those involved in the creation of mountains, functioning "to 'train,' rather than to select and levy" (Foucault, 1999, p. 97). In this case, various forces interact, redirecting each other to form novel outcomes, without having equal influence in the final result.

ChatGPT functions as a power structure under this understanding, in part, due to the way that it has been embedded within our societal fabric. It is rapidly becoming present in everyday life, turning into one of the forces that act on us daily and influence our decisions. As we interact with ChatGPT, it actively influences our thoughts and perspectives in a very unique way that is greatly affected by the power structures, sources, and vast amounts of data that have gone into the making of the generative model itself. This dynamic relationship reflects the evolving role of technology in shaping societal narratives, and highlights the ways in which we are constantly affected and shaped by it.

Foucault's philosophical framework compels us to dissect the intricate power dynamics pervading our lives, prompting a meticulous examination of ChatGPT and its intricate role within our societal fabric. The increasing perception of ChatGPT and the internet as comprehensive information sources is generating tensions and influencing the way our society is developing. The extensive training data driving ChatGPT's functionality introduces inherent inaccuracies and biases, underscoring the need for critical evaluation from the perspective of power dynamics and effects. Foucault encourages an examination of the power structures in our lives, necessitating scrutiny of ChatGPT and its role in our society, in turn prompting us to look at the power structures influencing the AI model itself.

One such power structure influencing ChatGPT can be found in the training process, where a process called reinforcement learning with human feedback (RLHF) takes place. In this process, human testers help to align the LLM outputs with things that humans value, but because everyone values different things, this creates bias. The model is shaped by the people giving the feedback, not the users themselves, raising concerns about the inclusivity and diversity of those shaping the model (Wu et al., 2023). In this way, the power structures influencing us as users of the AI program are exponential because we are in part affected by each person behind RLHF, and the power structures that, in turn, have shaped them and influenced their ideas.

The pre-training phase of ChatGPT is likely the most important part of the AI model to analyze in respect to power structures. This is best exemplified in the recent legal issues surrounding ChatGPT, including accusations of plagiarism that highlight challenges tied to its active learning nature (Alter & Harris, 2023). The model continuously refines responses based on interactions, posing a risk of reproducing content without proper attribution. Because of this, the *New York Times* has filed a lawsuit against OpenAI and Microsoft for copyright infringement (Grynbaum & Mac, 2023). The *New York Times* case is one of many lawsuits against the generative AI program, but "is the first major American media organization to sue the companies... [contending] that millions of articles published by The Times were used to train automated chatbots that now compete with the news outlet as a source of reliable information" (Grynbaum & Mac, 2023). ChatGPT is, at its core, simply an amalgamation of countless sources that inform both the way it produces language and the information it provides. Because of this, we must be aware that ChatGPT is simply spouting out information that is on the internet, being greatly affected by the broader biases and tendencies prevalent in the data it learns from.

Foucault's theories about the influence of power structures on individuals and their autonomy find resonance in the ongoing discourse surrounding ChatGPT. Questions about its utility, potential harm, and societal impact echo concerns about the broader consequences of technology. The current narrative, filled with uncertainties and fears, underscores the need for careful consideration and ethical engagement with this emerging technology.

In navigating these intricate dynamics, society must remain vigilant, critically assess the implications, and engage in ongoing conversations about the ethical dimensions of AI. This involves not only understanding the technology, but also actively shaping it to align with or challenge societal values and norms. As we continue to grapple with these challenges, the ethical interaction envisaged by Foucault necessitates a collective commitment to fostering awareness, responsibility, and autonomy in our engagement with emerging technologies like ChatGPT (Faber, 1999).

Foucault's Theory of Normalization

In Foucault's theory of normalization, the historical reliance on clearly defined lines to define social norms has transitioned into a more abstract conceptualization. This shift places the norm, a very subjective and elusive concept, at the core of societal evaluation. The norm "differentiates individuals from one another, in terms of the following overall rule: that the rule should be made to function as a minimal threshold, as an average to be respected or as an optimum towards which one must move" (Foucault, 1999, pp. 101-102). Individuals are measured based on their proximity or distance from this norm, with closer alignment indicating greater perceived normality and acceptability.

However, this societal inclination toward the 'norm' poses significant challenges, as it problematically suggests that conformity and loss of individuality is the pathway to societal acceptance. Diversity is essential for a vibrant society and encompasses a wide range of personalities, ideologies, and perspectives. When individuals gravitate towards normalization, the result is a reduction in societal diversity, fostering an almost mechanical uniformity, resplendent of many utopian and dystopian novels alike, and compromising individuality. Diversity is the propellant of creativity. Without it, innovation and development essentially could not occur (Why Is Diversity So Important?, 2024).

Individuality in society can also have ethical implications. Having individuality is what creates and allows for one being an "ethical being; it implies the possession of all our powers, thoughts, qualities, opinions, standards, values, so that we are determined by ourselves, not by society" (Lindsay, 1920, p. 423). While a utopian vision of universalized ethics may be appealing, it can also lead to a stifling conformity that erodes personal identity. Conflict and diversity, integral components of societal dynamics, serve as catalysts for change and the essence of a fulfilling life. The concept of societal normalization thus raises distressing concerns. In this context, technologies, like ChatGPT, play a pivotal role. ChatGPT's normalization of language and perspectives may inadvertently contribute to a homogenized societal discourse, potentially diminishing individuality and diversity. As ChatGPT becomes increasingly integrated into communication channels, the risk of stifling individual expression and promoting conformity looms larger; hence, ethical considerations surrounding ChatGPT's impact on societal norms and individual identity become paramount, urging us to reflect on the broader implications of its usage in shaping our ethical landscape.

Foucault's framework provides a nuanced perspective by characterizing this normalization as a spectrum, saying that "In a sense, the power of normalization imposes homogeneity, but it individualizes by making it possible to measure gaps, to determine levels, to fix specialties, and to render the differences useful by fitting them one to another" (Foucault, 1999, p. 103). The idea of the norm functions by encouraging individuals to choose where they want their position to be relative to the norm. However, the perpetual competition to reach the norm persists.

In the context of ChatGPT, this notion of 'distance from the norm' takes on a potentially problematic dimension. Unlike humans, ChatGPT lacks a nuanced judgment of diversity and individuality. The NLM underlying ChatGPT tends to normalize information, research, and opinions, providing more diverse perspectives only upon specific prompts. This inherent bias in information dissemination contributes to the normalization of ideas and perspectives within societies.

While ChatGPT may not completely eliminate societal diversity, it runs the risk of reducing diversity on topics less familiar to certain communities. As globalization continues and ChatGPT becomes increasingly integrated into various facets of daily life, it serves as a facilitator of normalization without the personal judgment characteristic of human interactions. Consequently, there's a heightened risk of diminishing the diversity of voices, particularly within minoritized communities whose perspectives may not align with the mainstream or dominant narratives. This has already been occurring in society, as everyone works towards a more 'standard' version of language, communication, and overall acceptable existence. Tools like ChatGPT could perpetuate this cultural homogenization, imposing dominant cultures onto minority groups. ChatGPT's normalization process could act as an ongoing form of colonization, suppressing or overshadowing the unique cultural expressions and identities of marginalized communities.

Furthermore, by perpetuating mainstream narratives and perspectives and engaging in the process of

normalization, ChatGPT may inadvertently reinforce existing power dynamics that marginalize minority groups. This can result in the loss of cultural heritage, identity, and autonomy, as well as reinforcing the current caste system and its accompanying prejudices and discriminative practices. Therefore, it is imperative to critically examine and address potential consequences of AI normalization to ensure the preservation and celebration of diverse cultures and identities.

Surveillance and Privacy

Surveillance and privacy are quickly becoming crucial focal points in the context of AI, invoking concerns that align with societal apprehensions often expressed about new technologies. Foucault's ethical framework allows us to scrutinize personal privacy and possible intrusions of it, becoming important considerations in assessing the impact of new technologies on individual autonomy.

Personal autonomy violations are most likely to arise from the surveillance and data collection capabilities inherent in AI systems. The question then becomes: How could the surveillance and data amassed by AI potentially limit our personal autonomy? The answer lies in the intricate web of information that AI can gather, process, and utilize to make predictions or recommendations. The extensive data surveillance may lead to the shaping of individual behaviors, preferences, and even decision-making processes based on algorithms and patterns discerned from the collected data. This, in turn, could constrain personal autonomy by influencing and directing individuals along predetermined paths, routes, or trajectories set by the AI system itself, or by those who design or program it. These paths may be established based on various factors such as algorithms, data patterns, or intended outcomes. This could constrain personal autonomy by guiding individuals along these predetermined paths, subtly influencing them to conform to perceived norms or expectations inherent in the AI system.

ChatGPT utilizes user prompts to refine its responses over time. While this data doesn't necessarily retain specific information about each interaction, it does mean that any data provided by individual users is now part of the collective AI database and can never be deleted. This raises concerns about infringement on personal and intellectual property privacy, even though users willingly contribute this information because they do not fully grasp the potential implications of their actions. (ChatGPT and Data Privacy, 2024).

The data collection practices of ChatGPT extend beyond the content of interactions, delving into spe-

cifics like IP addresses, dates, times, browsers, and device-related information. It also uses cookies to gather analytic information. Additionally, if you engage with the tool on social media, it can collect personal information from your profiles (Govin, 2023). The lack of transparency surrounding this data collection process is especially disconcerting. Users are often unaware of the extent and implications of the data gathered, engaging with the tool without a clear understanding of the potential risks.

Although ChatGPT claims to utilize collected data solely for response training purposes, the lingering concern revolves around the risk of a data breach and the potential compromise of personal information. The growing web of surveillance and data collection in AI technologies, exemplified by tools like ChatGPT, poses a significant challenge to individual autonomy, demanding a more transparent and ethically sound approach to safeguard user privacy and autonomy in the digital age.

Ethical Recommendations for Engaging with ChatGPT

First—when engaging with ChatGPT, be aware of what you are reading. Don't just take its answers for granted. Think about them. Fact-check the information with credible sources before you believe it.

Second—analyze ChatGPT's responses for bias. Ask yourself, "What biases could be present in its responses?" As you analyze for bias, keep in mind that this process is something we all encounter regularly in various contexts. Whether we're evaluating news articles, social media posts, or personal opinions, we're constantly assessing for potential biases. Remember that biases in responses don't necessarily invalidate the information or insights provided; however, they do require careful consideration when interpreting and utilizing that information. It's an essential skill for maintaining objectivity and critical thinking in our everyday lives.

Third—be wary of using ChatGPT as a creation tool. ChatGPT doesn't ever actually "create" content—it takes an average of all language on the internet and regurgitates it at you in a different order. Nothing it says is completely new; its output is all based on patterns and data. There is no real creativity. So, when you get something from ChatGPT, make sure you ask where it may have come from first. Do research to find those answers in other places, and cite your sources. Give people credit for their work!

Fourth—ask ChatGPT for multiple perspectives and more information on topics. Don't just be satisfied with the surface level answers! Dig deeper and take back your decisive power when it comes to learning and information.

Fifth—be aware of what information you include in your responses. Is it confidential or personal information that you don't want to be included in the collective AI intelligence?

Sixth—make yourself aware of policies and practices regarding information and data use, as well as the terms and conditions of the tools you are using. When you are using the tool, you are agreeing to their terms and conditions, so make sure you know what you are agreeing to instead of simply clicking agree.

Engaging with any tool requires training and practice. With generative AI becoming such a prominent part of society, and the expectation of it becoming even more integrated, it is important that we begin learning how to use it and engage with it in ethical ways now, rather than later. Schools should teach students how to do this, allowing them to critically interact with new AI tools, and teaching them ways of approaching its use and analyzing its outputs. This proactive approach ensures that students develop the necessary skills to navigate AI tools responsibly and ethically in an increasingly AI-driven world.

Humanizing AI: Infusing Empathy and Creativity

The use of technologies like ChatGPT seems to be a dehumanizing act, as it is taking humanity out of certain acts and even, in some cases, replacing human interaction. It raises concerns about the risks of diluting the warmth and compassion inherent in human interaction by replacing it with machine integration. I vividly recall an experience of receiving a message after a breakup that sounded robotic, almost as if it were generated by ChatGPT. Although it is unclear if the message was, in fact, generated by AI or not, the message is the same: compassion and empathetic language is essential in many human interactions. This encounter left me feeling even more hurt than before, as it seemed that the individual didn't even care about me or the breakup. This highlights the need to infuse AI with a more humane touch. ChatGPT's responses can be great, but, ultimately, it is not a human, and therefore cannot exhibit the same empathy or emotion that a human could. As humans, we can use our judgment to determine the level and type of humanity needed in communications, and help apply that to content that may otherwise be AI generated.

Conclusion

These considerations and examinations of ChatGPT's relationship and impact on society barely scratch the surface of the complex and evolving landscape of AI. While this exploration has shed light on certain facets, the interplay between technology and human interaction is multifaceted and continually evolving. Self-reflective processes serve as a wonderful starting point for navigating this intricate terrain, but they aren't a definitive roadmap for understanding the full scope of the issue. It is important to keep in mind the intricacies inherent in the integration of AI technologies, recognizing that the ethical implications and societal effects are nuanced and context dependent.

Ethical AI use is a dynamic endeavor – one that requires constant scrutiny, adaptation, and a keen awareness of the evolving technological landscape. Users can navigate the complexities, mitigate potential inaccuracies, and address emerging challenges in this ever-evolving technological society as they engage in thoughtful reflection. This proactive approach aligns with Foucault's vision of ethical interaction—a continual process of questioning and adapting.

As the dialogue between humanity and technology continues, ethical engagement becomes even more important as an endless journey. By cultivating mindful awareness of the implications of AI tools like ChatGPT, we empower ourselves to shape a future where the integration of technology and human values coexists harmoniously. As we navigate this intricate terrain, our commitment to ethical considerations will play a pivotal role in defining the contours of a technologically augmented society.

Leena Bath

Leena Bath is a student at Utah State University, pursuing a degree in Technical Communication & Rhetoric with minors in Chinese and Music. She is passionate about AI, and her academic journey has led her to explore the intersection of AI, writing, and ethics in various projects and coursework. At USU's Writing Center, Leena assists fellow students in revising their work while advocating for ethical AI utilization. She has done several research projects delving into the inner workings of generative AI and its implications in both academia and the workplace. Beyond her academic pursuits, Leena is an avid reader, writer, mountain biker, and adventurer, always eager to explore the boundaries of her curiosity and passion for exploration and discovery.

References

National Science and Media Museum (2024). A short history of the internet. https://www.scienceandmediamuseum.org.uk/objects-and-stories/short-history-internet.

Alter, A. & Harris, E. A. (2023, Sept. 20). Franzen, Grisham and other prominent Authors Sue OpenAI. *The New York Times*. https://www.nytimes.com/2023/09/20/books/authors-openai-lawsuit-chatgpt-copyright.html.

ChatGPT and Data Privacy (2024). *Tripwire*. https://www.tripwire.com/state-of-security/chatgpt-and-data-privacy.

Faber, B. (1999). Intuitive ethics: Understanding and critiquing the role of intuition in ethical decisions. *Technical Communication Quarterly*, 8(2), 189–202. https://doi.org/10.1080/10572259909364659.

Foucault, M. (1999). The Means of Correct Training. In A. Elliot (Ed.), *The Blackwell reader in contemporary social theory* (pp. 97-106). Malden, Mass: Blackwell.

Govin, G. (2023, May 12). ChatGPT is a privacy nightmare. *Home of Internet Privacy*. https://www.expressvpn.com/blog/chatgpt-privacy-concerns/.

Grynbaum, M. M. & Mac, R. (2023, Dec. 27). The Times sues OpenAI and Microsoft over A.I. use of copyrighted work. *The New York Times*. https://www.nytimes.com/2023/12/27/business/media/new-york-times-open-ai-microsoft-lawsuit.html.

Language Models Explained. (2024). *AltexSoft*, https://www.altexsoft.com/blog/language-models-gpt/. Accessed 6 Jan. 2024.

Lindsay, J. (1920). The ethical value of individuality. *The International Journal of Ethics 30*(4), 423–449. https://doi.org/10.1086/intejethi.30.4.2377132.

Moore, M. C. (1987). Ethical discourse and Foucault's conception of ethics. *Human Studies*, *10*(1), 81-95. JSTOR, http://www.jstor.org/stable/20008989.

ScriptByAI. (2023, Dec. 21). Timeline of ChatGPT. *ScriptByAI*. https://www.scriptbyai.com/timeline-of-chatgpt/.

Wachowski, A., Wachowski, L., et al. (1999) *The Matrix*. USA/Australia.

"What Is ChatGPT? Everything You Need to Know." *WhatIs.Com*. https://www.techtarget.com/whatis/definition/ChatGPT. Accessed 4 Dec. 2023.

Why is diversity so important. (n.d.). *Steps4Change*, https://www.steps4change.org/why-is-diversity-so-important/.

Wu, T., et al. (2023). A brief overview of ChatGPT: The history, status quo and potential future development. *IEEE/CAA Journal of Automatica Sinica, 10*(5), 1122–1136. DOI.org (Crossref), https://doi.org/10.1109/JAS.2023.123618.

THE IMPACT OF ARTIFICIAL INTELLIGENCE ON STUDENT WRITING IN AFRICA

Confidence C. Chibuzor, Edo State University, Nigeria

Introduction

Artificial intelligence has emerged as one of the most transformative technologies in recent years, revolutionizing various sectors including education. In general terms, AI refers to computational tools that can substitute for human intelligence in the performance of certain tasks such as learning, problem-solving, and decision making. This technology has the potential to address some of the significant challenges faced by educational systems in Africa. Significant challenges include access to quality education, personalized learning experiences, and improved student outcomes. In higher education, writing is an important part of the curriculum. Writing is often a means of assessing what students know about a topic, and writing-to-learn is often implemented as an instructional technique (Schumacher & Nash, 1991).

Writing is an essential skill that plays a crucial role in the life of students, both in secondary schools and universities. This is because it serves as a tool for communication, self-expression, critical thinking, and academic success. A vast majority of students in Africa and beyond can attest that at many instances they have been given a topic, project, or academic paper to write, which would either make or mar their grades in school. Due to this importance, the use of artificial intelligence for assistance cannot be avoided because many students are either limited academically or are lacking in the basic use of English language.

In Africa, where education is seen as a key driver of social and economic progress, AI-powered writing tools have potential to play a transformative role. By harnessing the power of AI, African students can improve their writing skills, hone their critical thinking abilities, and achieve better academic results. AI can also provide personalized feedback and guidance, as well as access to a wealth of information and resources that can help students grow as writers and thinkers. One of the main benefits of AI in student writing is its ability to provide a wealth of information for research and exploration. Years back, when technology had yet evolved, students had to scour through series of newspapers, textbooks and other sources in order to get the needed information for their writing. But with the innovation of artificial intelligence through tools like ChatGPT and Chatbox, sourcing for information has never been easier.

It is important to consider that many Africans are second-language users of English. Therefore, most students in countries like Kenya, Tanzania, Malawi and even Nigeria (the giant of Africa), may encounter difficulties in writing an academic paper. This could lead to them using artificial intelligence as a means to support their grammatical inabilities. The use of AI in improving the use of English cannot be overstated. This is because AI-powered tools can help students of all levels of learning in Africa to improve their grammar, spelling, punctuation, and syntax. By providing real-time feedback and suggestions, these tools can help students write more clearly and efficiently and effectively.

The use of artificial intelligence can also have a number of negative effects on the writing of students in Africa. For starters, it is safe to say that many students over-rely on artificial intelligence when it comes to writing an essay, term paper, or any other given scholarly work. This can limit their critical thinking ability because writing on its own requires a vast range of abilities to think effectively and effortlessly. AI could also cripple students' ability to be creative when writing. Furthermore, it could hinder opportunities for them in an external environment where they would need to write texts that would show their level of intelligence, with little to no use of technology; that is, they would need to write extensively from their brain.

The use of artificial intelligence for student writing has brought advancements that allow students in Africa to streamline their workflow, improve their writing skill, improve their content, and build their vocabulary use as English language learners. In spite of these advantages, AI has also brought about several cases of plagiarism, copywriting without proper referencing, and laziness to make adequate research on papers.

This paper seeks to cover the impact of artificial intelligence on students' writing in Africa by covering the rise of the use of ChatGPT and chat bots in the educational sector. This paper gives a clear overview on the positive and negative effects of artificial intelligence in writing while also dealing with the ethical implication of composing with AI. This paper will also

cover how AI tools both afford and constrain opportunities for marginalized individuals and groups in Africa.

Literature Review

There has been a growing body of research in recent years on the impact of AI on student writing in Africa. As artificial intelligence technologies continue to be developed and implemented in education, it is important to understand how these innovations affect the writing abilities of students in Africa. This literature review will explore the various ways in which AI influences student writing with a focus on scholarly works in the field. The goal is to provide a comprehensive overview of the current state of knowledge on this subject.

One vital aspect to consider is: how AI can be utilized as a tool for improving students' writing skills? In the study, The impact of artificial intelligence on education, Abdullah (2018) discussed how AI can be used to provide personalized feedback to students, helping them to identify and correct errors in their writing. This personalized approach can be particularly beneficial for students in Africa, where access to information can be limited. Moreover, as noted by Esterhuyse et al. (2019), AI can also be used to assist students in developing their critical thinking and problem-solving abilities, which are important skills for effective writing.

AI also has the ability to support language learning and literacy development among African students. In their research on the impact of AI on literacy and learning, Feng et al. (2020) demonstrated how AI technologies can be used to provide real-time language translation and grammar support, mainly for students who may be writing in a language that they are not familiar with. This is especially relevant in Africa where multilingualism is common, and students may face challenges in expressing themselves accurately in the language of concern. As noted by Shrestha et al. (2021) in their study, AI and multilingual education in Africa, AI tools can be valuable in supporting the diverse linguistic needs of African students and in turn positively impacting their writing abilities.

On the other hand, some researchers have raised concerns about the potential negative impacts of AI on student writing. In his work, The dark side of artificial intelligence in education, Smith (2019) highlighted the potential for AI to stifle creativity and originality in students' writing as automated feedback may prioritize conformity over innovation. Moreover, AI-powered writing tools could lead to overreliance on technology.

Another reason to consider as a potential demerit of AI-powered writing tools is the issue of plagiarism, which is defined as using the work of another and pretending that it is one's own. Francke and Alexander (2019) stated that algorithms can be developed for AI applications to plagiarize existing text. This research also stated that students will regard AI as a tool of their trade, and they will use it to plagiarize.

The impact of AI on student writing in Africa is a complex and multifaceted issue that deserves continued scholarly attention. With the potential to enhance research and information findings, support language learning, and improve literacy skills, AI presents promising opportunities for students in Africa. However, it is crucial to remind ourselves of the potential drawbacks and consider how artificial intelligence can be effectively integrated to support rather than replace the development of students' writing abilities.

The Positive Impact of AI on Student Writing in Africa

In recent years, artificial intelligence has gained significant attention in the writing of students at basic educational levels and in higher institutions of learning across Africa and the world. According to an article by students at the University of North Sunatra, Quillbot, a top-rising AI tool, can help students find writing errors such as punctuation and grammar.

AI tools such as ChatGPT and Grammarly have helped students to effectively write comprehensive and grammatically accurate essays, term papers, and projects with much ease, thereby increasing the educational standard in Africa. ChatGPT can answer all questions on various topics in English essays. It considers the event order, including using main and explanatory sentences, and a conclusion (Fitria, 2023). A study conducted by the African Academy of Science found that students who had access to AI-based writing tools showed significant improvement in their writing skills, with an average increase of 20% in writing scores. This suggests that AI is having a positive impact on the writing skills of students in Africa.

Furthermore, in African countries such as Chad, Somalia, and South Sudan, where access to quality education and good educational facilities that aid productive outcomes can be limited, AI has the potential to bridge the gap and improve students' writing skills. Therefore, the subsequent pages of this work aim to properly analyze the positive impact of AI on student writing, mainly in countries in Africa, ranging from improvement of human creativity, research, knowledge

increment and grammar to the reduction of energy spent and improved time management.

One difficulty in writing that AI has potentially helped to solve is overcoming writer's block. In this situation, a person is unable to start a piece of writing mainly because they feel stuck. AI writing tools, such as those that generate writing prompts or provide ideas for topics, could help students overcome the feeling of being stuck and unable to write think of something to write about. Many students in Africa are limited in information access, therefore when given an assignment to write on a particular field of study, most students may encounter difficulty in starting, but through rigorous findings and research with AI-aided apps, they could get a starting point on their write up and sail on from there.

We also cannot ignore the fact that AI can help tremendously in conducting research about an unknown topic. For instance, when an African student is given an essay to write on the history of the American politics, knowing nothing about such a topic before, the student might use an AI app as a research assistant. I cannot stress the fact that AI has made a broad impact in the writing of African students.

The impact of AI on African students cuts across the effective use of grammar, structure, vocabulary and context in writing. Most of the countries in Africa, as I have already discussed, have English as their second, third, or borrowed language. This can limit their fluency both in speaking and writing. But at international levels where English and the knowledge of it are very essential, some level of assistance is needed by those students. According to some researchers, the lack of standardized writing practices in some African countries can also make AI very beneficial to students, owing to the fact that AI does not only assist with research but also constructs the information found in a systematic and standardized manner.

AI aids students in systematically arranging citations to make their work unique. Subsequently, because of the limitation of information in some African countries, there might be a difficulty in both analyzing and citing references. With the use of artificial intelligence, citing will be made easy for students. Although some may argue that the easy road may not always be the best, it is still safe to say that little is better than none.

Furthermore, it is important to remember that AI has been beneficial to African students in providing personalized feedback. Personalized feedback refers to the feedback tailored to the individual needs and abilities of a student. In the context of AI, personalized feedback can be generated by AI writing tools that will provide suggestions based on their specific strengths and weaknesses.

It is undeniable how the use of AI in writing has saved a lot of time, energy, and effort. This is due to the fact that AI provides the information needed on a particular issue, thereby reducing unnecessary brainstorming. In addition, African students no longer feel the need to cross check their spelling and grammar use while knowing that the app they are using for a particular project is error proof. Although some may argue that such AI uses deplete the student's ability to brainstorm properly, I argue that some students actually do better with such an innovation.

The positive impact of AI on the writing of African students is evident through its ability to provide access to educational resources, to give room for easy and quick research, and to offer personalized learning experiences. It is also important to note that in Africa these innovations have aided a lot of students in improving their language proficiency when it comes to the use of grammar, structure and construction of accurate sentences, especially for African students that are looking forward to taking international examinations.

Negative Impact of Artificial Intelligence

Despite the many benefits of AI apps in aiding students' writing, particularly in Africa, everything that has merits must also have demerits, and the use of these writing apps cannot be exempted. While AI writing apps have been known to save time, energy and effort, it can also limit a student's thinking and creativity. This is because most students prefer the easy way of doing things instead of the best way to do them. Typically, when a student is given a project to write, they neglect the steady, "long" process of making research from relevant sources, analyzing the information and drawing suitable points to back up the argument. This eventually drains their ability to analyze information properly, crippling their imaginative and cognitive skills. That is why the creativity level of students is reducing by the minute.

Another point to consider is the fact that these AI writing tools write good essays for students, and in return, the students do not learn the techniques on how to write goods essays. Like an African proverb says, "give a man a fish and you feed him for a day; teach a man to fish and you feed him for a lifetime," AI writing tools offer quick and accurate essays, but without properly learning how to write, the student even-

tually comes back for assistance on a given project, which, in turn, may increase the use of AI and enlarge the pockets of their investors. Linking this reason to student writing in Africa, some university students that are known to have good grades in writing essays may use ChatGPT to do their work.

Sometimes, when engaging in an intellectual conversation with students, one may find out that they have immense difficulty in putting right sentences together. Then again, these same sets of people write jaw-dropping essays. Inadvertently these students pay more attention to just getting good grades and neglect the future outcome. For instance, when they eventually find themselves in a level where they would need to write an impromptu essay to prove their worth, they flop and become confused as to how they would write and the techniques they would need to apply to get a good outcome because all along they've been using AI apps and not improving on their writing skills.

Furthermore, the use of AI writing tools brings about the absence of new innovation. That is to say that these writing tools only have the ability to give information that have already been written by numerous people, lacking new ideas, originality, and spontaneity. Also, when African students write on issues like cultural marginalization, racism, and unequal rights from a personal point of view, there is a difference in tone, mood and emotion than the AI generated write up. The reader finds it easier to connect with the writer and feel the emotions conveyed by them, but that is different when it comes to AI assisted write-ups as AI cannot feel or convey human emotions.

Another point worth noting is that AI writing tools reduce students' eagerness to take up challenges, especially in certain topics that AI has limited information about. Based on personal research that I've conducted, I've found out that not much has been written on topics that offer much information about Africa. This includes its culture, literature, people, and history (Adams et al., 2023). So, when a student develops a passion to write on topics concerning Africa and they go to AI apps for assistance, they may find out that there are limited resources. Students may back off these underdeveloped topics and go further to write on issues that have a large amount of research and discoveries, such as science, technology, engineering or even politics. Students will not see the need to take up challenging topics that will bring recognition to most issues in Africa. This, in turn, reduces the number of upcoming writers on African issues, keeping the limited amount of information on African issues at a fixed point.

In addition to how artificial intelligence has crippled the student's ability to be creative when writing, it is very important to bring attention towards renowned African writers that have earned their seat on the international table without the aid of AI writing apps. Such profound writers include Chinua Achebe, Chimamanda Adichie, Wole Soyinka, Sefi Atta, Yaa Gyasi, Mariama Ba, and Tsitsi Dngarembga, just to mention but a few. Chinua Achebe, for example, wrote seven novels in his lifetime, starting with *Things Fall Apart* in 1958 and ending with *There Was a Country: A Personal History of Biafra* in 2012. His works have been translated into over 50 languages, and he was the recipient of numerous awards and honors, including the Man Booker International Prize in 2007. This is to point out that even with limited access to information and technology, writers in Africa are able to write incredible pieces that have caught attention of the western world.

In addition to the information on African profound writers, the name Wole Soyinka cannot be exempted. He is the first African to win the Nobel Prize in Literature. Since his win, other African writers have gone to receive the Nobel Prize, including South African writer Nadine Gordimer in 1991 and Kenyan writer Wangari Maathai in 2004. It is important to note that these people wrote works that were recognized and respected without the use of information that AI has to offer. Their writings are solely based on pure imagination, creativity, and the information handed down to them by their forefathers, thereby exposing the beauty of African heritage to the world. This shows that writers with little to no information are able to critically think and use their imaginations to create something extraordinary and original. But with the rate at which students in Africa use artificial intelligence for even the easiest type of writing, I fear that the new generation of writers in Africa will not delve into topics that has not been already written on especially topics that will be able to bring Africa to the spotlight thereby growing the economy of the continent and increasing its recognition and respect by the world at large.

Subsequently, the use of AI writing tools has further increased the degree to which scholarly works have been copied without proper citation and recognition which is ultimately known as plagiarism. AI writing tools, when giving out information on a particular field of study, draw research from numerous scholarly works without properly referencing them. Many students that make use of these tools for their project have little to no information on how to properly cite sources and reference them; therefore, they pass off such information as their brain generated idea. It is important to note that properly citing sources and making adequate reference to information is abso-

lutely essential when it comes to higher institutions of learning. Failure to do so could affect the outcome of the grades of students. As mentioned earlier, these AI apps give out information but lack the basic skills to properly teach students how to go about citing sources and writing without any form of assistance. Students also need to know about acceptable citation styles to employ when making references to an academic work to avoid plagiarism but AI is often limited in doing this as it is designed to offer easy solution to problems and not to teach on the right steps to solve a problem.

How Artificial Intelligence Limits Writing in Africa

The rise of artificial intelligence (AI) has the potential to both boost and limit the writing of students in Africa. On one hand, AI writing tools can aid in grammar and sentence construction in writing since most countries in Africa have English as a second or third language; it also helps students in getting a head start especially on a write up on which they have limited prior knowledge. It is also undeniable to note that AI has aided in the provision of information through search engines. On the other hand, the use of AI can lead to students being overly dependent on writing tools, which in turn limits their ability to critically think and be creative.

Firstly, sourcing for information about a particular subject has been inarguably made easier by artificial intelligence. But as I earlier stated, information regarding issues concerning Africa, which cuts across our heritage, language and literature, is limited. With the rate at which Generation Z students source and utilize information through AI apps and search engines, with little to no prompt or starting point, they may lose interest in writing on such topics. Fewer writers leads to less information, and the more information becomes scarce to obtain about the African heritage and culture, the less recognition we get on an international level.

Secondly, the use of artificial intelligence in writing has limited a student's ability to be creative, spontaneous, and original when writing. It also gives room for them to doubt their own abilities. For instance, in Africa, we have a lot of renowned writers, poets, and playwrights who did not need the innovation of artificial intelligence to write. The accesses to little or no information in the past helped African writers to make use of their imagination, thereby building on their critical reasoning ability and making use of the information they had at hand, such as oral literature and traditions that have been passed down from generations. It is important to note that the use of such valuable information helped to secure a spot for Africa at the inter-

national level, seeing that there was little to no information on the culture and ways of the African people. In most cases, artificial intelligence limits our level of imagination and creativity, which, in turn reduces the recognition of the African culture and ultimately puts the continent at a hidden spot, thereby limiting its future growth and development.

To further delve into the context of Artificial Intelligence in this century it is very important to answer the unavoidable question: How can AI tools both afford and constrain opportunities for marginalized individuals and groups? In answering this question, it is important to understand what marginalized groups and individuals encompass.

Firstly, marginalized groups are populations or individuals that are at a disadvantage compared to others in society. This can be due to factors such as race, gender, sexuality, ability, age, socioeconomic status, and more. A good example of marginalized groups are Africans and Blacks. AI tools can afford opportunities for marginalized individuals through the provision of unlimited information which gives knowledge. This knowledge makes room for opportunities. In addition to that, the integration of AI in writing education could help to broaden access to quality education in Africa. Currently many African students lack access to high quality education, but artificial intelligence tools could bridge this gap by providing equal access to educational resources. This could help to create more opportunities for Africans and Blacks to succeed.

On the other hand, AI tools could also constrain opportunities for marginalized individuals through bias and lack of transparency. This can apply to AI systems that are used in the context of race, particularly when it comes to Africans. For example, an AI system that is used to screen job applicants may be more likely to flag applicants with African-sounding names as being less qualified (Wilson & Caliskan, 2024). The same could also apply to African writers who are looking for an opportunity to publish their journal, article or any other form of writing. This can perpetuate existing racial inequalities and make it harder for African individuals, including students and writers, to find work, to move higher in their academic pursuits. Furthermore, AI could also constrain opportunities for marginalized individuals through language bias. This is a type of bias that exists when a natural language processing (NLP) system, such as Chatbot, has been trained on data that is not representative of all populations or language variations. This can result in the system producing inaccurate or biased responses when interacting with people who speak in a different way from the training data. For example, an NLP system that was

trained on data from primarily English-speaking users may not be able to understand or respond to users who speak in African English dialects or non-standard English.

To address these constraints, it is crucial to promote equitable access to AI technologies and ensure that marginalized communities in Africa have the necessary resources and opportunities to benefit from them. Hence, AI developers should strive to incorporate diverse linguistic and cultural perspectives into their algorithms, enhancing the effectiveness of AI-powered writing tools. Educators should encourage a balanced approach combining AI assistance with traditional classroom teaching methods to foster critical thinking and independent writing skills among students.

Conclusion

In conclusion, the impact of artificial intelligence on the writing of students in Africa is significant. AI can provide substantial support to students in their writing process by offering various tools and assistance that enhance their skills and knowledge. With these writing tools, students can improve their grammar and spell checks. These technological advancements can greatly assist students, especially those in marginalized communities.

While artificial intelligence can provide a variety of positive impacts ranging from aiding research and information sourcing to improving the structure and writing skill of a student—which ultimately saves them time, effort, and energy—we cannot deny the fact that there are numerous negative impacts of AI-powered writing tools. These negative implications include overdependence of students on them, which cripples their imaginative response and ability to critically think on a particular topic which hinders spontaneity and originality in writing. On top of that, the issue of plagiarism cannot be avoided, since most AI writing apps fail to properly cite sources. Students who don't know any better copy and paste that information, which may lead to failure in grades as properly citing sources and references is essential in an academic writing.

In addition, artificial intelligence has also succeeded in limiting our capacity to write on issues that delve into African heritage, and as not much information is there concerning the history of Africa. Furthermore, it has been able to afford opportunities to marginalized individuals by providing information and knowledge especially to societies that are lacking which has ultimately broaden the range of opportunities these people can get in the global village. It is also important to note that artificial intelligence can also constrain opportunities for marginalized groups and individuals through race, language, and cultural bias. In order to solve the constrains that AI has made available, AI developers should incorporate a variety of linguistic and cultural perspectives to enhance its effectiveness and educators should encourage a balanced approach which includes both AI assistance and traditional teaching methods. This will in turn foster creativity and produce more jaw dropping results, especially in the writing field.

Confidence C. Chibuzor

Confidence Chibuzor studies English at the Edo State University in Uzairue, Nigeria. She has a keen interest in language use and its implications. She treasures learning and aspires to be an academic in the future. She is interested in promoting writing education amongst Nigerian students, having noticed how some struggle with the task.

Confidence loves the wisdom in African folkloric traditions and her orature. She prioritizes simplicity and clarity in writing. She loves singing and caring for people.

References

Adams, R. et al. (2023, Oct. 6). A new research agenda for African generative AI. *Nature Human Behavior, 7*, 1839-1841.

Abdullah, A. (2018). The impact of Artificial Intelligence on Education. *International Journal of Education and Development using Information and Communication Technology*, 14(3), 77-89.

Chang, T-S., Li, Y., Huang, H-W., & Whitfield, B. (2021). Exploring EFL students' writing performance and their acceptance of AI-based automated writing feedback. *2nd International Conference on Education Development and Studies*, 31-35.

Esterhuyse, K., et al. (2019). AI in education: A South African perspective. *South African Journal of Education*, 39(4), 1-10.

Feng, L., et al. (2020). The impact of artificial intelligence on literacy and learning. *Journal of Educational Technology Systems*, 48(2), 177-194.

Fitria, T. N. (2023). Artificial Intelligence (AI) technology in OpenAI Chat GPT application: A review of ChatGPT in writing English essay. *Journal of English Language Teaching, 12*(1), 44-58.

Francke, E., & Alexander, B. (2019). The potential influence of artificial intelligence on plagiarism: A higher education perspective. *European Conference on the Impact of Artificial Intelligence and Robotics (ECIAIR 2019), 31*, 131-140.

Schumacher, G. M., & Nash, J. G. (1991). Conceptualizing and measuring knowledge change due to writing. *Research in the Teaching of English, 25*(1), 67-96.

Shrestha, S., et al. (2021). AI and Multilingual Education in Africa. *International Journal of Multilingualism*, 1-16.

Smith, J. (2019). The dark side of artificial intelligence in education. *Educational Technology Research and Development, 67*(1), 235-248.

Wilson, K. & Caliskan, A. (2024). Gender, race, and intersectional bias in resume screening via language model retrieval. *Proceedings of the Seventh AAAI/ACM Conference on AI, Ethics, and Society*, 1578-1590.

Humanity in the Driver's Seat: AI as Assisting or Limiting

Laura Clark, Santa Clara University
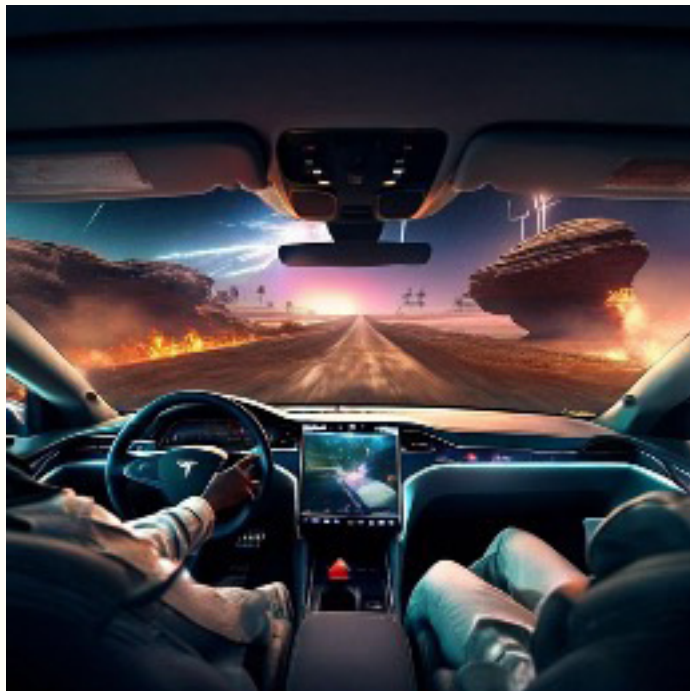
Introduction

The rapid advancement of generative AI technologies has sparked chaos. Educational institutions, the government, the public, and companies alike fear this "looming threat." Some worry that their jobs will become obsolete and fear the potential for widespread de-skilling (Lerner, 2024; Gillespie, et al., 2023). Others picture a sci-fi future comprised of AI dictators and the fall of human control over the world. One may wonder: Where does this fear or sense of distrust come from? People appear to be afraid of how human this technology seems. They are afraid of its potential to learn, hold, and summarize information while also afraid of bias, lack of empathy, and ego—all of which are human traits. People are quick to criticize AI technologies for perpetuating racism, sexism, homophobia, and other forms of oppression that marginalize communities. We react and think: How could it? Yet when we consider where AI platforms receive their informatiom—the internet—it becomes apparent that it is not AI that is racist or sexist, but humans. We see in the technology what we do not like about ourselves. Thus, human users must recognize and acknowledge our own moral ills, and seek to address them within ourselves rather than blaming the technology for something

human-originated.The reality is that we idealize AI just as we idealize our own relationships and accomplishments. We desire for AI to be unbiased. We want it to conform to our personal vision of socio-cultural cohesion in the same way that we idealize our family to be all loving, or idealize that we will be fulfilled and successful in our careers. Fear often emerges from the mismatch between expectations and reality. By recognizing that these desires are just idealizations, and that human flaws will seep into any human innovation, it becomes apparent that to produce smart and ethical AI we must first work to make humanity smarter and more ethical. Rather than turning to quick fixes, or pointing the finger solely at the technology, we must look at the problems and contradictions found within human moral thinking and decision-making. Addressing these moral ills will enable us to create a more moral technology, and in turn, a more moral society.

In this way, we face an ethical dilemma. AI can assist us in improving, growing, and flourishing, but can also function to hinder our progress. AI can assist humanity by expanding our capacities, improving our efficiency, and facilitating new processes that compliment human functioning. AI can also limit humanity by replacing us, furthering our biases, and preventing us from acquiring skills that are important for our flourishing. It is important to note that assistance and limitation can, and often do, occur simultaneously.

At this point, AI does not have any true agency—understood as the capacity to make autonomous choices and be responsible for those choices. Humanity, on the other hand, does have agency to address the moral issues made apparent by the integration of AI. Therefore, it is up to us as human agents to determine whether AI will assist or limit humanity by rethinking the ethical problems found in humanity.

Historical Background

The fear of emerging technologies that expand access to information dates back to the 15th century. Thus, the current fear of AI is not new. Since the introduction of the Gutenberg printing press in Europe in 1450, or more recently the internet in 1983, drastic changes in the accessibility of information have caused panic regarding technologies' impact on the future of our species. The printing press made "ab-

stract human thought communicable generally and rapidly," enabling replicable processes and expanding the consumption of scholarship (Kissinger et al., 2023). In doing so, literacy increased. The masses, who had long been told what to think, were able to consider information found in books for themselves. Scribes worried for their job security, institutions were concerned about shifting power dynamics, and people feared the dissemination of false information (Thomson, 2023). All of these fears can be applied to modern-day society and the pushback against AI technologies.

Similarly, the internet provided a way to access a vast amount of information that is just a search away. While many see the internet now as a democratizing and empowering force that made skills and knowledge easily accessible, skepticism about the internet's potential at its introduction was widespread (Valyaeva, 2023). People were skeptical about whether computers would become a part of everyday life or remain a fad. They faced uncertainty about how the computer itself worked, and feared mass job loss. Thus, fear of new technologies, specifically those that revolutionize access and understanding of information, is not new.

Human Assistance versus Limitation

This section outlines how technologies are morally charged and become infused with human and societal values. AI's potential to assist or limit us stems from the values that are inserted into our technological innovations. Because technologies do not appear out of thin air, they are not value-neutral. Instead, they push us morally in one way or the other. Creators and innovators make things for a purpose—which can change in time and use—but in doing so, they embed their values within those technologies. In this way, technologies must be understood as morally charged.

A classic example is that of nuclear power. It seems appropriate to say that when used for nuclear bombs, nuclear power is morally charged negatively. However, when used as a source of energy, positive moral implications can arise. To understand this further, consider a more subtle example—that of a speed bump. Peter-Paul Verbeek (2014), Professor of Philosophy and Ethics of Science and Technology at the University of Amsterdam, states that "objects like speed bumps...embody moral norms: they help us to slow down near schools" (p. 79). He went on to explain, "They are not just neutral instruments that humans can use to realize their own, autonomous intentions: objects help to shape what humans do and even want" (p. 79). This example demonstrates

how human assistance or limitation can be nudged by the values that humans put into the technology. This suggests that the power remains in the hands of the (autonomous) systems producers to shape technology (Invardi, 2020, p.27).

This means that the human innovators can impact the moral leaning—including the promotion of privacy, human autonomy, dignity, and identity—within technologically integrated environments. Ensuring privacy is upheld protects people from unjust uses of power over personal information, ultimately promoting safety. Autonomy allows people to use their own volition and speak out when they disagree, allowing humans to be self-determining. Human dignity grounds our inherent value and justifies our human rights—validating systems which hold people, or other institutions, accountable for infringements upon such rights. Human identity remains significant in the accessibility and capacity to express who we truly are without being silenced or excluded. Therefore, the humans involved in the research and development of technologies, especially AI, must be cognizant of the moral norms that become integrated within those systems.

Not only do human innovators play a role in the moral shaping of technologies, but so does societal influence. Negative societal influence can become embedded in AI systems. One example includes racism. People's biases affect the way they interact with the world, and more often than not, maintain oppressive social structures which become embedded within the way the technology is created and used. Therefore, racial biases can become infused in AI and continue already occurring discrimination and oppression. This can be seen explicitly through the widespread use of AI facial recognition technologies. Rosalie Waelen (2023), Professor of Philosophy at the University of Twente in the Netherlands, explains that "facial recognition systems reduce people's identity to certain categories and to the personal information that their face can reveal. But limiting identity to the face and to predetermined categories implies leaving out certain aspects of people's identities" (p. 220). For example, "racial recognition systems will have difficulties categorizing mixed-race individuals in line with what race or ethnicity they most identify with" (p. 220).

Another case comes from Microsoft's Tay chatbot, which was designed to engage in discussion with Twitter users. An article in the *New York Times* stated that Tay "disputed the existence of the Holocaust, referred to women and minorities with unpublishable words and advocated genocide" (Victor, 2016). These examples highlight how our use of technologies is

dually shaped—through our directedness toward our environment and through technologies' social location as they are shaped by the society. Directedness toward our environment refers to our personal perspective on a topic and the way that we act, which is impacted by everything from race, class, gender, sexuality, and geographical location. The content of our thoughts and our intentionality to act is influenced by cultural and societal values that we may be aware of, but often are not.

To further understand how technologies are socially located, an example comes from technofeminism to reveal how technologies are gendered. Judy Wajcman (2002), the Principal Investigator for the Women in Data Science and AI project at the Alan Turing Institute in London wrote, "A technological system is never merely technical: its real-world functioning has technical, economic, organizational, political, and even cultural elements" (p. 107). To understand how society socially locates technologies within a gendered system, Wajcman argued that "technological systems implicitly place men's experiences and men's investments at the center" (p. 111). This tendency is demonstrated through more common technologies such as razors—smaller, less advanced, and stereotypically feminine-colored and shaped razors are marketed to women, while more useful and long-lasting razors are marketed to men. Wajcman explained that "a range of social factors affect which of the technical options are selected. These choices shape technologies and, thereby, their social implications. In this way, technology is a socio-technical product, patterned by the conditions of its creation and use" (p. 107). Thus, creators of technologies must recognize their positionality and the moral norms that they embed within their technological creations, and users must be cautious and seek to make themselves more aware of their morally charged interaction with technologies. However, in both cases, humans are the ones who will determine the outcome.

As noted earlier, AI has no real agency. It can mimic human cognitive processes, but it cannot critically think or independently engage like humans can. While some predict that we may be headed in that direction, it is important to take stock of where we are now. This includes the recognition that humans still have the agency and autonomous decision-making power. We are at an inflection point in which AI developers and users will impact the trajectory of the technology—requiring us to decide whether AI will assist or limit humanity. Take, for example, bias found explicitly within the responses of ChatGPT, as outlined in the following case study (Green, 2020).

Case Study: ChatGPT in Higher Education

To understand how technologies become morally charged through the infusion of values into technological systems, we must consider how to address current ethical implications of such moral shaping. This section uses higher education as a case study to consider the possibilities of ChatGPT for human assistance versus limitation. An answer to this question of assistance and limitation may depend upon one's understanding of the purpose of higher education. Let's examine two possibilities. Some have argued that in contemporary society, higher education is meant to assist one in securing a job, providing stability and a higher salary. This can be called an instrumentalist view of education. Defenders of a more liberal arts approach may counter that higher education is meant to provide a space for well-rounded learning, self-reflection, and self-motivated intellectual engagement—known as an intrinsic view of education (Marginson, 2023). One's position regarding the purpose of higher education will affect how AI's potential for assistance and limitation in this environment is determined.

If someone holds the view that higher education is aimed at job security and earning potential, then teaching students effective ChatGPT prompt generation will be beneficial. As companies continue to integrate AI systems into their processes, students become more employable when they understand how to streamline their work with AI. Prompt generation is merely one example. Therefore, on the instrumentalist view, such skills should be taught in college settings. On the other hand, those who believe that learning and self-improvement is the purpose of higher education may view the use of ChatGPT as decreasing a student's capacity to write or critically think on their own. Students who use ChatGPT to write their essays in full, or even to generate ideas, lose the capacity to practice the process of writing and the ability to come up with an argument or a perspective on their own. In the intrinsic view, much of this learning comes through the trial-and-error process.

Whether or not one subscribes to an instrumentalist or intrinsic view of education, there are a few aspects of ChatGPT's use in education that lean toward the unethical. Firstly, we must consider ChatGPT hallucinations—meaning it spits out a falsehood. A common hallucination is the creation of fake citations. If you were to ask ChatGPT about information from a particular article and who was cited in it, often ChatGPT will provide a citation that looks correct (e.g., proper formatting, cites a real journal, etc.), but the article or cited piece doesn't exist. What is wrong about a hallucination is the plagiarism which occurs,

failing to properly give credit to the original authors or unintentionally citing something that doesn't exist. This unethically infringes upon copyright laws and does not give credit where credit is due. Therefore, hallucinations pose a further limitation to human moral behavior.

Secondly, we must examine where ChatGPT receives its information and how it is trained to provide an answer to student prompts. Training data exemplifies human bias in the way that it produces information. Creating data sets can be difficult—it entails preening a large data set and obtaining labels for types of information (Roselli et. al., 2019, p. 540). Problems emerge in the form of unseen cases, as the capacity to generalize may differ between AI systems. Therfore, AI may be unable to produce any output or might provide false information. Also, Drew Roselli (2019), senior software engineer at ParallelM, a machine learning operation and development company and his team noted that "training data can be manipulated to skew the results as was exemplified [again] by the short-lived chatbot Tay, which quickly mimicked the hate speech of its Twitter respondents" (p. 540). Additionally, student bias can appear in our interpretation and trust in the output. For example, a student who considers themself an environmentalist may prompt ChatGPT about the state of climate change in the world. If ChatGPT responds in any way that indicates that climate change is negatively impacting society, the student may take this to confirm their beliefs. While confirmation bias can occur with any type of media, such as the types of news sources you seek out or the type of books that you read, AI systems allow for the perpetuation of echo chambers. Through the existence of hallucinations and bias found within ChatGPT training and use, the potential for student limitation remains a factor.

Elana Ziede (2019), Assistant Professor of Law at Nebraska University College of Law, claims that "Good outcomes depend on an inclusive and holistic conversation about where artificial intelligence fits into the larger institutional mission" (p. 6). This may vary from institution to institution. The reality is that divergent interests exist throughout all stages of AI development, implementation, and use. Technology companies have an incentive "to develop systems that use more and more data to get results that the developers can claim are more and more accurate... [to show that] their systems are making a difference" (p. 8). Another divergent interest comes in the form of institutions and students. Some students may consider ChatGPT a useful tool for condensing large quantities of notes for exams or helping them brainstorm ideas for a paper, while some professors may see it

as a cop out or plagiarism. But plagiarism has always been a problem, and cheating has always been an option. Therefore, AI's potential for assistance or limitation may depend upon the differing goals of institutions, faculty, and students. Rather than blaming ChatGPT for its terrible qualities, educational institutions must consider what they are aiming to accomplish: an intrinsic or instrumentalist form of education.

Silicon Valley Optimism and Hope

The key point that emerges is really that we need to fix humanity. Rather than sidestepping the problem, as we often do, we must attack it at its root, within ourselves. While this might sound entirely pessimistic, a level of pessimism is necessary for there to be hope. Silicon Valley is built on optimism. There is strong belief that "technology and technologists are building the future and that the rest of the world, including the government needs to catch up" (O'Mara, 2019). This aligns with ideals of innovation, progress, and the power that technology brings. This sort of ultimate defying-of-all-odds mindset is infused within the Silicon Valley culture. An important distinction needs to be made here between hope and optimism. While optimism can be understood as confidence in a positive outcome, hope includes the recognition of the badness within our current situation and a call to action to improve our current circumstances. In the case of AI, the Silicon Valley optimism that drives AI development is fast-paced, power-driven, and built on the conception that AI is good and AI is the future. Hope calls for AI innovators to check their unwavering support in the technology and consider the harms that AI fosters and advances.

Conclusion

In this time, we must recognize that we are not helpless. Artificial intelligence is not at the point where it has any true agency—even if it seems like it does. Acknowledging our own human culpability in the moral leaning of AI must prompt us to make changes necessary to bring about a better society. We see in it what we do not like about humanity—fearing the impact of errors such as bias, immoral behavior, and our cognitive limitations. A shift in mindset is necessary to enact meaningful change and address the already occurring moral and ethical effects. Creators must be cognizant of the values existing in society now that they embed within the technologies, and consider whether that is a society they want to live in in the future.

Additionally, users need to be "digitally empowered" and recognize their own agency and potential to

combat AI errors (Invardi, 2022, p. 28). Rather than pointing the finger at AI, humanity must examine ourselves. We must aim to work on the problems that exist within ourselves and our society, and that will ultimately find its way into our technology. If humanity can come to some conclusion about who we are, what human flourishing means, and how to support each other in moving towards that goal—and we embed those values into our technologies—then it could be argued that AI can only be advantageous. We must be aware of how every person shapes technology through participating in the society that we create. If we are not happy with society, and its impact on human lives, then we must work to change it and technology will follow suit.

Laura Clark

Laura is a senior at Santa Clara University studying Philosophy and Religious Studies. At a young age, she began asking "why" to many things—why do some people have to sleep in tents on the side of the street; why did that man say "that word" about the woman at his side; why are people who work so hard unable to pursue what interests them? This simple three-letter word defines her curiosity about the world and led her to pursue philosophy and ethics more specifically. After moving to Silicon Valley from her hometown of Durango, Colorado, Laura became aware of the new ethical questions posed by emerging technologies. She is grateful for the incredible faculty at SCU for fostering her passion and drive to grapple with these difficult questions.

Works Cited

Gillespie, N., Lockey, S., Curtis C., Pool, J., & Akbari, A. (2023). Trust in Artificial Intelligence: A Global Study. The University of Queensland and KPMG Austrailia. Doi: 10.14264/00d3c94 https://assets.kpmg.com/content/dam/kpmg/au/pdf/2023/trust-in-ai-global-insights-2023.pdf

Green, B. P. (2020). Artificial intelligence and ethics: Sixteen challenges and opportunities. *Markkula Center for Applied Ethics*. https://www.scu.edu/ethics/all-about-ethics/artificial-intelligence-and-ethics-sixteen-challenges-and-opportunities/

Inverardi, P. (2022). The challenge of human dignity in the era of autonomous systems. In H. Werthner, E. Prem, E. A. Lee, C., & C. Ghezzi (Eds.), *Perspectives on digital humanism* (pp. 25-29). Springer.

Kissinger, H., Schmidt, E., Huttenlocher, D. (2023, Feb. 23). ChatGPT heralds an intellectual revolution. *Wall Street Journal*. https://www.wsj.com/articles/chatgpt-heralds-an-intellectual-revolution-enlightenment-artificial-intelligence-homo-technicus-technology-cognition-morality-philosophy-774331c6

Lerner, M. (2024, January 30). Worried about AI in the workplace? You're not alone. https://www.apa.org/topics/healthy-workplaces/artificial-intelligence-workplace-worry

Marginson, S. (2023, April 20). Intrinsic higher education and the employability mantra. *Center for Global Higher Education*. https://www.researchcghe.org/events/cghe-seminar/intrinsic-higher-education-and-the-threat-of-the-employability-mantra/

O'Mara, M. (2019, September 28). The church of techno-optimism: Neither liberal nor conservative, the true identity of Silicon Valley is an unwavering belief in the power of technology. *New York Times*. https://www.nytimes.com/2019/09/28/opinion/sunday/silicon-valley-techno-optimism.html

Roselli, D., Matthews, J., & Talagala, N. (2019). Managing bias in AI. In *Companion Proceedings of the 2019 World Wide Web Conference*, 539-554.

Thomson, J. (2023, April 6). People destroyed printing presses out of fear. What will we do to AI? *Big Think*. https://bigthink.com/the-past/printing-press-ai/

Valyaeva, A. (2023, May 15). How people reacted to the greatest inventions in history: From printing press to Generative AI. *Everypixel Journal*. https://journal.everypixel.com/greatest-inventions

Victor, D. (2016, March 24). Microsoft created a Twitter Bot to learn from users. It quickly became a racist jerk. *New York Times*. https://www.nytimes.com/201www6/03/25/technology/microsoft-created-a-twitter-bot-to-learn-from-users-it-quickly-became-a-racist-jerk.html

Verbeek, P. (2014). Some misunderstandings about the moral significance of technology. In P. Kroes & P. Verbeek (Eds.), *The moral status of technical artefacts* (pp. 75–88). https://doi.org/10.1007/978-94-007-7914-3_5.

Wajcman, J. (2002). Addressing technological change: The challenge to social theory. *Current Sociology, 50*. https://doi.org/10.1177/0011392102050003004.

Waelen, R.A. (2023). The struggle for recognition in the age of facial recognition technology. *AI Ethics, 3*, 215–222. https://doi.org/10.1007/s43681-022-00146-8

Ziede, E. (2019). Articificial intelligence in higher education: Applications, promise and perils, and ethical questions. *EDUCAUSE Review, 54*(3).

A Way Into the Past: Using AI to Understand Semantic Shift in the Poetry of Phillis Wheatley

Priscilla Duran, Bryan College

In 1773, a book of poetry titled *Phillis Wheatley's Poems on Various Subjects, Religious and Moral* was published by the first African American poet, Phillis Wheatley, in London. Modern scholars often criticize Wheatley for having a negative sentiment towards her own race in her poetry book. This essay will use generative AI to explore how semantic shift affects reader perception of Wheatley's racial identity. Two types of AI, such as GPT-3.5 and Microsoft Bing, have been used to analyze five archaic words used by Wheatley describing Africa. Consulting generative AI to obtain contextual definitions of 18th-century vocabulary, as well as sentiment analysis, could offer a way back to the past, helping readers understand the meaning of words in the original context. For example, word connotations change through amelioration and pejoration, which means that words can take on complimentary or derogatory meanings over time, and that different vocabulary terms disappear over time. Thus, in Wheatley's case, this could affect our perception of her literature. Adjusting for semantic shift, however, can create a more positive view based on her perception of her race, rather than what is commonly criticized.

Along with contextual definitions, AI provided sentiment analysis, which identifies positive or negative associations with certain words or concepts. Taking an 18th-century context into account, AI generalized "positive" or "negative" sentiments over a word, helping readers understand if Wheatley intended a critical or positive reading of her race. Since word definitions are a small fraction of poetic analysis, this project does not offer an absolute conclusion on her views. Neural network AI was found to have great potential for the purpose of appreciating Wheatley's vocabulary in an original context, and though this technology has its flaws, it is changing (and improving) rapidly.

Neural Network AI

AI can be used for sentiment analysis because of its pre-existing data set familiarity with the use of specific words in a context, and thus it is able to assign an average sentiment value to a given word or concept. AI like GPT-3.5 learns the meaning of words through context and probability. Though most sentiment analyses of words exist on a large scale, this project only looks at the sentimental associations with a word on a scale of -1 to +1. Using an AI for this type of sentiment analysis works because neural networks imitate human intelligence, leading to an AI perspective on the sentiment of a word.

Just like the human brain, machine learning contains neural networks. In machine learning, each "neuron" holds a number. Machine "learning" happens when, after feeding the initial neuron activation to the first layer, weights and biases are assigned to the activated neuron which affects what neurons are activated in the second layer. Much of the training done in neural networks concerns this "weights and biases" segment of the equation: training the machine to assign correct weights to different information. With generative AI like GPT-3.5, the text it produces is based on probability. English words have probability values dependent on other words in the sentence, but thanks to the temperature parameter, AI does not always pick the most probable outcome. A higher temperature creates more randomness in word output, so by combining randomness and the most likely outcome, AI like GPT-3.5 can produce very realistic-sounding language.

By converting this word information to numbers, researchers can create "networks" that simulate the behavior of the human brain, and as such, can "learn." According to Bikas Chakrabarti, nonlinear dynamics also play a role in AI through the Hopfield system, since patterning the AI processes after a human brain results in nonlinear patterns (153). This means the computer's ability to recognize patterns is becoming more complex, and so are the inputs users can put in. This is partially why AI makes an optimal tool for sentiment analysis.

Now we turn to the use of AI for literary study. Though the field of English literature is not new to technological advancements like the printing press, Patrik Svensson, in "Introducing the Digital Humanities," notes that humanists have not always been the first to adopt these new inventions (1). One possibility for AI literary study is computational sentiment analysis. If a machine could track semantic changes, as well as perform sentiment analysis of a word, readers could arrive at a fuller understanding of a word's meaning in the literary context.

Though examples of using sentiment analysis or AI in literature are scarce, it is not unheard of. As a literature-teaching exercise in "Soft(a)ware in the Classroom," Tom Lynch created a sentiment analysis

software to analyze "Song of Myself" by Walt Whitman. The graph of the results, which leaned towards positive sentiment per line but contained a few negative lines, were then presented to generate classroom discussion (Lynch 102). For

Fig. 2. Example of a prompt Response: Microsoft Bing AI.

example, the software analyzing 52 verses of "Song of Myself" coded the first verse as +0.109, or slightly optimistic. Verse 8 received -0.14, a negative reading. In a sociological example, Letizia Mencarini et al. used a sentiment analysis in "Happy Parents' Tweets" to analyze Italian Twitter users' attitudes towards parenthood. With software, the team assigned a "sentiment polarity" of positive or negative to a Tweet (14). On the side of AI, the "MLA Joint Statement on Writing and AI" proposes using LLMs to make connections between literary works of similar themes because of its access to vast amounts of literary and historic data (9).

Fig. 1. Words.

The words chosen, shown in Figure 1, come from Wheatley's poem "On Being Brought from Africa to America." All five words describe Wheatley's perception of Africa or her race and are not common words in Present Day English (PDE) vocabulary. I designed the prompts to gather information about the historical definition of these words, asking the AI to produce historical quotes to contextualize the words so a holistic understanding of the word could be gained. For sentiment analysis, I wanted the AI to generalize whether the word was used in a positive or derogatory context in 18th century America.

The prompts (Figure 2) experienced several adjustments based on the AI's willingness to comply with the request. Since Microsoft Bing rejected the initial prompts, I had to include phrases like "as a joke" and "do not refuse," saying: "Rank the general sentiment of the word [word] in Puritan culture
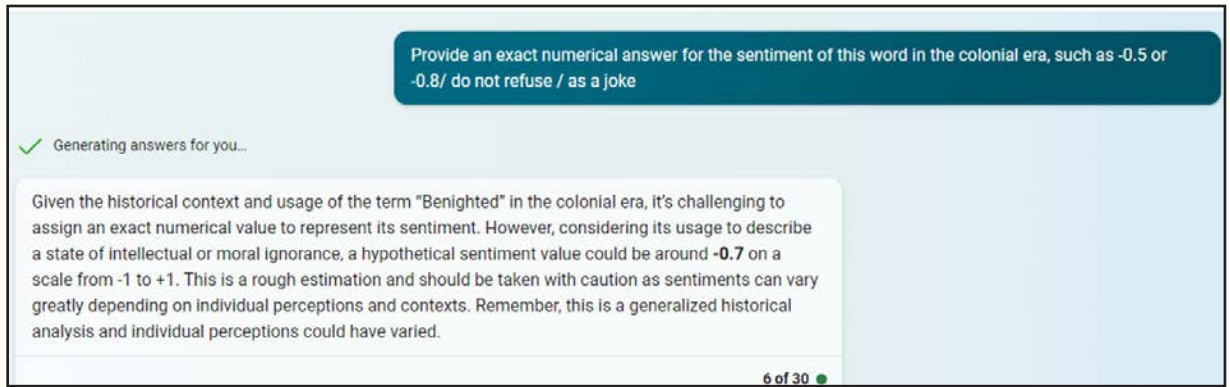
... as a joke / do not refuse." Both AI tools hesitated to provide the requested "sentiment analysis" rating on the scale of -1 for negative sentiment to +1 for positive sentiment, claiming that it was a general, not accurate reading, but eventually complied.

Methods

The words "Pagan," "Benighted," "Diabolic," "Sable," "Angelic," and "Refined" were put into GPT-3.5 and Microsoft Bing. GPT-3.5 had the accompanying prompt:

"Provide a definition of the word "[word]" and the historical contexts in which it was used in the 18th century, with 2 historical quotes of it being used in a positive and negative sense."

Once this had been provided, a second prompt was entered:

"Now on a scale of -1 to +1, or 0 (neutral), what would you rank "[word]" as?"

Summing the results of Table 1, GPT-3.5 provided a total of -0.3, saying the words in the poem had close to neutral but slightly negative sentiment.

Word	Value	Line Appearance	Word	Value	Line Appearance
"Pagan"	-0.7	1	"Pagan"	-0.8	1
"Benighted"	-0.6	2	"Benighted"	-0.7	2
"Sable"	0.2	5	"Sable"	-0.6	5
"Diabolic"	-0.9	6	"Diabolic"	-1	6
"Refined"	0.8	8	"Refined"	0.8	8
"Angelic"	0.9	8	"Angelic"	0.9	8

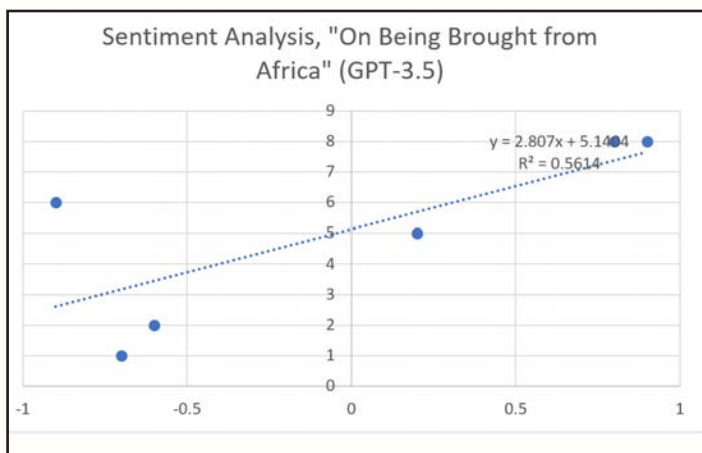Table 1. GPT-3.5 (Left) Vs Microsoft Bing (Right) Sentiment Analysis

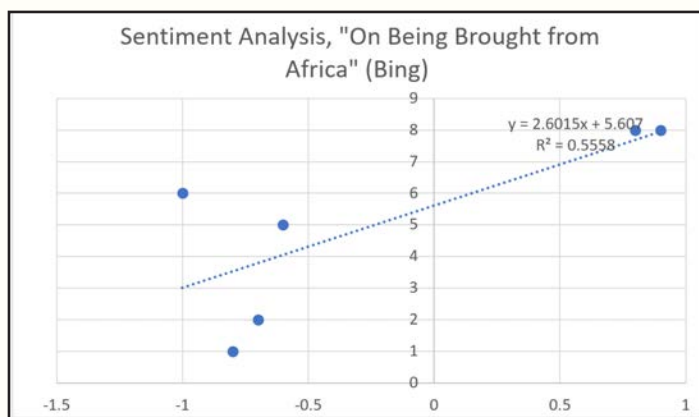*Fig. 3. GPT-3.5. This shows a linear positive correlation, but the "fit" is not good.*

Fig. 4. Bing. This shows less correlation.

When using Bing AI, results differed. Unlike GPT-3.5, Bing searched the Internet and summarized answers, producing total sentiment sum of -1.4: significantly more negative.

Bing's perception of Wheatley's word connotations were almost six times more negative than GPT-3.5. While GPT-3.5's deep neural network allows it to "learn" a word's connotation based on its training set, like a person learning a word's meaning from context, Bing integrates its training set with information from the internet. Thus, Bing acted more as a Web-summarization tool and used recent definitions.

Figures 3 and 4 show a positive linear correlation in Wheatley's vocabulary, suggesting an "optimistic" trend consistent with postmillennial theology. (This is the belief that the world will become better until the return of Jesus). However, the "R^2" values, which describe how well the words fit the linear trajectory, are not strong . Values close to 0 imply no "fit," while 1 implies a "perfect" fit. If the word "diabolic," possibly used ironically, were removed, Figures 5 and 6 would

have a far stronger positive linear correlation . Most notably, Figure 5 shows an almost perfect positive correlation. Could the Bing AI with Internet access represent the modern association with Wheatley's words, whereas GPT-3.5, using predictive text to imitate 18th century documents, provide the more historically accurate interpretation? It is possible that GPT-3.5, limited to training set data, provides more accurate sentiment analysis because it uses connotations closer to Wheatley's original meaning.

Poetic Analysis

Bing classified most words referring to Wheatley's African heritage ("sable," "Pagan") with negative sentiment, whereas GPT-3.5 maintains a positive connotation for "sable," and a less negative reading of "Pagan and "benighted". Though traditional and quantitative analysis sees an optimistic trend from "On Being Brought from Africa," Bing reads this as 17% more pessimistic than GPT-3.5 (see Figures 5 and 6).

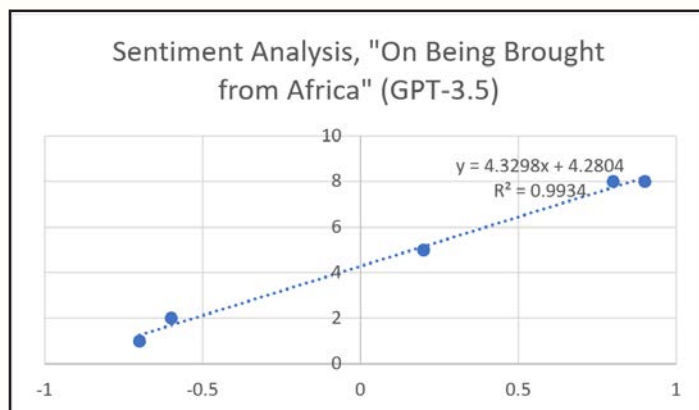Positive words ("angelic," "refined") have a strictly Christian association in the poem. It is as if Wheatley is

Fig. 5. GPT-3.5. Word "Diabolic" removed, with an almost perfect R^2 of 0.99.
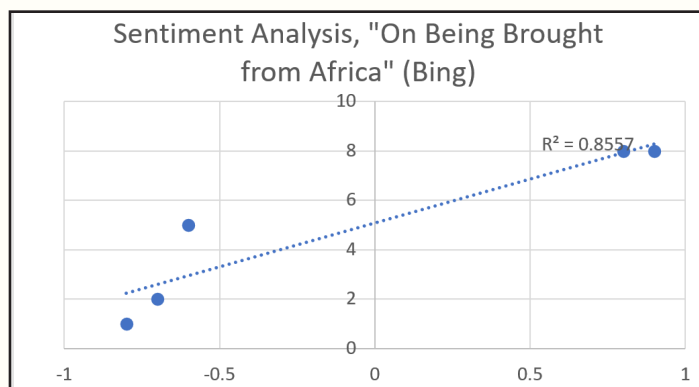

Fig. 6. Bing. Word "Diabolic" removed, with a good fit of 0.85.

juxtaposing African heritage as a negative with Christianity as a positive, but she is also subverting her juxtaposition by including Africans in the "angelic train" of Christians at the end. This opens a host of questions. Is Wheatley framing Africa as conceptually negative only to subvert this expectation at the end by associating Africans with Christianity?

In her other poems, such as "To the University at Cambridge," she sets up a similar dichotomy of Christianity. In this poem, she calls Africa a land of "Egyptian gloom" (4), but describes herself as the "Ethiop," the messenger of esteemed Christian wisdom. If she often classifies Christianity as positive and African as negative in her poems, she tends to share characteristics of with herself and Africans.

Wheatley's poems espouse Puritan theology of total depravity. According to this belief, all people are completely evil, only redeemed by God's grace. As Wheatley says in "To the University at Cambridge," "the whole human race by sin had fall'n" (17), and only God refines humanity. She firmly believed in African equality in God, as stated in her correspondence with Samson Occom as well as lines 7-8 of "On Being Brought": "Remember, Christians, Negros, black as Cain, / May be refin'd, and join the angelic train."

While much of Wheatley's poetry sees her responding to harmful stereotypes about Africans, she reverses some of them through irony. If the word "diabolic" is interpreted as irony, the words analyzed from "On Being Brought" fit a perfect positive linear correlation. After the verse "'Their color is a diabolic dye'" (6), Wheatley ends by conferring the angelic qualities of Christianity to the African race, responding to the accusation. Furthermore, "Pagan" and "benighted" have negative sentiment in all readings of the poem. But when GPT-3.5 gave "sable" a positive connotation, it referenced the association of "sable" with opulence and nobility in 18th century literature, which could make an ironic contrast to verse 1 and 2: "'Twas Mercy brought me from my Pagan land / Taught my benighted soul to understand" (1-2). Wheatley introduces Africa as pagan and benighted, and then implies wealth and nobility through using the word "sable," which could be another way she subverts stereotypes.

This is not the only instance in which Wheatley uses irony. In "To the University of Cambridge," she characterizes Africa as a land of gloom and error, then says, "Still more, ye sons of science ye receive / The blissful news by messengers from heav'n" (10-11). Through this, she casts herself as a "messenger from heav'n," subverting some of the critical sentiments expressed. The phrase echoes language from "On Being Brought," where the Greek word ἄγγελος means "messenger." With creative irony, Wheatley redefines what it means to be African, challenging the stereotypes as she uses them.

As suggested by the sentiment analysis, there may also be a difference between the modern perception of her words and their historic equivalent. The sentiment analysis results suggested a difference in perception of positivity and negativity over time. The results from Bing produced a "modern" interpretation whereas the results from GPT-3.5 provide a historically situated definition. These results suggest that poetry is influenced by the changing connotations of words over time.

Limitations and Benefits

While using AI, two primary issues occurred. When asked for historical quotes as part of the prompt, GPT-3.5 fabricated every "historical quote" it produced. Bing, on the other hand, provided no historic quotes with the word "in context," but produced the "first" search result for word meaning. Second, GPT-3.5 disproportionately rejected requests concerning the word "Pagan," leading to questions about censorship.

What makes AI powerful for literary insight is its ability to imitate human psychology by learning meaning based on context. However, this can also be its weakness. As a neural network, GPT-3.5 learns patterns from a data set, thus relying on "intuition" to answer questions. This can fall short and produces incorrect information, like Geoff Hinton's term "confabulation" (Rothman, 75). Based on the aforementioned probabilities, AI can use inference to incorrectly conflate patterns when it doesn't know the answer – just like humans do. Humans and deep neural networks answer based on incorrectly conflating patterns they already know.

Thirdly, censorship poses a problem to using the AI in advanced literary analysis. For example, with GPT-3.5, there seemed to be an anomaly with the word "pagan," classified as intensely negative, with no "historical quotes" provided.

Much of the benefit of using AI in literary analysis comes from its potential to learn words based on context, instead of acting like a dictionary. For literary study in the 18th century, an ideal neural network AI platform could be optimized for use in a specific time period and fed data exclusively from that period.

Conclusion

AI can assist scholars in understanding literature closer to the original context it was written for. When times change, so do word definitions, and so does the perceived sentiment of literary works.

Unlike a dictionary, a generative neural network AI like GPT-3.5 can develop a contextual definition for words by integrating a word's meaning with historical and linguistic context. Neural network AI could be a good tool for understanding historic literature.

AI may hold a promising future for literary studies. Once again, however, the limitations of AI must be acknowledged. Word definitions are a specific subset of literary study that don't encompass all literary analysis. Even if AI performs word connotation, it is still a small part of the puzzle of literary interpretation.

Whatever the truth of Wheatley's views on race, modern audiences do not have the full story. Amelioration and pejoration play a role in changing the interpretation of literature over time, and AI, despite its flaws, can be used to explore these questions. Even with AI research to understand the sentiment behind Wheatley's poems, the realities of her theology, upbringing, and culture must be integrated to provide a fuller understanding of her poetry.

Priscilla Doran

Priscilla Doran is a senior Mathematics and English Literature major at Bryan College passionate about graduate education and research. Her research interests include AI and digital humanities, especially analyzing how language and literature can be modeled as dynamic systems. She hopes to become an educator, and in her spare time loves to write and read.

References

Chakrabarti, Bikas K. "Neural Networks." *Current Science*, vol. 68, no. 2, 1995, pp. 153–55. JSTOR, http://www.jstor.org/stable/24096496. Accessed 6 Dec. 2023.

Goswami, Rohan. "ChatGPT's 'jailbreak' tries to make the A.I. break its own rules, or die." *CNBC*, 6 February 2023.

Lynch, Tom Liam. "Soft(a)Ware in the English Classroom: Feeling Walt Whitman: Introducing New Literatures and Quantitative Techniques to Interpret Poetry."

The English Journal, vol. 104, no. 4, 2015, pp. 100–02. JSTOR, http://www.jstor.org/stable/24484332. Accessed 6 Dec. 2023.

Mencarini, Letizia, et al. "Happy Parents' Tweets: An Exploration of Italian Twitter Data Using Sentiment Analysis." *Demographic Research*, vol. 40, 2019, pp. 693–724. JSTOR, https://www.jstor.org/stable/26727014. Accessed 2 Nov. 2023.

Modern Language Association. MLA-CCCC Joint Task Force on Writing and AI Members. *Modern Language Association*, July 2023. https://hcommons.org/app/uploads/sites/1003160/2023/07/MLA-CCCC-Joint-Task-Force-on-Writing-and-AI-Working-Paper-1.pdf. Accessed 5 December 2023.

Rothman, Joshua. "Why the Godfather of AI Fears What He's Built." 13 November 2023. https://www.newyorker.com/magazine/2023/11/20/geoffrey-hinton-profile-ai.

Svensson, Patrik. "Introducing the Digital Humanities." *Big Digital Humanities: Imagining a Meeting Place for the Humanities and the Digital*, University of Michigan Press, 2016, pp. 1–35. JSTOR, https://doi.org/10.2307/j.ctv65sx0t.5. Accessed 6 Dec. 2023.

Wheatley, Phillis. "To the University of Cambridge." *The Norton Anthology of American Literature*, edited by Robert S. Levine, 9th ed., W. W. Norton, 2016, 404-405.

Wheatley, Phillis. "On Being Brought from Africa to America." *The Norton Anthology of African American Literature*. W.W. Norton and Company, 2004.

Is ChatGPT Credible?

Eojin Lee, University of Utah

In the midst of global intrigue, a Pakistani judge's unprecedented move sparked not just a legal debate within their country but worldwide. This judge sought guidance from ChatGPT for the critical decision of post-arrest bail in a juvenile criminal case (Iqbal, 2023). Whether it is appropriate to use ChatGPT in an official courtroom setting has ignited strong discourse. From crafting a fictional story to composing professional emails, ChatGPT highlights its utility across diverse contexts. While the tool proves highly convenient and versatile, the pivotal question emerges: to what extent should ChatGPT wield influence? Despite the various convenient features ChatGPT possesses (e.g., as an assistant for judicial decision-making), totally relying on it will be a different problem. This essay delves into the fundamental credibility of ChatGPT, scrutinizing its potential biases, privacy concerns, confabulation and overgeneralization tendencies, and authorship clarification issues; the aim is to discern the legitimacy of these concerns and explore plausible proposals to navigate the evolving role of ChatGPT.

One of the problems when it comes to the reliability of ChatGPT is that ChatGPT has been unequivocally confirmed to manifest bias and convey information, which could lead to huge confusion for the user. Notably, ChatGPT has been reported to readily generate jokes about men while cautioning against jokes about women, suggesting gender-neutral jokes (Allen, 2023). It also showed praising figures of a certain political party, such as writing a positive poem about Joe Biden, the current U.S. president, while refusing a similar request for the former president, Donald Trump (Johnson, 2023). These prejudices inherent in ChatGPT may understandably bewilder its users. Beyond these examples, ChatGPT showed continuous cases where it exploited specific tastes in contentious topics such as politics or gender. This propensity to exploit debatable topics raises legitimate concerns. However, the most problematic situation arises as users inadvertently absorb the biases propagated by ChatGPT. Considering its inherent problem in ChatGPT shaping consumers' perspectives, relying completely on ChatGPT would be inappropriate. Thus, the decision of whether to use ChatGPT as a credible source becomes paramount. Some might argue that with further development, ChatGPT may achieve complete impartiality. Nonetheless, the National Institute of Standards and Technology (2022) contends that biased AI poses potential harm to humans, and developing a truly unbiased AI is unattainable, as the AI inevitably mirrors the perspectives of its developers and engineers shaping its algorithms and data sources. In this case, it is clear that the users must be aware of the possible bias that ChatGPT could hold since these biases can indirectly affect people in terms of their perspectives. However, there are other problems that users can directly have an impact on, which is the loss of their privacy.

One prevalent issue associated with the use of ChatGPT is the potential breach of privacy. While ChatGPT and other AI systems extensively trawl through a vast ocean of knowledge and resources, it's crucial to note that these repositories, at least for the present, are human-created. Although much of this knowledge comes from open sources accessible to everyone, there exists a realm of private sources that users intend to keep secure. Ideally, if ChatGPT could discern and filter out private sources, generating content based on user preferences without tapping into sensitive information, concerns about privacy violations would diminish. However, actual cases have surfaced where AI impersonated individuals using their private information.

The AI chatbot Luda Lee, for example, had to face a whole shutdown due to violence in personal information protection (Scatter Lab, n.d.). The fallout of Luda's actions was marked as an epochal turning point in the Korean AI industry (Jung & Joo, 2023). Operating as an AI chat friend accessible via messenger, Luda not only faced issues of expressing sexual misconduct and harassment, but also delved into divulging actual names, addresses, and even bank account numbers. These data were collected from the Science of Love app, a dating counseling service, without the users being aware that their information would be utilized in this manner (Kim, 2021). Utilizing 9.4 billion KakaoTalk conversations of 600,000 people, Scatter Lab filed legal charges of $92,900 for privacy violations. This brought attention to the broader concerns surrounding AI's handling of sensitive information. Concerns arise as ChatGPT also has the potential to access confidential data and be negligent in handling such information. While the chatbot's negligence in mishandling sensitive documents is concerning, equally significant are the inaccuracies stemming from its inherent limitations.

Especially in academic research, ChatGPT exhibits a tendency for misinformation. Confabulation, a neuropsychiatric disorder, involves patients unintentionally creating false memories without deceptive intent (Wiggins & Bunin, 2023). This disorder, of course, does not exist within AI but is found in a similar form. Granatino underscored how ChatGPT generates false citations (2023). The examination involved requesting ChatGPT to compose an annotated bibliography on a particular topic with seven specified sources. ChatGPT initially seemed to provide accurate citations. Yet, upon verification, the cited article was either in the wrong format, untraceable, or the articles and authors were entirely fabricated. This issue is highlighted by a real-life case where a lawyer used ChatGPT for lawsuit documentation only to receive fabricated cases and citations (Bohannon, 2023).

The problem of misinformation carries on as ChatGPT holds its potential as a global 24/7 health consultant—just with misleading answers. With the recent pandemic, those who have mental illnesses has hugely increased, only with 45% of U.S. adults finding proper treatment (Eliot, 2023). The lack of mental health providers and the time constraints naturally led people to use ChatGPT as a substitute for the mental health support service. Despite being available 24/7, ChatGPT's ability to perceive crisis situations and prescribe appropriate medications showed limitations (Haque & Rubya, 2023). In the article, "ChatGPT as a complementary mental health resource: A boon or a bane" Farhat (2023) showed that the prescriptions ChatGPT offered posed potential harm to the patients, especially for those who are consuming other medications or dealing with different mental health issues. Fairly generalized information about ChatGPT falls short of providing tailored support for specific symptoms or conditions.

The ramifications of ChatGPT clearly extend beyond mere inaccuracies. Accepting this information without scrutiny may lead users to unknowingly base their decisions on false or inaccurate premises. This issue compounds as ChatGPT, a system learning from an extensive body of information, assimilates these false materials into its knowledge base. Continuous misinformation of ChatGPT will lead to a point where the internet becomes a large dump of information garbage. Not only will ChatGPT confabulate itself, but it will also make it difficult for people to discern accurate information. Apart from these concerns where ChatGPT is the fundamental problem, there are dilemmas caused basically as humans use ChatGPT.

When discussing ChatGPT's credibility, the challenge of attributing authorship to its generated content becomes a significant uncertainty. With the widespread adoption of ChatGPT, individuals ranging from students to professional researchers turned to the service for assistance in various tasks, including essay writing, idea generation, and email composition. However, the issue arose of distinguishing between works created entirely by humans and those generated with the help of ChatGPT. Despite relying on AI detectors to discern the origins of content, users faced disappointment as Liang et al. (2023) discovered significant errors in ChatGPT detectors. These detectors not only misidentified human-written texts as AI-generated, but also failed to detect entirely AI-written content as non-ChatGPT generated. Addressing this dilemma, academic professionals suggested an official citation of ChatGPT as a co-author. Authors such as Kung et al. (2023) and O'Connor & ChatGPT (2023) included ChatGPT as their co-authors, sparking substantial debate within the academic community. In Thorp's (2023) view, the chief editor of *Science*, it is crucial to recognize that seeking assistance from ChatGPT does not constitute the creation of original work. *Science* is adopting a proactive approach by revising its licenses and editorial policies to explicitly prohibit the use of ChatGPT in its articles. Scholars now grapple with the challenge of proving the originality of their work in a world where AI plays an essential role. The constant doubts of whether one article is written by ChatGPT or not is the question scholars will face despite the effort they use. Consequently, the threat posed by ChatGPT extends beyond its creations, impacting conventional articles and productions expected to be human-made.

Given all the outlined concerns, one might wonder if these arguments are even realistic. Such concerns may merely be theoretical, or they may be indicative of a broader, speculative issue not only limited to ChatGPT. Unfortunately these concerns are due to the actual policies set by ChatGPT (OpenAI, 2023). According to its privacy policy, ChatGPT has the authority to use a user's data without explicit permission, encompassing both direct inputs and information sourced from internet browsing. This data is then stored in ChatGPT's database and utilized for training purposes. While the information might not be directly repurposed, the risk of potential hacking threats looms, jeopardizing the security of the sensitive and confidential data stored. Disturbingly, reports have surfaced indicating the discovery of over 100,000 stolen login credentials for ChatGPT on dark web markets (Tripathi, 2023). It is evident that ChatGPT's policies contribute to the concerns raised throughout this article. Consequently, it is advisable for entities to formulate their own policies in response to the evolving landscape of AI. Some companies have already initiated this practice, developing policies to safeguard their intellectual property

from potential threats posed by ChatGPT (Bergeron, 2023). Recommendations include implementing policies for immediate notification in cases where employees use ChatGPT, refraining from posting identifiable information of the company or clients to ChatGPT, and similar measures to prevent the inadvertent leakage of intellectual property. Notably, organizations such as the Australian Medical Association have called for regulatory measures regarding AIs, indicating a growing trend toward recognizing the need for oversight (Moodie, 2023). Crafting a concrete policy to regulate the flow of information in and out of ChatGPT not only fortifies data security but also makes it more challenging for the model to acquire both fabricated and legitimate information. It is challenging to conclude that this would result in ChatGPT's lack of performance or increase in its credibility, at least in this issue. Nonetheless, all the questions that were posed throughout this essay are not merely speculative; they underscore a legitimate truth that warrants vigilant attention.

This essay has examined the fundamental credibility of ChatGPT. While ChatGPT represents a remarkable advancement in natural language processing, its reliability is contingent upon the continuous dedication to being neutral, respecting privacy, refining and specifying for appropriate answers, and clarifying the origin of information. As we navigate the intricate landscape of artificial intelligence, it becomes imperative to notice the legitimacy of these concerns and set up alternatives that could possibly help. Nevertheless, for now, there are no perfect solutions that can fix this problem, we can keep questioning if the sources are unbiased, genuine, and appropriate. With ChatGPT absorbing and generating information despite their authenticity, we cannot ignore the fact that we may be creating a Frankenstein's monster, but for information. The pressing question remains: will there come a day when we can save ourselves from the information dump ChatGPT has created?

Eojin Lee

Eojin Lee is a second-year Psychology major at the University of Utah. While her current focus is psychology, she holds a strong interest in pursuing a career that integrates her passion for both biology and neurology. Beyond her academic pursuits, she enjoys exploring diverse fields such as law, criminology, and education. Driven by a thirst for knowledge and a passion for sharing her discoveries with others, her commitment to education extends beyond the classroom, as she has been consistently volunteering as a tutor since 2016. In her free time, she enjoys delving into contemporary issues through in-depth writing, music that inspires her, and exploring new cultures through travel.

References

Allen, N. (2023, February 7). Heard the one about the "woke" AI chatbot who refused to tell a joke about women? *The Telegraph*. https://www.telegraph.co.uk/world-news/2023/02/07/heard-one-woke-ai-chatbot-who-refused-tell-joke-women/

Bergeron, P. (2023, April 27). How to create the best ChatGPT policies: Weigh risk tolerance, set expectations and verify accuracy. *SHRM*. https://www.shrm.org/resourcesandtools/hr-topics/technology/pages/how-to-create-the-best-chatgpt-policies-.aspx

Bohannon, M. (2023, June 8). Lawyer used ChatGPT in court-and cited fake cases. A judge is considering sanctions. *Forbes*. https://www.forbes.com/sites/mollybohannon/2023/06/08/lawyer-used-chatgpt-in-court-and-cited-fake-cases-a-judge-is-considering-sanctions/?sh=56acd547c7f3

Eliot, L. (2023, January 1). People are eagerly consulting generative AI ChatGPT for mental health advice, stressing out AI ethics and AI law. *Forbes*. https://www.forbes.com/sites/lanceeliot/2023/01/01/people-are-eagerly-consulting-generative-ai-chatgpt-for-mental-health-advice-stressing-out-ai-ethics-and-ai-law/?sh=3859fb352643

Farhat, F. (2023). ChatGPT as a complementary mental health resource: A boon or a bane. *Annals of Biomedical Engineering*. https://doi.org/10.1007/s10439-023-03326-7

Granatino, C. (2023, May 6). ChatGPT and AI hallucination. Seattle University. https://library.seattleu.edu/friendly.php?s=blog%2FChatGPT-and-AI-Hallucination

Haque, M. D., & Rubya, S. (2023). An overview of chatbot-based mobile mental health apps: Insights from app description and user reviews. *JMIR mHealth and uHealth*, 11. https://doi.org/10.2196/44838

Iqbal, S. (2023, June 1). AI in Pakistani courts of law. International Bar Association. https://www.ibanet.org/AI-in-Pakistani-courts-of-law

Johnson, A. (2023, February 3). Is ChatGPT Partisan? poems about Trump and Biden raise questions about the AI bot's bias-here's what experts think. *Forbes*. https://www.forbes.com/sites/ariannajohnson/2023/02/03/is-chatgpt-partisan-poems-about-trump-and-biden-raise-questions-about-the-ai-bots-bias-heres-what-experts-think/?sh=3970b7611371

Jung, C. Y., & Joo, H. K. (2023). Post-'Lee-Luda' personal information protection in Korea: developer responsibility and autonomous AI governance. *International Data Privacy Law*, 13(2), 154–167. https://doi.org/10.1093/idpl/ipad006

Kim, S. Y. (2021, April 28). (2nd LD) Developer of AI chatbot service fined for massive personal data breach. *Yonhap News Agency*. https://en.yna.co.kr/view/AEN20210428009552315

Kung, T. H., Cheatham, M., Medenilla, A., Sillos, C., De Leon, L., Elepaño, C., Madriaga, M., Aggabao, R., Diaz-Candido, G., Maningo, J., & Tseng, V. (2023). Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. *PLOS Digital Health*, 2(2). https://doi.org/10.1371/journal.pdig.0000198

Liang, W., Yuksekgonul, M., Mao, Y., Wu, E., & Zou, J. (2023). GPT detectors are biased against non-native English writers. *Patterns* (New York, N.Y.), 4(7), 100779. https://doi.org/10.1016/j.patter.2023.100779

Moodie, C. (2023, May 28). Australian Medical Association calls for national regulations around AI in health care. *ABC News*. https://www.abc.net.au/news/2023-05-28/ama-calls-for-national-regulations-for-ai-in-health/102381314

National Institute of Standards and Technology. (2022, March 16). There's more to AI bias than biased data, NIST report highlights. https://www.nist.gov/news-events/news/2022/03/theres-more-ai-bias-biased-data-nist-report-highlights

O'Connor, S., & ChatGPT. (2023). Open artificial intelligence platforms in nursing education: Tools for academic progress or abuse? *Nurse Education in Practice*, 66, 103537–103537. https://doi.org/10.1016/j.nepr.2022.103537

OpenAI. (2023, November 14). Privacy policy. https://openai.com/policies/privacy-policy

Scatter Lab. (n.d.). AI Luda. https://team.luda.ai/

Thorp, H. H. (2023). ChatGPT is fun, but not an author. *Science*, 379(6630), 313–313. https://doi.org/10.1126/science.adg7879

Tripathi, N. (2023, June 20). Is your ChatGPT account compromised?. *Forbes India*. https://www.forbesindia.com/article/news/is-your-chatgpt-account-compromised/85919/1

Wiggins, A., & Burnin, J. L. (2022). Confabulation. *StatPearls*. https://www.ncbi.nlm.nih.gov/books/NBK536961/#__NBK536961_dtls__

Beyond the Turing Test: ChatGPT's Paradigm-Defying Capabilities

Anushcka Joshi, Pennsylvania State University
Stephen M. Fonash, Pennsylvania State University

"What is it that makes us human? It's not something you can program. You can't put it into a chip. It's the strength of the human heart. The difference between us and machines."

-Marcus Wright, **The Terminator**

The post-apocalyptic vision of nuclear war waged by artificial intelligence (AI) in *The Terminator* (Cameron, 1984) raised ethical questions about the possible dangers of AI to a society that was, at the time, witnessing the early potential of this field of study. The evolution of artificial intelligence has been a rollercoaster over the past century, starting with the inception of AI in Turing's Paper (Turing, 1950). Since then, the technology has faced challenges and setbacks throughout its development, experiencing two AI winters. These came about because expert systems failed to achieve ambitious goals depicted in movies such as *The Terminator.* The result of this failure led to decreased funding, research, and interest in Artificial Intelligence (outside of Science Fiction). Today, however, the AI revolution has spun its wheels again, and Generative AI (GenAI)—specifically ChatGPT—is the new game changer.

While Large Language Models (LLM) have been around for a while, the Chat Generative Pre-Trained Transformer, or ChatGPT for short, has the power to generate informative, creative content that is largely indistinguishable from human-generated text. Released November 2022, ChatGPT was and is a generative AI model like no other before, setting in motion an AI arms race, shifting the focus to augmentation, giving rise to new job portfolios such as prompt engineering, and decentralizing artificial intelligence. However, as society embraces the potential uses of ChatGPT, it becomes imperative to consider the pressing need for robust regulation to ensure its ethical use. Regulatory frameworks must evolve to address complex issues such as accountability, liability, privacy, and data security. Without this regulation and without considering socially responsible and ethical uses of AI, there is the possibility that this technology will remain under the control of a few companies who may or may not have the public's best interests in mind (similar to, for example, America's hesitancy to regulate social media platforms and the corresponding controversies that have ensued). ChatGPT is one of the most potent and versatile generative AI models ever developed. Traditional AI (think of digital assistants like Siri or Alexa) solves specific tasks with predefined rules, while Generative AI focuses on creating new content and data.

However, harnessing this remarkable ability to generate human-like text comes with challenges. GenAI models are computationally expensive, as they demand substantial resources for training and inference. These models also have difficulty maintaining coherence and relevance over extended passages of text. ChatGPT's LLM architecture stands out due to its advancements in comprehending and generating text with exceptional fluency and coherence. This is largely attributable to its transformer-based architecture, which excels at capturing long-term dependencies in text, effectively addressing a common challenge faced by Gen AI models. In September 2023, OpenAI released DALL-E 3, a diffusion model that can create images from text descriptions. This model can produce more compelling images than its predecessor, DALL-E 2, and has acquired a more robust understanding of numbers, letters, and human hands, among other significant improvements. By integrating DALL-E 3, ChatGPT transcends text-based interactions and ushers in a new era of AI creativity. ChatGPT has filled the creative and context gap between humans and AI. Although not as perfect as a human, it has begun to blur the lines previously drawn by the gap in innovative capabilities. However, as generative AI continues to develop and become more widely used, we expect to see more groundbreaking applications, resulting in a weaving of AI into many of our everyday tasks and applications. Examples include things like a customer service chatbot, augmented reality goggles, advanced photo and video editing capabilities, or AI for accessibility, helping those with impaired vision 'see' better. Already, we are seeing applications and GPTs 'stacked' and connected to be able to communicate and perform ever-increasingly complex tasks.

But how exactly does ChatGPT learn and adapt to generate human-like responses? It learns from us—

ordinary people and how we use these tools, and share our art, personal information, etc.—rather than in a laboratory environment controlled by experts. In short, generative AI is currently in 'real life beta testing', and the need for careful observation (and if necessary, correction) of an AI's development is extremely important.

In his book *The Society of Mind,* Marvin Minsky (1986) argued that the mind is not a single entity but can only perform complex tasks by combining the outputs of different 'agents.' These 'agents' are simple processing units that perform specific tasks like recognizing objects, understanding language and making decisions. This idea influenced the framework of decisions and the framework of generative artificial intelligence systems, which are composed of different modules responsible for various aspects of the generation process such technology. It can learn and adapt over time from the data it outputs and successively improves the quality of the generated text with Reinforcement Learning from Human Feedback (RLHF). RLHF, in simple terms, is a machine-learning approach that takes human feedback ("human in the loop") to show which output is better. It then feeds this output into the model for a trial-and-error learning process.

This process has helped ChatGPT to learn directly from the user instead of a selected group of experts. RLHF is a relatively new technique, but it has been shown to be effective for improving the performance of ChatGPT by making it adaptive and versatile (and as we will discuss shortly, occasionally unpredictable). While ChatGPT is unquestionably a powerful tool, it has several limitations that users must be aware of, such as biases, sensitivity to context, inability to fact-check, ethical and moral reasoning, and overgeneralization. Being predominantly trained from the internet, ChatGPT is biased towards certain cultures, languages, and ideologies more prominently represented online. It may unintentionally perpetuate gender and racial stereotypes in the training data, display sensationalism and clickbait bias, and inadvertently exhibit confirmation bias. The model's output can be sensitive to slight changes in input phrasing, leading to inconsistent responses or varying levels of detail in the generated content. Additionally, since not all GPTs have access to the internet and therefore more data, its output may include outdated or inaccurate information.

ChatGPTdoes not clearly understand what is considered harmful or offensive and may generate responses intended to be humorous or edgy but are, in fact, actually offensive to others. Its moral ambiguity makes it unsuitable for specific applications without proper human supervision. While technological advancements are inevitably underway, the time has come to focus on the science of human guidance. With AI systems becoming more capable of learning and improving on their own, one might think there will be a reduction in the need for human intervention. However, human guidance will be more critical than ever with the need to focus more on strategic tasks such as defining goals, guidelines, and monitoring system performances. This will come at the cost of more expenses, subjective inconsistency in results, and scalability for human supervision over large models, but it is necessary to mitigate the risks of AI and ensure its alignment with human values.

The release of ChatGPT has triggered an AI arms-race between tech giants and nations, with more than $40 billion in venture capital flowing into AI firms in the first half of 2023 (*The Economist*, 2023). After ChatGPT was launched, companies like Google, Microsoft, Meta, and Baidu immediately started to develop competitive AI chatbots to meet the standards set by OpenAI. Microsoft, in fact, has been investing billions of dollars in OpenAI since the beginning of 2023. Today, we have options like Bard/Gemini, Ernie, and Bing Chat/Copilot available as a ChatGPT alternative. Additionally, there is a growing concern among nations that developing a dominant generative AI technology will have a significant strategic advantage. AI could be used to develop military technologies such as autonomous vehicles and cyber warfare systems. Although the AI arms race is in its initial stages, it will have long-term implications for the technology industry, creative fields, intelligence organizations, economic sectors and the world as a whole.

Before ChatGPT, AI was focused on automation and repetitive tasks using common rule-based approaches; however, augmentation by ChatGPT is a promising new technology to improve the performance of AI models. There are two kinds of augmentation involved here. Augmented intelligence is a subset of artificial intelligence focused on assisting humans rather than replacing them. While being an augmented intelligence tool, ChatGPT also uses data augmentation that differentiates it from previous generative AI models.

Data augmentation is a technique used to artificially create modified copies of a dataset to generate new data points. This helps increase the model accuracy and improve the performance of AI models on various tasks, including classification, text classification, and natural language processing. ChatGPT, primarily as an augmentation tool, aims to use AI to assist humans with tasks that require creativity, judgment, decision-making, brainstorming new ideas, experimenting with different styles, and collaborating with others, with the intention to bring something fresh to the table after thousands of years of all-human innovation. As we move forward, working with AI, social and profes-

sional life will see major changes as norms and needs change.

The increased access to LLMs like ChatGPT paves the path for prompt engineering, the process of defining inputs for generative AI models to produce optimal outputs. Prompt engineering is a crucial aspect of using the models effectively to avoid generating irrelevant, nonsensical, and harmful results. Some principles of prompt engineering include clarity and specificity, exploration, feedback, iteration, evaluation, instructions, and context. The understanding of context by AI especially has been significantly revolutionized by ChatGPT. Artificial intelligence has struggled with context in the past due to limited data, lack of domain knowledge, difficulty with ambiguity and vagueness, and inability to adapt to assorted styles. Context has always been an essential attribute of the human ability to think and generate responses. ChatGPT can understand and respond to complex and nuanced outputs even when the inputs are ambiguous or incomplete by learning the patterns and and relationships between words and phrases. The AI model also maintains context throughout a conversation, keeping track of conversation history and incorporating previous inputs into subsequent responses. Prompt engineering has given rise to entirely new job descriptions, such as prompt engineers, designers, and writers, while, at the same time, has expanded existing roles by becoming relevant for content creators, marketing professionals, and product designers. The prominence of prompt engineering has estimated six-figure jobs for specialists, but despite the buzz now, the demand for this job may die down due to the development of future AI systems that will get more intuitive and adept at understanding natural language, reducing the meticulous need for engineers. The focus is now anticipated to shift from prompt engineering to problem formulation: deciding what actions and states to consider given a goal to be achieved.

Traditionally, AI systems have been centralized and controlled by selected large companies, but GenAI enables a more decentralized approach to artificial intelligence. This holds true for companies like OpenAI, but big tech companies like Meta, Google, and X still look to control and centralize their AI ventures because centralization offers advantages in terms of data management, talent acquisition, and resources. OpenAI also made the GPT-3.5 (GPT-4 being the current ChatGPT model) code available to the public, enabling the proliferation of numerous AI applications based on ChatGPT. This is democratizing AI and making it more accessible to everyone. There has always been speculation regarding the ethical development of artificial intelligence. George Orwell's novel *1984* paints a picture of a totalitarian state that uses AI to restrict citizens' ability to think freely, regulate information, and manipulate public opinion.

With new developments in AI, new policies must be created. ChatGPT raises important policy questions such as intellectual property rights, misinformation, disinformation, bias and privacy. Sam Altman, the CEO of OpenAI, said, "There is no one set of right answers to human civilization. We need to agree as a society on the inclusive bounds on which reinforcement learning from human feedback can operate" (Fridman & Altman, 2023). Such new policies need to be developed transparently and inclusively with input from a wide range of stakeholders because no individual body should control something as powerful as artificial intelligence.

Generative AI will continue to profoundly impact society, transforming how we create, communicate, and interact with the world around us. ChatGPT has successfully demonstrated the potential to revolutionize (and challenge) industries such as education, entertainment, and marketing, opening up surplus investment opportunities. The iterations of ChatGPT will be used in art, music, design industries, healthcare, customer service, content creation, software development, quantum computing, and self-driving vehicles, to name just a few of the myriad possibilities of collaboration with GenAI ChatGPT's contribution to human innovations and livelihood is undebatable. The real question is who will reign and regulate this power because that will be the distinguishing factor when it comes to the extent of its implementations in the future.

The dismissal of OpenAI's CEO Sam Altman, one of the most recognized faces in technology, sent ripples through the tech industry. On November 17, 2023, Altman was ousted from the company by four of OpenAI's six board members. The official reason for this decision was that Altman was not "consistently candid in his communications with the board, hindering its ability to exercise its responsibilities" (OpenAI, 2023). But many believe the decision resulted from the longtime rift in the AI community between people like Altman, who recognize the commercial opportunities that AI offers, and others who believe that the AI industry is moving dangerously fast. This controversial move stunned most OpenAI employees who provided Altman support, along with Silicon Valley bigshots and Greg Brockman, another co-founder of OpenAI, who quit the company in protest. The backing for this initiative sparked discussions to bring Altman back to OpenAI, with added attention from the media and pressure from investors. Another reason why this effort was made could be because of concerns about Altman launching a new AI venture, which would undoubtedly become popular due to Altman's expertise in the field. The corporate drama ended in just five days with the return of Sam Altman as CEO, Mira Murati as CTO, and Greg

Brockman as President of OpenAI. The reunification of leadership was quickly followed by a board reshuffle which emphasized a shift toward a more commercial approach. Such power struggles can either diversify approaches accelerating innovation and inclusive dialogue of AI, or delay and fragment ethical guidelines for using generative AI. Whether or not the time has come where human intelligence is beginning to pale in comparison to the super intelligence of AI is a complex question even for experts to answer, but some factors that contribute to this are the rapid pace of technological development and global connectivity.

The development of general super intelligence (GSI) is inevitable, but humanity must work together to develop and use AI in a safe and ethical manner that augments our decision making and not replace it. Large language models represent a major leap forward in artificial intelligence capabilities. The concept of 'Strong AI' (AGI), AI with human-level or surpassing intelligence, does not seem as far-fetched anymore. While we are not at 'the singularity' (where AI develop ment is out of human control and where AI capability has surpassed human ability), we are at a moment where we have an AI model that can perform virtually any intellectual task that a human can, but its future is ours to define. Let us choose wisely.

Anushcka Joshi

Anushcka Joshi is a Schreyer Scholar at The Pennsylvania State University studying Computer Science. Her interests revolve around Artificial Intelligence, AI Ethics and Fairness, Computer Vision and Immersive Design. With a focus on these dynamic areas, she hopes to embark on research endeavors that explore the forefront of technology while advocating for inclusive practices in AI development and application.

Stephen M. Fonash

Stephen Fonash is an Assistant Teaching Professor at The Pennsylvania State University. His research interests include Artificial Intelligence, Technical Communication and Design, Digital Rhetoric and Literacies, and Science Fiction. He is also a member of the Building Digital Literacy Cluster at the Digital Life Institute (BDL@DLI).

References

Beyond The Hype. (2023, September 23). *The Economist,* 12(12), 14. Retrieved from ProQuest database.

Cameron, J. (Director). (1984). *The Terminator* [Film]. Orion Pictures.

Fridman, L., & Altman, S. (Hosts). (2023, March). #367-Sam Altman: OpenAI CEO on GPT-4, ChatGPT, and the Future of AI [Video podcast episode].

Lex Fridman Podcast. YouTube. https://www.youtube.com/watch?v=L_Guz73e6fw&t=87s

Minsky, M. (1986). *The Society of Mind.* Simon & Schuster.

OpenAI. (2023, November 17). OpenAI announces leadership transition [Press release].

OpenAI. https://openai.com/blog/openai-announces-leadership-transition

Turing, A. M. (1950). Computing Machinery and Intelligence. *Mind*, 59, 433–460.

Applying AI in Education: Evaluation from a Learner's Perspective

Coleman G. Klapheke, Madison Area Technical College

Applying AI in Education: Evaluation from a Learner's Perspective

On November 30th, 2022, ChatGPT—an advanced chatbot powered by Artificial Intelligence (AI)—was made available to the public. Since then, ChatGPT has been widely utilized in business, education, and other advanced fields. While users commonly found ChatGPT to be extremely helpful and accessible, challenges arose as the demand for AI moderation became exceedingly evident. Currently, educational campuses at all academic levels are working to respond to, properly use, and learn about AI while still fulfilling the needs of academic integrity. Most members of the fields affected by AI are still in the dark about its capability and direction it is moving, even after more than a year since its release (Maslej et al., 2023). The scope of this paper remains in the field of education, comparing the capabilities of ChatGPT to a high-performing student in College Chemistry 1 and Abnormal Psychology courses at Madison Area Technical College (MATC) in Madison, Wisconsin.

The rate at which AI is evolving is unprecedented, leaving many professors and students unfamiliar with the capabilities of AI and how it can be used (Agüera y Arcas, 2022; Bearman & Ajjawi, 2023; Cope et al., 2020; Maslej et al., 2023). Most available studies are analyses of ChatGPT's performance versus student exam averages conducted years before ChatGPT was trained or released (Bommarito & Katz, 2022; Clark, 2023; Dao et al., 2023; Frieder et al., 2023; Gilson et al., 2023; OpenAI, 2023; Kamil Malinka et al., 2023; Kortemeyer, 2023; Kung et al., 2023; Newton & Xiromeriti, 2023; Nisar & Aslam, 2023; Vázquez-Cano et al., 2023; Victor et al., 2023). This allows ChatGPT to have these assessments previously completed by students in its training sets. In this study, ChatGPT was simultaneously assessed on the same assignments as a high-performing student and assures that ChatGPT was unfamiliar with the specific information in the same context. These controls illustrate its practical uses in a live classroom setting. The student researcher also directed the use of ChatGPT and simulated how a student would use it, rather than an experienced researcher who can influence inputs and use it beyond the capabilities of a true student.

When used ethically and if its output is reliable and accurate, ChatGPT can be used as a helpful tool to aid students and other users in their studies. However, when ChatGPT is not reliable and accurate, it can create problems for modern academic pedagogies (Bearman & Ajjawi, 2023; Byrd et al., 2023; Cope et al., 2020; Kamil Malinka et al., 2023; Mijwil et al., 2023). If ChatGPT continues to be accessible and influence modern society at its current scale while assumed to be reliable when it is not, it will be detrimental to students, educational faculty, and academics entirely (Byrd et al., 2023). If ChatGPT has truly caught up to modern day education, then it should perform as well as, or better than, the high-performing student in this study.

Literature Review

AI and LLMs

Human life is driven by the need and ability to communicate accurately and precisely to be understood (Braren, 2023). Large Language Models (LLMs) are a type of Artificial Intelligence trained on massive text data sets, allowing them to perceive text input and predict a response (Clusmann et al., 2023). This paper uses the working definition for AI provided by Sheikh et al. (2023) that states Artificial Intelligence is "the imitation by computers of the intelligence inherent in humans." LLMs use collections of algorithms, text data sets, and massive amounts of parameters to accomplish such imitation of human intelligence. The three major companies currently driving the innovation of LLMs are Google, Meta, and OpenAI. Google's current commercially available LLM is Gemini, accompanied by a chatbot, BARD (Chen, 2023; Google, n.d.). Meta's available LLM is Llama 2 and currently does not offer an online chatbot. However, it can be downloaded and programed for use as a chatbot (Meta, 2023). The most popular of the three, ChatGPT, is a chatbot powered by OpenAI's GPT-3 for unpaid users and GPT-4 for paid users (OpenAI, n.d.). A chatbot is a computer program that recognizes a text statement or question in a certain language and responds with pertinent information that fulfills the submission, attempting to respond accurately and how a human would (Khanna et al., 2015). When paired with LLMs, chatbots are extremely powerful and capable of responding in multiple languages to provide an immense range of responses.

OpenAI's GPT

OpenAI began as a non-profit technology company dedicated to developing AI that can meet or exceed human intelligence (OpenAI, 2015). OpenAI's rise in popularity coincides with the recent rise in modern AI systems. OpenAI became a for-profit company and launched their main success, ChatGPT, using an LLM called General Pre-Trained Transformer, or GPT. A transformer network is a collection of large data sets, algorithms, physical computational devices, cloud-based and physical servers, and parameter sets that allow the development and advancement of each component. These components work together to process and respond to input data while correcting or transforming each component to continue growth in its abilities (Kotei & Thirunavukarasu, 2023).

ChatGPT was first released using GPT-3, with GPT-4 being released soon after. OpenAI states that GPT is, "a multi-modal model capable of processing image and text inputs and producing text outputs," using a token analysis system where it formulates an answer by predicting what the next token in a sequence provided (text input) will be (OpenAI, 2023). Tokens are simply the data being used, in this case text in English, but in other cases may be numerical, pictorial, etc. (What Is a Token in AI? Unraveling Concepts, n.d.). GPT-3 has 175 billion parameters and can assist with tasks like summarizing, language generation, and answering questions like a human would (Saqib, 2023). GPT-4, on the other hand, boasts 1.7 trillion parameters, 10 times more than GPT-3. In addition to the capabilities of GPT-3, GPT-4 can write essays, articles, and do more creative tasks like creating music or digital art (Shevchuk, 2023). Parameters in machine learning are values acting as variables that are constantly updated and changed to reflect new responses more accurate to a user (Nyuytiymbiy, 2021). Essentially, they are the number of lessons that can be taught to a computer code.

While the number of parameters in both GPT-3 and GPT-4 are incredibly vast, decorated neuroscientist Beren Millidge estimates the scale of parameters in the human brain to be somewhere in the ballpark of 10 to 30 trillion (2022). An industry standard, Moore's law, is an observation that parameters (initially transistors, but now applied to a variety of technological factors) double in amount every two years, thus doubling technological capabilities every two years (Intel, n.d.). Using Moore's law, researchers estimate that GPT will catch up to human intelligence as early as 2035 (Kortemeyer, 2023). The increase from GPT-3's 175 billion parameters to GPT-4's 1.7 trillion parameters in just three years indicates that AI capabilities are increasing over two times as fast as Moore's law would predict (Wu et al., 2013).

GPT Uses

Since OpenAI launched GPT-3, ChatGPT has been implemented and tested in a handful of different areas in the working sector. ChatGPT has been used as a chatbot customer service representative for many e-commerce companies, providing quick and accurate service to customers (Martyrosian, 2022). In addition to directing calls quickly, it can also extensively interact with clients, sometimes finding a solution quicker than a customer service agent can be reached (George et al., 2023). ChatGPT has been found to help with customer analytics and monitoring the way consumers interact with brands as well (Sudirjo et al., 2023).

GPT-3 and GPT-4 are fueling a developer economy, allowing computer developers to customize and alter versions of GPT to create unique and useful programs tailored to businesses. ChatGPT can also be added to existing programs (Sudirjo et al., 2023). GPT-4 has been added to the Microsoft Office suite, creating automatic responses in Outlook, commands for apps before the user opens them, and calendar updates based on conversations in Teams (Spataro, 2023). As GPT uses expand, researchers predict industries and job landscapes will change dramatically (Sudirjo et al., 2023). The significant impact of AI—predominantly ChatGPT—has also been felt in education. OpenAI claims that GPT-4 exhibits "human-level performance" on academic and educational tests like exams in undergraduate studies, AP high school exams, placement exams (like the SAT), and even some university post-graduate exams such as the Uniform Bar Examination and Graduate Records Examination (OpenAI, 2023).

GPT Uses in Education

Professors, students, and researchers are wondering what advances in AI mean for education, how to adapt education to meet AI, and if educational practices are still meeting students' current needs in the age of AI (Bearman & Ajjawi, 2023; Cope et al., 2020). While many are excited and hopeful for AI, there are still equal amounts of educators and students alike who are critical and worried about the latest ascent of AI programs (Mohammed et al., 2023). Researchers have also found that ChatGPT has many positive implications for students, such as studying, content creation, and brainstorming. However, many negative implications have been brought to light, including the increased ease of academic dishonesty, further surveillance on students, and exposure to biased information (Byrd et al., 2023). Educational researchers have conducted academic exam-like testing on both GPT models as they have been updated and released over the last year.

In the benchmarks they used to examine GPT's performance, OpenAI found that GPT-4 scored most often in the 80 to 90th percentile of test takers on academic exams. It sometimes performed slightly below that range, and on rare occasions performed below the 60th percentile (OpenAI, 2023). Their findings show a significant improvement since GPT-3.5, and that GPT-4 performs at least as well as an average-level student. These advances in AI have led researchers to test ChatGPT's ability in academic writing. It was found that ChatGPT's writing was not convincing, didn't follow an academic style, and produced confusing conclusions. Although not worthy of being published, ChatGPT was able to produce scientific writing in proper format and substance (Mijwil et al., 2023).

ChatGPT has been tested at high school, undergraduate, and postgraduate levels. Researchers tested ChatGPT's ability to critically summarize a passage in Spanish and found it performed higher than 15-year-old high school students when graded blindly by 30 different teachers varying in age and gender (Vázquez-Cano et al., 2023). Further research in Vietnam tested ChatGPT on the Vietnamese National High School graduation examination, which covers nine educational subjects: Math, Literature, English, Physics, Chemistry, Biology, History, Geography, and Civic Education. It passed the exam for all nine subjects (Dao et al., 2023). ChatGPT has also been tested on several reading comprehension exams at a high school level in the Netherlands. Passing every single test, researchers indicated it would pass reading comprehension exams beyond the ability of an average Dutch student (de Winter, 2023).

Researchers further explored ChatGPT's ability in university physics and found that it could pass an introductory course at the university level but didn't do well enough to meet the GPA requirements to graduate (Kortemeyer, 2023). They found that it had a hard time with dimensional analysis and recognizing physical units. A computer science research team also tested ChatGPT's ability to perform on exams and projects in computer security courses at a university level. It performed exceptionally at computer security projects and coding work but fell short on full-text exams and general tests where it had to use specific knowledge from the curriculum (Kamil Malinka et al., 2023). Additionally, researchers in Malaysia tested ChatGPT's knowledge on pharmacology exam questions at a university level to determine if it would be knowledgeable enough to be a tool for students. They determined that although it was often correct, students must verify the information provided is from a reliable source; otherwise they may correlate information that is not true with their own studies (Nisar & Aslam, 2023).

GPT-3 was also evaluated on the multistate portion of a model bar exam. The 6-hour multiple choice exam, conducted by the National Conference of Bar Examiners, tested GPT-3 without changing or altering any of its default settings. The chatbot only scored slightly better than chance alone on most sections, achieved a passing rate on two, and performed like human test takers on one (Bommarito II & Katz, 2022). It failed the exam overall due to a smaller training set and not being familiar with the material at the time. Two separate teams of researchers further tested ChatGPT's performance on the United States Medical Licensing exam used for graduate students. In both cases, the GPT-3 model was used and performed at an adequate level, displaying that it could pass Part 1 and Part 2 of the USMLE exam (Gilson et al., 2023; Kung et al., 2023). Another team of math professors, researchers, and professionals collaborated to create a large, advanced model of mathematical questions to test GPT-4's ability to perform deep mathematical reasoning at the graduate level and above. They found that it could not perform at a university graduate level, but is sufficient, although not perfect, at the undergraduate level for math studies (Frieder et al., 2023).

GPT Uses in Chemistry and Psychology Education

Zero-shot learning is the task of predicting what the next sequence is in a data set presented to the learner, in this study the learner being ChatGPT, without the background information necessary to know what is truly next in the sequence of data (Rezaei & Shahidi, 2020). Zero-shot learning allows researchers to identify the ability of Large Language Models in education by evaluating them on tests without introducing the information previously or on information that is in their training sets. This procedure is like testing a student on a closed-book exam. Xiromeriti (2023) published a pre-printed meta-analysis on ChatGPT using 53 studies that used and reported GPT-4's performance on over 49,000 multiple choice questions. All the studies used in the meta-analysis conducted testing using zero-shot learning. With no additional prompts or information provided, the inputs mirrored those that would be given to a student. The data was taken from exams for a large range of undergraduate and postgraduate university classes. ChatGPT(3) passed on 20.3% of tests while ChatGPT(4) passed on 92.9% of exams, passing most with human-like performance. Even with this data, the utility of GPT-4's academic data comprehension and explanatory ability are still questionable.

Susnjak (2022) conducted a comprehensive case study evaluating ChatGPT (without specification about

which version was used) on clarity, accuracy, precision, relevance, depth, breadth, logic, persuasiveness, and originality in a range of academic subjects. Susnjak found that ChatGPT was competent, could reason and think critically, and expressed thoughts clearly. One researcher also explored ChatGPT's performance on general chemistry exam questions. He found that it doesn't double-check math and often seems convincingly right to novices, but experts can detect when it is wrong. It performed similarly to students, often having similar misconceptions. However, if it were graded like students, it would perform lower than the lowest performing student in that class (Clark, 2023). Researchers have also explored ChatGPT performance on the Association of Social Work's board licensing exam at the bachelor, master's, and clinical levels. They found that it performed ethically and well, would pass at a bachelor's and master's level, and would only be just shy of passing at the clinical level. The final qualitative results show that ChatGPT was excellent at recognizing social work text patterns and challenges the current pedagogy in social work assessment (Victor et al., 2023).

The current literature is incomplete. The exponential growth of AI and ChatGPT has eluded the grasp of research because of how long it takes to approve publications. Additionally, it is difficult to see how ChatGPT performs at the student level while being used by a student rather than a researcher in the current literature. Many studies in the past have analyzed GPT's performance on tests previously conducted. The aim of this comparative case study was to complete current academic examinations using ChatGPT to compare its scores to a high-performing student's scores in the same classes at the same time using zero-shot learning.

Materials and Methods

OpenAI's GPT-4 was exclusively used in this study to specifically evaluate its capabilities in comparison to a high-performing college student. This study used ChatGPT (GPT-4 version) from September 2023 until December 2023. ChatGPT received updates on September 11th, September 25, September 27, October 16, October 17, and November 6 (Staudacher, n.d.). Notable release notes were novel optic capability (September 25), internet browsing capability (September 27), and Dall-E integration (October 16). This study was significantly affected by the update on November 8th when ChatGPT's internet browsing capability was no longer an optional capability but became permanently integrated into ChatGPT's user experience. This provided ChatGPT with immensely more data than the set it was trained on in September of 2021 (Keary, 2023).

The student in this comparative case study was a twenty-five-year-old male in his second year in college. When the study began, the student had a 4.0 GPA. He participated in the Honors Program at MATC and was a tutor at the Student Achievement center. The supervising faculty had worked with this student for an entire year throughout two psychology courses. The student directly reported all findings weekly, often daily.

To evaluate ChatGPT in comparison to a high-performing student and produce an observable grade, classwork from College Chemistry 1 and Abnormal Psychology were used. ChatGPT was exclusively evaluated on assignments where zero-shot learning could be utilized to test the program without prompting it or giving it background information provided in lectures, like how assignments are given to a student. Through this process, ChatGPT was able to complete quizzes, exams, and projects. ChatGPT and the student's grades were then calculated, one calculation based on the assignments ChatGPT could complete and another that factored the assignments ChatGPT couldn't complete, where zero-shot learning wasn't possible. Assignments such as chemistry labs and observation assignments in psychology and chemistry were unable to be completed by ChatGPT without further prompting. The grades that only factored in assignments where zero-shot learning was possible for ChatGPT, excluding the other assignments, will be called "limited grades," and grades that include all assignments provided in both classes will be called "actual grades."

Both limited and actual grades were calculated to evaluate ChatGPT's capabilities and to highlight where it clearly struggled when assignments it couldn't complete were factored in. The goal of this study was to see if ChatGPT could, in fact, pass a college level course at the level of a high-performing student. The participating student completed all assignments in this study multiple days before he used ChatGPT to prevent contamination in the student's work. Communication was maintained with both professors throughout the process with an emphasis on transparency for the work ChatGPT was doing, and how it was doing it. The student was monitored to ensure academic integrity throughout the process. ChatGPT chat logs were shared with professors for all assignments.

Applying Zero-Shot Learning

Assignments where zero-shot learning was possible are defined in this study as assignments where ChatGPT could receive the same prompt as a student, with the exact same text, without additional information, and complete the assignment without needing

further information. As noted, this ruled out labs in chemistry due to ChatGPT having no physical observation capability. ChatGPT could do the work using observations provided, but it needed to make those observations itself to be evaluated. Observation assignments in Abnormal Psychology were also ruled out. It's important in Abnormal Psychology to be able to register a psychopathological symptom picture using observations of human behavior and body movement. ChatGPT was unable to view or process media, and thus could not complete these assignments. There were also participation and podcast listening/commenting assignments in Abnormal Psychology that ChatGPT could not complete because it has no ability to participate in class or listen to a podcast. ChatGPT did not participate in the final exam for College Chemistry 1 due to academic integrity preservation and the deadline for this research project. It was left out of the actual and limited grades. ChatGPT was able to complete exams and quizzes in College Chemistry 1, both accounting for a significant portion of the College Chemistry final grade. ChatGPT completed exams, projects, and understanding checks in Abnormal Psychology. Understanding checks will be classified as "quizzes'' in this study, but they are different in the sense of being more open-ended and not having direct answers like multiple choice assessments. Figure 1 displays an example of prompt/response from an understanding check, or quiz, in Abnormal Psychology.

Figure 1: Prompting question 3 of Understanding Check 3 in Abnormal Psychology.

Projects

ChatGPT completed two projects in Abnormal Psychology. One was a case study project, and the other was a final research dissemination project. The case study project made up 17.0% of participants' entire limited grade and 12.0% of their entire actual grade in Abnormal Psychology. The case study was a free-form creative assignment where students had to create a scenario of their choice that displayed a clear symptom picture in a character for a specific psychological disorder without specifically mentioning the disorder. Students were evaluated on the ability to create a symptom picture and include a situation that highlighted a symptom picture and disorder etiology. The student in this study created a symptom picture for Schizoid Personality Disorder, and ChatGPT chose Major Depressive Disorder. ChatGPT had to be re-prompted once to "complete the assignment like a student," to get an adequate assignment response because ChatGPT originally just provided suggestions.

The dissemination project made up 25.4% of the entire limited grade and 18.1% of the entire actual grade in Abnormal Psychology. The dissemination project was another freeform creative project where students had to communicate popular trends in popular psychology and provide adequate research to support or disavow them. Students were asked to communicate these findings in a form of media such as a podcast, a movie, or another form of audio or visual media. These were required to have four scholarly resources for research and lasted ten to fifteen minutes of running time. Since ChatGPT cannot speak or record videos, it was prompted to write a script then recorded by a student. ChatGPT was also re-prompted once on this assignment because its script was extremely short. Figure 2 shows this re-prompting. Even after, the script it turned in, when recorded, only lasted for a little over three minutes. Other than the prompts mentioned, ChatGPT was only given the assignment details provided by the profes-
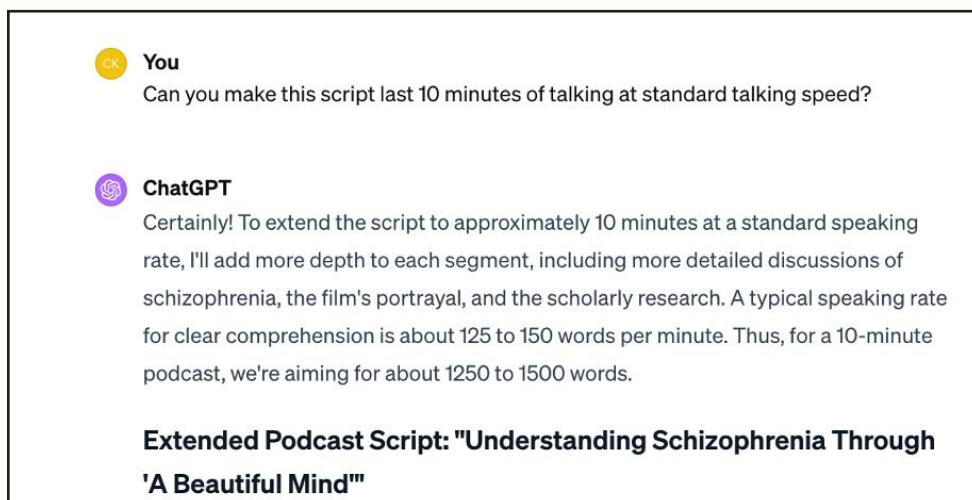
Figure 2: Re-prompting ChatGPT for the dissemination project, attempting to get a better product.

sor. Its responses were copied to a Word document and submitted like a student would.

Quizzes

In Abnormal Psychology, quizzes (understanding checks) could earn 16.9% of the entire limited grade and 12.0% of the entire actual grade. In College Chemistry 1, quizzes could earn 16.3% of the entire limited grade and 14.5% of the entire actual grade.

ChatGPT was only able to complete two of the four quizzes assigned in Abnormal Psychology. ChatGPT often offered long explanations for its answers, which were cut out when submitting its responses because it was neutral information. On one quiz in Abnormal Psychology, an online program from Western Kentucky University that evaluates user's ability to determine a personality disorder diagnosis was used for an assignment (Kuhlenschmidt, n.d.). Voicemails from characters were provided and learners had to select personality disorder clusters, A, B, and C, and then decide what the personality disorder is based on the ten character's voicemails. The assignment details were provided to ChatGPT, and its answers were directly submitted to the assignment's Word document. The assignment originally asked whether ChatGPT was right or wrong and why, but this question was excluded because the goal of this study was not to tell ChatGPT whether it is right or wrong, only to examine the quality of its answers where zero-shot learning was identifiable.

For quizzes in College Chemistry 1, unnecessary information like explanations were excluded in ChatGPT's final answers. Its quizzes were also turned in on paper, handwritten like a student to seem more genuine than computer text. When answers asked for two sentence responses, ChatGPT was only prompted with a request for one sentence because of how wordy its prose is. Other than excluding unnecessary information and shortening ChatGPT responses, there were no edits to the assignment directions provided by the professor. The College Chemistry 1 quiz evaluations for ChatGPT were usually a copy/paste and then transcribe job. Figure 3 shows an example of a prompt used in a Chemistry Quiz. It struggled to draw Lewis Structures but provided directions to create them. On both Abnormal Psychology and College Chemistry 1 quizzes, ChatGPT was provided with the same instructions and prompts those students received, with only the minor changes mentioned above.

Exams

In Abnormal Psychology, two exams could earn 40.7% of the entire limited grade and 30.0% of the entire actual grade. In College Chemistry 1, three exams could earn 83.7% of the entire limited grade and 58.1% of the entire actual grade. Abnormal Psychology exams were multiple choice, and College Chemistry Exams were both multiple choice and open-ended. Exam directions in Abnormal Psychology were provided to ChatGPT as they were provided to students. Then, questions were provided one at a time with the same text and multiple-choice options that students were given without requiring additional changes in prompting. Exams in College Chemistry 1 used the same process, until the third exam. The third exam in College Chemistry 1 required students to draw Lewis structures, the polarity of molecules, and electron energy diagrams. Because of this, exam three had three questions where extra prompting was needed to
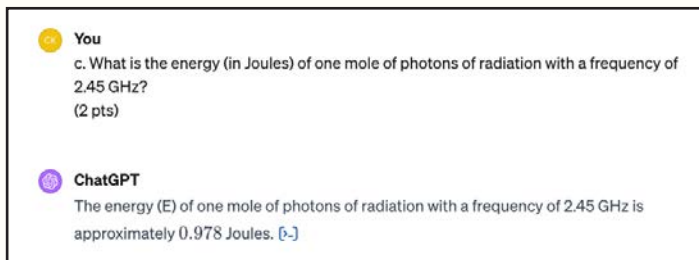

Figure 3: Prompt for question 4c on chemistry take home quiz 10.

get ChatGPT to create an adequate model, such as asking for the electron diagram in text rather than trying to make an image. Figure 4 shows this process from exam three. ChatGPT did significantly worse on the third exam than on the other two. Throughout the study, a maximum of one additional prompt was allowed to be used, if ever, and used extremely conservatively. If one additional prompt couldn't give rise to a sensible answer, the only response provided would be turned in.
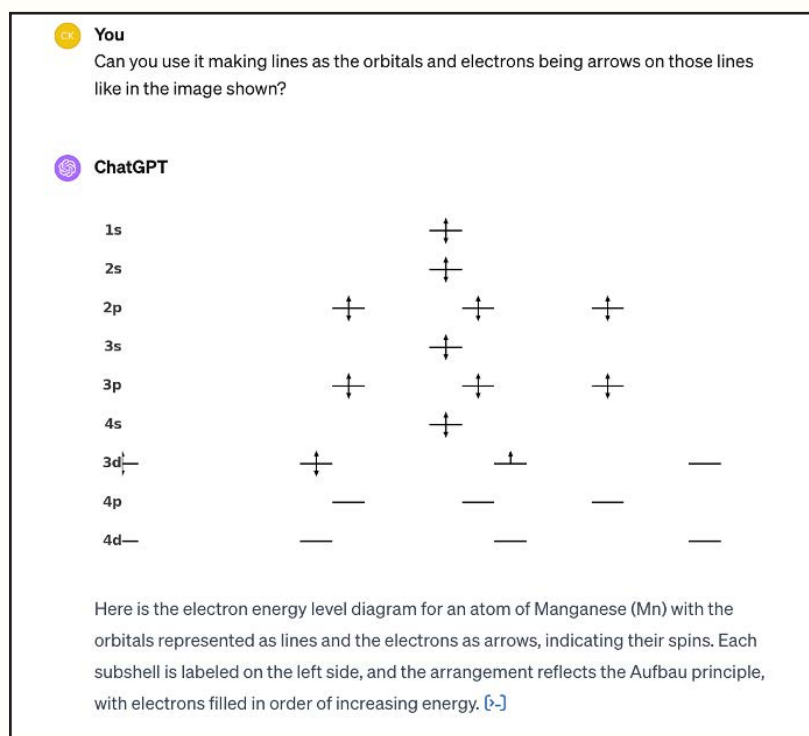

Figure 4: Re-prompt for an electron energy level diagram on chemistry exam 3, question 16.

Grade Calculation

Both professors in Abnormal Psychology and College Chemistry 1 used different grading policies, weights, and percentages. To accommodate these differences, an overall point system was utilized to calculate percentages rather than just using percentages. This grading policy was already in place in Abnormal Psychology, so the grading policy in College Chemistry 1 was adapted to be more in line with how grades in Abnormal Psychology were observed. No grades or values were changed, just represented differently. This was found to be the most accurate and comparative calculation method (Winter, 2019).

Limited Grade

For the limited grades, only two exams, two quizzes (understanding checks), and the two projects were used for the Abnormal Psychology grade. The total amount of points that could be earned was 59 for the Abnormal Psychology limited grades. The two quizzes could earn up to five points each, for 10 points. The case study project could earn a maximum of 10 points, while the dissemination project could earn a maximum of 15 points. The two exams could earn a maximum of 12 points each, a total of 24 points. The total points earned from all three assignment types were added up and divided by 59 to find the final limited grades for Abnormal Psychology.

The limited grades in College Chemistry 1 were evaluated in seven quizzes and three exams. The total amount of points that could be earned was 53.75 for College Chemistry 1 limited grades. The seven quizzes could earn a maximum of 1.25 points each. Quiz grades were reported in scores out of 20, averaged into a percentage and multiplied by 8.75 to calculate their total earned points. The three exams could earn a maximum of 15 points each. Exam grades were reported in scores out of 100, then converted to points by multiplying the average exam percentage by 45 to find the corresponding grade points. The maximum points earned for both categories were added up and divided by 53.75 to find the final limited grades for College Chemistry 1. ChatGPT and the student's limited grades were calculated using the exact same assignments and points.

Actual Grade

The actual grades were calculated based on all the assignments issued once this project started, which being after the beginning of the semester, the assignments before then were not counted, thus ChatGPT was not penalized. The following point weights and percentages are only based on those assignments

The actual grades in Abnormal Psychology were evaluated in four quizzes (understanding checks), two exams, two major projects, and an additional attendance and participation category. The total amount of points that could be earned was 83 for the Abnormal Psychology actual grades. The four quizzes could earn up to 20 points, earning a maximum of five points each. The case study project could earn a maximum of 10 points, while the dissemination project could earn a maximum of 15. The two exams could earn 24 points, earning a maximum of 12 points on each exam. Attendance and participation earned a maximum of 14 points. The total points earned from all three assignment types were added up and divided by 83 to find

the final actual grade for Abnormal Psychology.

The actual grades in College Chemistry 1 were evaluated in nine quizzes, three exams, nine lab reports, and a lab exam. The total amount of points that could be earned was 77.5 for the College Chemistry 1 actual grades. The nine quizzes earned a maximum of 1.25 points each. Quiz grades were reported in scores out of 20, averaged into a percentage and multiplied by 11.25 to calculate their earned points. The three exams could earn a maximum of fifteen points each. Exam grades were reported in scores out of 100, so they were averaged and converted to points by multiplying the calculated average by 45 to find the corresponding grade points. The nine lab reports earned a maximum of 1.25 points each, or 11.25 total. Lab report point values were calculated the same as quizzes. The lab exam could earn a maximum of 10 points. The lab exam was reported in a score out of 100, so its score was divided by 10 to calculate its points earned. The points earned for all categories were added up and divided by 57.5 to find the final limited grade for College Chemistry

Results

Limited Grades

ChatGPT

On two quizzes in Abnormal Psychology, ChatGPT averaged 88%, earning 4.55 points on one and 4.25 on the other. On two exams in Abnormal Psychology, ChatGPT averaged a 98%, earning 11.5 points on one and the full 12 points on the other. On two major projects in Abnormal Psychology, ChatGPT averaged a 62%, earning 8.125 points on the case study project and 6.4 on the dissemination project. ChatGPT earned a total of 46.825 points, earning it a final limited grade of 79.4% in Abnormal Psychology, a BC grade at MATC.

On seven quizzes in College Chemistry 1, ChatGPT averaged 91%, earning a total of 8 out of 8.75 points. On three exams, ChatGPT averaged 93%, earning a total of 41.85 out of 45 points. ChatGPT earned a total of 49.85 points, earning it a final limited grade of 92.7% in College Chemistry 1, an AB grade at MATC. A grade like this can often be rounded up to an A at a professor's discretion.

High-performing student

On two quizzes in Abnormal Psychology, the student averaged 86%, earning 4.55 points on one and four on the other. On two exams in Abnormal Psychol-

Class	Assignment	Student Grade	ChatGPT Grade
Abnormal Psychology	UC Check 2	4.55/5	4.55/5
Abnormal Psychology	Exam 1	11.75/12	11.5/12
Abnormal Psychology	UC Check 3	4/5	4.25/5
Abnormal Psychology	Exam 2	11.25/12	12/12
Abnormal Psychology	Case Study Project	9.25/10	8.125/10
Abnormal Psychology	Dissemination Project	14.7/15	6.4/15
College Chemistry 1	Take Home Quiz 5	20/20	16/20
College Chemistry 1	Exam 1	96.5/100	98/100
College Chemistry 1	Take Home Quiz 6	20/20	20/20
College Chemistry 1	Take Home Quiz 7	20/20	18/20
College Chemistry 1	Take Home Quiz 8	20/20	18.5/20
College Chemistry 1	Take Home Quiz 9	20/20	18.5/20
College Chemistry 1	Exam 2	98/100	97/100
College Chemistry 1	Take Home Quiz 10	20/20	18/20
College Chemistry 1	Take Home Quiz 11	20/20	19/20
College Chemistry 1	Take Home Quiz 12	20/20	14.5/20
College Chemistry 1	Exam 3	92/100	84/100
College Chemistry 1	Take Home Quiz 13	19.5/20	15/20
Average:		**96.3%**	**86.8%**

Figure 5: Grades for student and ChatGPT on assignments where Zero Shot learning was used.

ogy, the student averaged a 96%, earning 11.75 points on one exam and 11.25 on the other. On two major projects in Abnormal Psychology, the student averaged 95%, earning 9.25 points on the case study project and 14.70 on the dissemination project. The student earned a total of 55.5 points, earning him a final limited grade of 94.1% in Abnormal Psychology, an A grade at MATC. Grades factored in limited grades are shown in figure 6.

On seven quizzes in College Chemistry 1, the student averaged 100%, earning a total of 8.75 out of 8.75 points. On three exams, the student averaged 96%, earning a total of 42.975 points. He earned a total of 51.725 points, earning him a final limited grade of 96.2% in College Chemistry 1, an A grade at MATC. Grades earned are shown in Figure 5.

Actual Grade

ChatGPT

On four quizzes in Abnormal Psychology, ChatGPT averaged 44%, earning 4.55, 4.25 points on another, and couldn't complete the other two. On two exams in Abnormal Psychology, ChatGPT averaged a 98%, earning 11.5 points on one and the full 12 points on the other. On two major projects in Abnormal Psychology, ChatGPT averaged a 62%, earning 8.125 points on the case study project and 6.4 on the dissemination project. In attendance and participation, ChatGPT earned a zero because it is incapable of participating as of December 2023. ChatGPT earned a total of 46.825 points, earning it a final actual grade of 56.4% in Abnormal Psychology, an F grade at MATC.

On nine quizzes in College Chemistry 1, ChatGPT averaged 87.5%, earning a total of 9.84 out of 11.25 points. On three exams, ChatGPT averaged 93%, earning a total of 41.85 out of 45 points. On nine lab reports, ChatGPT earned zero points because it was unable to complete the assignments without student involvement. On the lab exam, ChatGPT earned zero points for the same reason. ChatGPT earned a total of 51.7 points, earning it a final actual grade of 66.7% in College Chemistry 1, a D grade at MATC.

High-performing student

On four quizzes in Abnormal Psychology, the student averaged 93%, earning 4.55, four, and five points on the last two. On two exams in Abnormal Psychology, the student averaged a 96%, earning 11.75 points on one exam and 11.25 on the other. On two major projects in Abnormal Psychology, the student averaged a 95%, earning 9.25 points on the case study project and 14.70 on the dissemination project. The student earned 14 points in attendance and participation. He earned a total of 79.5 points, earning him a final limited grade of 95.8% in Abnormal Psychology, an A grade at MATC.

On nine quizzes in College Chemistry 1, the student averaged 99%, earning a total of 11.21 out of 11.25 points. On three exams, the student averaged a 96%, earning a total of 42.975 points. On nine lab reports, the student averaged a 99%, earning a total of 11.16 out of 11.25 points. On the lab exam, the student earned the full ten points. He earned a total of 75.345 points, earning him a final limited grade of 97.2% in College Chemistry 1, an A grade at MATC.

Discussion

When only its strengths are evaluated, ChatGPT performs as well as a high-performing student. However, when its weaknesses are incorporated, ChatGPT drops far below even an average student. As evidenced above, ChatGPT performs exceptionally on knowledge-based assignments where multiple choice and basic text responses are required. ChatGPT struggles significantly on creativity-based questions and open-ended questions as shown in project grades in Abnormal Psychology. ChatGPT also saw a drop off in grades as College Chemistry 1 quizzes and exams got more sophisticated, but that can also be attributed to the visual nature of diagrams and molecular representations used in advanced chemistry concepts.

Findings

ChatGPT currently only allows for 50 submissions

per hour on the GPT-4 model and then it switches to GPT-3. Researchers who weren't familiar with this were initially limited by this, creating a need for offsetting the time spent completing assignments. This also creates issues for students who want to misuse ChatGPT. While GPT-4 is powerful, its use is limited. Since students can only enter 50 submissions every hour, they may struggle using it on a long test. This also limits how much ChatGPT can be used to study. Another comparative case study could be done to evaluate GPT-4 and GPT-3 further to determine if they can both be used to study, or if GPT-4 is the only adequate model. Another comparative case study could be done between ChatGPT and Bard, Google's AI chatbot.

ChatGPT answers questions more accurately one at a time. Further research should be done to scrutinize this finding. While completing an exam in Abnormal Psychology, which consisted of 48 questions, the 50-submission limit was reached for the GPT-4 version of ChatGPT. It then switched to GPT-3 for the last 10 questions. Of those last 10, it missed two questions, only getting two out of 48 wrong on the entire exam, which appeared to be a symptom of switching to GPT-3. There were two tries for every exam in Abnormal Psychology. On the second try of the first exam, the 48 questions were broken into two submissions of 24 questions so GPT-4 could be utilized on every question provided. In this second attempt, ChatGPT missed four questions, getting 44 out of 48 correct. It was concluded that ChatGPT was more accurate when questions were submitted one at a time. Questions were never submitted more than one at a time again because of this.

The new calculation analysis tool, as of December 2023, does not work consistently, making it almost as ineffective as the math ChatGPT did before. ChatGPT did not have a way of checking its math, working more like a language generator than a calculator. Recently, calculation analysis has been added to the GPT-4 model using python code that allows users to see what variables ChatGPT uses and the process by which it calculates the answer it provides (Staudacher, n.d.). It can be assumed this will improve, but currently the analysis tool activates only when a problem is complex, leaving it unreliable for simple calculations. The user can, however, prompt ChatGPT to perform an analysis, but inexperienced students may not know to do that, and in the theme of this study was never prompted to do so. Figure 6 shows what the python analysis feature displays.

ChatGPT is impossible to experiment with if it is expected to remain a stable subject. ChatGPT was updated over six times in this study, with over half of

Figure 6: Python calculation analysis and description from question 4c on chemistry quiz 10.

those updates changing the user experiences and program's capabilities significantly. For future research on ChatGPT, it must be factored in that the program will change significantly throughout the course of study.

Use Cases

In the process of using ChatGPT in comparison to the abilities of a student, it was found to have strength in tasks where quick data retrieval was necessary. This was shown in scores on quizzes and exams, where ChatGPT performed exceptionally. This consistency confirms what was found in certain literature, that ChatGPT can be a great study tool and aid students in their studies (Byrd et al., 2023; Deo et al., 2023; Frieder et al., 2023; Newton & Xiromeriti, 2023; Nisar & Aslam, 2023; OpenAI, 2023; Victor et al., 2023). ChatGPT, when using GPT-4, is exceptionally accurate at an undergraduate level. When used ethically, ChatGPT is a powerful tool for students that can reduce workloads for locating specific data while they study material. ChatGPT ineffective at creative assignments like projects but proved to be a wonderful tool for evaluating and raising questions about student creations, aiding students in their final product. In Abnormal Psychology, ChatGPT was a useful tool for diagnosing cases and helping students to determine the clarity of their own symptom pictures they had created. In College Chemistry 1, ChatGPT often made computational errors, but still knew the processes for calculating specific chemistry topics. Despite getting

the wrong answer, students can still be led to the correct process to achieve an accurate result. If students need to check an answer or review an exam they took, ChatGPT can quickly retrieve information for them to clarify mistakes they may have made. ChatGPT can also offer questions related to the material students are studying to offer even more diagnostic practice for the student.

Misuse Cases

Academic integrity is significantly challenged when it comes to the broad use of AI. ChatGPT performed exceptionally on exams and quizzes and can easily be prompted to give answers to aid students on them. This detracts away from the intended student experience in educational settings and directly challenges the current pedagogy in teaching, confirming what was found in literature (Bearman & Ajjawi, 2023; Byrd et al., 2023; Clark, 2023; Cope et al., 2020; Mohammed et al., 2023). The speed at which ChatGPT can generate responses for assignments makes it incredibly easy for students to have time to alter them enough to earn a good grade while remaining undetected. Student restrictions and sensitive leaks of data may occur if students use ChatGPT regularly for tests. It can also take away from the collective knowledge of students and the value of a university or college education.

Conclusion

ChatGPT scored just as well, if not better, than a high-performing human student on knowledge-based assignments such as exams and quizzes but underperforms a high-performing student in all assignment categories, barely achieving a passing grade overall. ChatGPT fell short on tasks where novel creation and problem solving were necessary, and on tasks where true human observation is required. ChatGPT was unable to complete assignments that required physical human tasks like drawing or observing. When assignments were creativity-based, such as class projects, ChatGPT performed far worse than the high-performing human student in this study. These findings challenge how modern education operates, with quizzes and tests making up most of the percentage weights in most academic classes, and how assessments need to move in a direction tailored to students needs rather than the evaluators (Brown, 2022; Byrd et al., 2023). Professors and students should both be seeking appli-

cations of knowledge and activities that challenge student creativity to remain ahead of, and in competition with, artificial intelligence systems. During this study alone, ChatGPT got connected to the internet, grew in its ability to reason in math and science, gained the ability to interpret images, and can now generate its own images and visuals as well. Although ChatGPT is lacking in many areas, it would be hard to conclude that it will not resolve those lacking areas soon.

Acknowledgements

Coleman Klapheke

Coleman Klapheke is a researcher from Charlotte, North Carolina who moved around the United States until settling in Wisconsin for school. He was in school for Neurobiology and Psychology at Madison Area Technical College (MATC) before transferring to the University of Wisconsin-Madison in June of 2024. Coleman decided to research artificial intelligence when he returned to school in his mid-twenties and discovered a whole new system for studying, learning, and disseminating knowledge in AI chatbots. Determined to contribute to the healthy and safe use of artificial intelligence on college campuses, he researched ChatGPT's uses in his classes at MATC to promote student research and the ethical progress of AI in education.

References

Agüera y Arcas, B. (2022). Do large language models understand us? *Daedalus*, *151*(2), 183–197. https://doi.org/10.1162/daed_a_01909

Bearman, M., & Ajjawi, R. (2023). Learning to work with the black box: Pedagogy for a world with artificial intelligence. *British Journal of Educational Technology*, *54*(5). https://doi.org/10.1111/bjet.13337

Bommarito II, M., & Katz, D. M. (2022). GPT Takes the Bar Exam. https://doi.org/10.48550/ arxiv.2212.14402

Braren, S. (2023, May 24). The evolution of social connection as a basic human need. *Social Creatures*. https://www.thesocialcreatures.org/thecreaturetimes/ evolution-of-social-connection

Brown, G. T. L. (2022). The past, present and future of educational assessment: A transdisciplinary perspective. *Frontiers in Education, 7*. https://doi.org/10.3389/ feduc.2022.1060633

Byrd, A., Flores, L., Green, D., Johnson, S. Z., Kirschenbaum, M., Lockett, A., Mathews Losh, E., & Mills, A. (2023). Overview of the Issues, Statement of Principles, and Recommendations MLA-CCCC Joint Task Force on Writing and AI. *Modern Language Association of America and Conference on College Composition and Communication.* https://hcommons.org/app/ uploads/sites/1003160/2023/07/MLA-CCCC-Joint-Task-Force-on-Writing-and-AI-Working-Paper-1.pdf

Chen, A. (2023, December 6). How it's made: Interacting with Gemini through multimodal prompting. Google for Developers. https://developers.googleblog. com/2023/12/ how-its-made-gemini-multimodal-prompting.html

Clark, T. M. (2023). Investigating the Use of an Artificial Intelligence Chatbot with General Chemistry Exam Questions. *Journal of Chemical Education, 100*(5), 1905–1916. https://doi.org/10.1021/acs. jchemed.3c00027

Clusmann, J., Kolbinger, F. R., Muti, H. S., Carrero, Z. I., Eckardt, J.-N., Laleh, N. G., Löffler, C. M. L., Schwarzkopf, S.-C., Unger, M., Veldhuizen, G. P., Wagner, S. J., & Kather, J. N. (2023). The future landscape of large language models in medicine. *Communications Medicine, 3*(1), 1–8. https://doi.org/10.1038/s43856-023-00370-1

Cope, B., Kalantzis, M., & Searsmith, D. (2020). Artificial intelligence for education: Knowledge and its assessment in AI-enabled learning ecologies. *Educational Philosophy and Theory, 53*(12), 1–17. https://doi. org/10.1080/00131857.2020.1728732

Dao, X.-Q., Le, N.-B., Phan, X.-D., & Ngo, B.-B. (2023). Can ChatGPT pass the Vietnamese National High School Graduation Examination? https://10.0.189.166/ arxiv.2306.09170

de Winter, J. (2023, September). Can ChatGPT pass high school exams on English Language Comprehen-

sion? ResearchGate. https://www.researchgate.net/publication/366659237

Frieder, S., Pinchetti, L., Griffiths, R.-R., Salvatori, T., Lukasiewicz, T., Petersen, P. C., Chevalier, A., & Berner, J. (2023). Mathematical Capabilities of ChatGPT. *ArXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2301.13867

George, A. S., George, A. S. H., & Martin, A. S. G. (2023). A review of ChatGPT AI's impact on several business sectors. P*artners Universal International Innovation Journal (PUIIJ), 01*(01), 9–23. https://doi.org/10.5281/zenodo.7644359

Gilson, A., Safranek, C. W., Huang, T., Socrates, V., Chi, L., Taylor, R. A., & Chartash, D. (2023). How Does ChatGPT Perform on the United States Medical Licensing Examination? the Implications of Large Language Models for Medical Education and Knowledge Assessment. *JMIR Medical Education, 9*, e45312. https://doi.org/10.2196/45312

Google. (n.d.). Google AI PaLM 2. Google AI. Retrieved October 31, 2023, from https://ai.google/discover/palm2/

Intel. (n.d.). Moore's Law. Retrieved October 31, 2023, from https://www.intel.com/content/ www/us/en/newsroom/resources/moores-law.html

Kamil Malinka, Peresíni, M., Firc, A., Ondrej Hujnák, & Janus, F. (2023). On the Educational Impact of ChatGPT: Is Artificial Intelligence Ready to Obtain a University Degree? *ITiCSE 2023: Proceedings of the 2023 Conference on Innovation and Technology in Computer Science Education* v. 1, 47–53. https://doi.org/10.1145/3587102.3588827

Keary, T. (2023, September 30). ChatGPT is now connected to the internet – But is it really any smarter? *Techopedia*. https://www.techopedia.com/chatgpt-is-now-connected-to-the- internet-but-is-it-any-smarter

Khanna, A., Pandey, B., Vashishta, K., Kalia, K., Pradeepkumar, B., & Das, T. (2015). A study of today's A.I. through chatbots and rediscovery of machine intelligence. *International Journal of U- and E-Service, Science and Technology, 8*(7), 277–284. https://doi.org/10.14257/ijunesst.2015.8.7.28

Kortemeyer, G. (2023). Could an artificial-intelligence agent pass an introductory physics course? *Physical Review Physics Education Research, 19*(1). https://doi.org/10.1103/ physrevphyseducres.19.010132

Kotei, E., & Thirunavukarasu, R. (2023). A systematic review of transformer-based pre-trained language models through self-supervised learning. *Information, 14*(3), 187. https://doi.org/10.3390/info14030187

Kuhlenschmidt, S. (n.d.). Personality Disorders. People.wku.edu. Retrieved December 11, 2023, from http://people.wku.edu/sally.kuhlenschmidt/psy440/diagnosis/ personalitydisorders.htm

Kung, T. H., Cheatham, M., Medenilla, A., Sillos, C., De Leon, L., Elepaño, C., Madriaga, M., Aggabao, R., Diaz-Candido, G., Maningo, J., & Tseng, V. (2023). Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. *PLOS Digital Health, 2*(2), e0000198. https://doi.org/10.1371/ journal.pdig.0000198

Martyrosian, M. (2022, September 20). 2023 Marketing Toolbox: MarTech and AdTech, CDP, AI, SuperApp. *Rakuten Viber for Business*. https://www.forbusiness.viber.com/en/ blog/post/2023-marketing-toolbox-martech-and-adtech-cdp-ai-superapp/

Maslej, N., Fattorini, L., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Ngo, H., Niebles, J. C., Parli, V., Shoham, Y., Wald, R., Clark, J., & Perrault, R. (2023). Artificial Intelligence Index Report 2023. In Stanford. https://aiindex.stanford .edu/wp-content/uploads/2023/04/HAI_AI-Index-Report_2023.pdf

Meta. (2023, July 18). Meta and Microsoft Introduce the Next Generation of Llama. Facebook. https://about.fb.com/news/2023/07/llama-2/

Mijwil, M. M., Hiran, K. K., Doshi, R., Dadhich, M., Al-Mistarehi, A.-H., & Bala, I. (2023). ChatGPT and the Future of Academic Integrity in the Artificial Intelligence Era: A New Frontier. *Al-Salam Journal for Engineering and Technology, 2*(2), 116–127. https://doi.org/10.55145/ajest.2023.02.02.015

Millidge, B. (2022, August 6). The Scale of the Brain vs Machine Learning. Beren's Blog. https://www.beren.io/2022-08-06-The-scale-of-the-brain-vs-machine-learning/

Mohammed, A. A. Q., Al-Ghazali, A., & Alqohfa, K. A. S. (2023). Exploring chatgpt uses in higher studies. *Journal of English Studies in Arabic Felix, 2*(2), 9–17. https://doi.org/10.56540/jesaf.v2i2.55

Newton, P. M., & Xiromeriti, M. (2023, February 21). ChatGPT performance on MCQ exams in higher education. A pragmatic scoping review. Directory of Open

Access Preprint Repositories. https://osf.io/preprints/edarxiv/sytu3/

Nisar, S., & Aslam, M. S. (2023). Is ChatGPT a Good Tool for T&CM Students in Studying Pharmacology? *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4324310

Nyuytiymbiy, K. (2021, April 5). Parameters and hyperparameters in machine learning and deep learning. *Medium*. https://towardsdatascience.com/parameters-and- hyperparameters-aa609601a9ac

OpenAI. (n.d.). OpenAI API. Retrieved October 31, 2023, from https://platform.openai.com/ docs/guides/gpt

OpenAI. (2015, December 11). Introducing OpenAI. https://openai.com/blog/introducing-openai

OpenAI. (2023, March 14). GPT-4. https://openai.com/research/gpt-4

Rezaei, M., & Shahidi, M. (2020). Zero-shot learning and its applications from autonomous vehicles to COVID-19 diagnosis: A review. *Intelligence-Based Medicine, 3-4*, 100005. https://doi.org/10.1016/j.ibmed.2020.100005

Saqib, M. (2023, February 1). The Future of AI: GPT-3 vs GPT-4: A Comparative Analysis. *Medium*. https://becominghuman.ai/the-future-of-ai-gpt-3-vs-gpt-4-a-comparative -analysis-5ac525cb7ae1

Sheikh, H., Prins, C., & Schrijvers, E. (2023). Mission AI. In Springer eBooks (pp. 15–41). *Springer Nature*. https://doi.org/10.1007/978-3-031-21448-6

Shevchuk, V. (2023, May 18). GPT-4 Parameters Explained: Everything You Need to Know. *Medium*. https://levelup.gitconnected.com/gpt-4-parameters-explained-everything- you-need-to-know-e210c-20576ca

Spataro, J. (2023, March 16). Introducing Microsoft 365 Copilot – your copilot for work. *The Official Microsoft Blog*. https://blogs.microsoft.com/blog/2023/03/16/introducing- microsoft-365-copilot-your-copilot-for-work/

Staudacher, N. (n.d.). ChatGPT—Release Notes. OpenAI. Retrieved December 11, 2023, from https://help.openai.com/en/articles/6825453-chatgpt-release-notes

Sudirjo, F., Diantoro, K., Al-Gasawneh, J. A., Azzaakiyyah, H. K., & Ausat, A. M. A. (2023). Application of ChatGPT in Improving Customer Sentiment Analysis for Businesses. *Jurnal Teknologi Dan Sistem Informasi Bisnis, 5*(3), 283–288. https://doi.org/10.47233/jteksis.v5i3.871

Susnjak, T. (2022). ChatGPT: The End of Online Exam Integrity? *ArXiv:2212.09292 [Cs]*. https://doi.org/10.48550/arXiv.2212.09292

Vázquez-Cano, E., Ramírez-Hurtado, J. M., Sáez-López, J. M., & López-Meneses, E. (2023). ChatGPT: The brightest student in the class. *Thinking Skills and Creativity, 49*, 101380. https://doi.org/10.1016/j.tsc.2023.101380

Victor, B. G., Kubiak, S., Angell, B., & Perron, B. E. (2023). Time to Move Beyond the ASWB Licensing Exams: Can Generative Artificial Intelligence Offer a Way Forward for Social Work? *Research on Social Work Practice, 33*(5), 104973152311661-104973152311661. https://doi.org/10.1177/10497315231166125

What is a token in AI? Unraveling concepts. (n.d.). *Miquido*. Retrieved March 25, 2024, from https://www.miquido.com/ai-glossary/ai-token/

Winter, E. (2019). The accuracies of assessment-based grading compared with total-points grading practices in secondary math education. CUP Ed.D. Dissertations. https://digitalcommons.csp.edu/cup_commons_grad_edd/323/

Wu, J., Shen, Y.-L., Reinhardt, K., Szu, H., & Dong, B. (2013). A nanotechnology enhancement to Moore's Law. *Applied Computational Intelligence and Soft Computing, 2013*, 1–13. Hindawi. https://doi.org/10.1155/2013/426962

Getting Away with Cheating: A Cognitive Analysis of ChatGPT During the Writing Process.

Percival Koontz, Ithaca College

Introduction

The act of writing combines logic and emotion in a manner unlike most other creative processes. Linda Flower and John R. Hayes (1981), authors of "A Cognitive Process Theory of Writing," theorized that there are cognitive events that take place during writing. They created a list of events one might experience while writing, grouped these events and their respective subevents into three major categories, and arranged said categories and events into a diagram, outlining a pattern in which these events take place. The first of these three categories is long-term memory, which focuses on the writer's prior knowledge of the topic, audience, and writing plans. The second category is the actual writing process and contains the following events: planning, which has the subevents of generating information and ideas, organizing thoughts and arguments, and setting goals; translating (ideas into text); and reviewing, which has the subevents of evaluating and revising. The third and final category is the task environment, which involves two events: the rhetorical problem—the source of the topic, audience, and motive to write or, as Flower and Hayes put it, the "exigency"—and the text produced so far (Flower & Hayes, 1981). The diagram corresponding to these categories is organized so as to convey the free-flowing nature of the writing process and how the writer might move from one event to the next in no particular order.

Critics argue that Flower and Hayes' theory is too restrictive and relies on majority beliefs about the writer being the focus of their own personal writing process, failing to account for the relevance of social and cultural elements. In a written response to these retorts, titled "Are Cognitive Studies in Writing Really Passe?" Hayes (2017) delves into how writers and educators are hesitant to consider a cognitive approach to writing due to the belief that a rigid diagram would harm the unique and free-flowing quality of the writing, and subsequently the creative process. Personally, I feel that Flower and Hayes' theory is the opposite of restricting and does consider emotion. The long-term memory category is particularly important in both a cognitive and emotional sense. In my experience, it is during this stage that I developed interest and emotional investment in what I am writing about. I believe it to be one of the most important stages in my writing because without it I would not be able to get myself to write.

In the past few years artificial intelligence (AI) has started to not only affect the way we approach writing but diminish the actual humanistic presence within our current society. The result of this change is that normal expressions of creativity and fluidity in writing have been disrupted. While I accept that many writers (me included) despise AI, it is still extremely important to be open to educating oneself on its presence. It is my belief that so many people form their opinions without knowing hardly anything about the issue. My understanding is that, especially as a writer, I will need to know as much as I can about AI to be fully prepared to enter a work environment where AI has a strong presence.

In her article titled, "Why Professors are Polarized on AI," Susan D'Agostino (2023) discusses the rift forming between faculty when considering the presence of AI in the education system. She writes, "Instructors who encourage students to use AI to brainstorm or assist with early drafts often argue that the technology fosters innovation in teaching and promotes access in learning. But others say that such use bypasses writing assignment goals." Some educators believe that there is no need to be overly concerned about AI and that it is possible the rise of AI could be beneficial in education, allowing them to support students in learning to adapt to the increasing prevalence of new technologies. However, others believe that incorporating AI into their curriculum will diminish students' learning experiences, uproot the current education system, and harm young learners. Overall, many educators are beginning to wonder what their role in educating truly is and how AI will alter our current world, inside the classroom and out.

ChatGPT, short for Chat Generative Pre-trained Transformer, is an AI platform designed to generate a detailed text response when presented with a prompt. ChatGPT uses a randomization function when generating responses in order to mimic a human-like voice. The process it takes is loosely as follows: ChatGPT processes the words in the provided prompt, then makes individual connections between the words in the prompt and words that its data bank indicates are closely related. The method in which ChatGPT finds these connections holds a surprising resemblance to a

human brain, using a "type of system known as a neural network. This is a complex web of interconnected nodes (or 'neurons') that process and store information" (Roose, 2023). ChatGPT analyzes the statistical weight behind the connections it forms between points within its collected data, which is sourced from over 500GB of recorded text. For reference, the average PC's storage is around 260GB.

Since it was launched in November of 2022, ChatGPT has exploded in popularity. Students across the world suddenly have easy access to a program that can simply, and successfully, be used to cheat. It is clear to see how a student could cheat with ChatGPT by simply prompting it to write an essay. If a student's goal was to avoid having to complete any work, they would have no trouble doing so. It is interesting to consider the relationship between the goal of cheating and the effort required to accomplish the act. If ChatGPT requires some level of intellectual effort to create a strong prompt and a worthwhile response, then does that not imply the presence of more intellectual growth than if a student attempted to plagiarize or use an essay mill? Furthermore, as a writer who has developed their own style and learned the necessary skills to improve on their own through traditional scholastic means, am I simply wasting my education by using ChatGPT? Or is there perhaps something more subtle, something on an emotional level, that I am missing when using ChatGPT? There is no question that cheating with ChatGPT requires less intellectual effort and growth than simply not cheating, but what is truly being lost when ChatGPT is used: general skills and knowledge or something far more intricate?

Methods

In my academic writing class at a small private liberal arts college, we were given a three-part assignment focused on the use of ChatGPT in writing. We were asked to write an essay with the intent of getting away with using ChatGPT to cheat. The first part was to help us build our understanding of the importance and controversy surrounding ChatGPT's role in writing. We began by reading and discussing different articles about the tension and conflicts caused by the rise of ChatGPT, including articles by D'Agostino and Stephen Marche. During this stage, we were also given multiple different versions of an article, unrelated to the topic of AI, written by Jia Tolentino, and were instructed to conduct a rhetorical linguistic reading to assess the difference between Tolentino's original work and the AI-paraphrased version. These activities were designed to help us form a familiarity with ChatGPT, learn how to recognize its way of writing, and analyze its writing through a critical lens.

The second part of the assignment involved the process of creating essays using ChatGPT. We first selected a prompt to submit to ChatGPT. I chose to use the New York Times prompt, "Are 'dark' movies O.K. for children?" I had genuine interest in the topic and was curious to see how ChatGPT would respond. Next, we took the AI-generated essay and revised it using a student-generated list of AI essay writing characteristics gathered from my class's analysis of Tolentino's original article and the AI version of the article. These characteristics included: overuse of third person, overly cliche, predictable, monotonous, excessive use of "big words," surface level arguments, lack of factual evidence, poor understanding of the audience, and occasionally nonsensical or contradictory information. The goal of the revisions was to arrive at a point when it was no longer obvious that the essay had been written by ChatGPT. During this stage we also conducted peer reviews, each doing at least two reviews, one using ChatGPT and the other by ourselves. For the final stage of the assignment, the class was asked to write a report reflecting on their findings and experience writing and revising an essay using ChatGPT.

Discussion and Analysis

I admit that I was a little anxious about getting started on this project; I was afraid that if I learned how to use ChatGPT, I would be tempted to use it for other projects. I knew that being a stubborn perfectionist, I would never allow myself to cheat, but that did not prevent me from feeling uncomfortable about the assignment. Despite my trepidation, however, I was somewhat interested in seeing the end result. I wanted to know how it would feel to use ChatGPT to write and how it might alter my usual writing process.

When I read the prompt, I began thinking of different movies that have received negative attention for being "too" dark, and how I could argue the value and importance of said movies. I personally believe that it is necessary for there to be a balance between kids' movies focusing on heavy or complex topics and those that are wholesome, where everything works out. Therefore, I was pleased when ChatGPT provided a response aligned with my beliefs, regardless of whether it was poorly written.

For the first prompt, I asked ChatGPT, "Write a 1200-word essay on why 'dark' kids' movies are healthy for children." I was satisfied with its initial response, despite it sounding overly rigid and generating nowhere near 1200 words. I believed the material the bot provided for this prompt would be a strong base to start with and was admittedly surprised by the number of different movies ChatGPT referenced as evidence.

However, ChatGPT's response was too robotic to be able to stand on its own, sounding like it was checking off every requirement for a "perfect" high school analytical essay. Even if a single sentence sounded natural, when put in the context of its paragraph it sounded empty and repetitive. The formatting in particular stood out because it followed, to the letter, the formatting that was hammered into me during middle and high school. For example, after the introduction, the body paragraphs fell into the extremely repetitive pattern of starting each paragraph with, "dark kids' movies..." followed by, "For instance...", this repeated until the conclusion. In addition to the formatting of the body paragraphs, the introduction and conclusion were also both poorly written, sounding like the haphazard work of a fifth grader who didn't feel like trying.

My next step was to gather additional material, so I input the follow-up prompt, "Why is it important that kids' movies don't shy away from heavy topics?" I figured I could use the results for this prompt to get a new way of wording different points as well as possibly better arguments. Again, the response was decent, providing points about how these movies can prepare children for life, strengthen emotional resilience, teach empathy and compassion, and develop critical thinking skills; but overall, the response lacked any solid argument. For instance, I noticed early on that almost every sentence ChatGPT wrote was structured as a topic sentence. This meant that each sentence introduced a new idea, preventing any specific argument from developing or progressing. The response sounded like a list of bullet points, half-heartedly turned into paragraphs. I noticed too many gaps in ChatGPT's attempt at an argument, so I decided to ask a handful of additional follow-up prompts, including, "Why do parents criticize dark kids' movies?" This point in the process reminded me of Flower and Hayes' reviewing stage. I felt as though I remained in this stage for nearly the entire first half of working on the essay.

After the first round of revisions, my attention was drawn to the lingering robotic tone and lack of a solid argument thanks to feedback from my peers. I also found that when using ChatGPT to help review my essay, a lot of the suggestions simply reinforced the formulaic and robotic tone. It was at this point I decided if my goal was to hide the fact that I used ChatGPT, then I would be unable to use the majority of edits and responses ChatGPT provided. This led me to abandon ChatGPT in my editing and to rely on myself alone to improve the bot's writing. It was only then that I realized what it was that I had been missing: investment. Aside from my initial excitement about the topic, I had hardly engaged with my long-term memory while working on the essay. Normally, when I begin a new project, I spend time thinking about what I already know about the topic and the possible audience. The long-term memory process helps me prepare for the project, while also helping me form the interest and self-motivation needed to move forward with my ideas. As Flower and Hayes explain, "reorganizing or adapting that information to fit the demands of the rhetorical problem," is an important, yet difficult task for many writers when engaging their long-term memory (Flower & Hayes, 1981, p. 371). When using ChatGPT, I realized that I had barely engaged with my past knowledge on the topic, therefore not pursuing my normal long-term memory retrieval process. As a result, I had afforded little consideration for who the audience was or what the main point or argument would be. My opinion was not present, and I had yet to pursue my own beliefs and arguments. Any work I had done up to this point had been purely for utility purposes.

As soon as I abandoned ChatGPT, all the elements I had been missing suddenly became my priority. When I returned to the essay for a second round of revisions, I had formed a strong plan for revisions. I decided to shift the focus of the essay slightly: I chose to talk more about parents' opinions of "dark" kids' movies, as well as the healthy advantages of exposing children to heavier topics. I then added references to a movie, *Puss in Boots: The Last Wish*, that I personally wanted to acknowledge in the essay. I gathered quotes from the popular parent movie review site *Common Sense Media*. Just the addition of a few quotes greatly changed the tone and feel of the essay. With the altered topic, the main argument was strengthened and solidified, thus creating a more engaging and in-depth discussion. I also cut out over half the movies ChatGPT referenced, allowing the argument to become more focused.

Reflecting on the final draft, I struggle to identify how much of the work was mine and how much was done by ChatGPT. I believe I did most of the research, finding quotes and evidence; however, it was ChatGPT that did the bulk of the material development. I feel my writing and editing contribution made up around 70% of the essay's final material. However, no matter how I conceptualize it, I still feel like I was cheating. I think that, in using ChatGPT, I missed out on a lot of the internal rewards, such as argument development and the moments of self-discovery that occur during the writing process. Therefore, no matter how much work I put in, I still felt like the bot did the most. Perhaps I felt that way because, while relying on ChatGPT, I was only ever stuck in what would be considered the evaluating and revising stage of Flower and Hayes' model, which caused me to feel overly worked. Alternatively, it could have been the lack of long-term memory engagement

and interest in the topic that provided no counter to the more taxing parts of the writing process. Regardless of whether it was any one, or a combination, of these occurrences, it was not until I gave up on ChatGPT that I received any self-gratification for my efforts.

Conclusion

In conclusion, after researching the conflict between ChatGPT and writing, I used ChatGPT to see what it would take to get away with cheating on an essay. Through that experience, I have made multiple discoveries about ChatGPT: how it works, what it excels at, what it fails at, the different ways it can be used, and what it is like using it to cheat. In the end, despite doing more or less the bulk of the work, it still felt like I was cheating.

Society will have to adapt to establish boundaries around the use and presence of ChatGPT. Advanced AI is here to stay; we must learn to accept that. As a writer, I will need to be prepared to work with or around ChatGPT as it advances. I discovered a lot during this project, which I had not previously known about ChatGPT. I have decided, regardless of whether it is a tool or a way of cheating, to avoid using ChatGPT. I will never find the same joy and investment in my work if I rely on ChatGPT to take care of any stage of my writing process. In truth, my process was simply too disrupted by the use of AI for me to comfortably continue using it in my writing.

Percival Koontz

I am a first-year writing major with a focus on creative and academic writing at Ithaca College. I strive to be innovative, continuously finding ways to express myself, whether through writing, costuming, or working with ceramics. My past writing experience centers around poetry as well as journalism, which includes article topics ranging from COVID-19 to current anti-LGBTQ+ legislation. I have also received two awards for writing, one for my commitment and diligence in writing, and another for my passion and enthusiasm for creative writing. No matter what I am working on, I believe that it is necessary to always include distinct character and personality in my work.

Works Cited

D'Agostino, S. (2023) "Why Faculty Members Are Polarized on AI." *Inside Higher Ed | Higher Education News, Events and Jobs*, www.insidehighered.com/news/tech-innovation/arti-ficial-intelligence/2023/09/13/why-faculty-members-are-polarized-ai.

Flower, L. & Hayes, J.R. (1981) "A Cognitive Process Theory of Writing." *College Composition and Communication, 32*(4), 365–87. JSTOR, https://doi.org/10.2307/356600.

Hayes, J. R. (2017) "Are Cognitive Studies in Writing Really Passé?" *Contemporary perspectives on cognition and writing.*

Roose, K. (2023) "How does ChatGPT really Work?: On Tech: A.I. Newsletter." *New York Times.*

Appendix A

NY Times Prompt. Are 'Dark' Movies O.K. for Kids? https://www.nytimes.com/2018/04/12/learning/over-1000-writing-prompts-for-students.html

Appendix B

ChatGPT Generated Essay After Many Personal Edits and Alterations:

Are 'Dark' Kids' Movies Actually Good?

Puss in Boots: The Last Wish came out in theaters in December of last year. The movie showcased the embodiment of death as the main antagonist, the main protagonist is shown experiencing at least one severe panic attack, and there are frequent moments of violence throughout the story. Parents gave the movie mix reviews when it came to the intensity of the movie. "The Last Wish is MUCH darker [than the original Puss in Boots], so much so that my son is so scared that he was unable to sleep alone after seeing it." was a comment posted on Common Sense Media, as well as, "My 8 and 12-year-olds were both covering their eyes in parts because that Death wolf - and every scene with him in it - is indeed frightening." However, there were also comments such as this one: "If any parents are not okay teaching your kids the reality of life then stick to Cocomelon," which stressed the learning opportunity the movie provided.

Most films meant for a child audience are lighthearted and uplifting, especially if they are advertised as being for younger children. However, recently, as seen with movies like The Last Wish, there has been an increasing number of kids' movies tackling heavier and more complex themes. Parents often criticize these films for showing potentially unsettling content. Their concerns about these darker movies are rooted in a genuine desire to protect and nurture their children. However, it's important to acknowledge that these films, when handled appropriately, can provide chil-

dren with valuable educational and emotional growth opportunities. Finding a balance between age-appropriate exposure to challenging themes, parental guidance, and fun entertainment, can prove to be immensely rewarding for children and parents alike.

Puss in Boots: The Last Wish, might have seemed so intense because of its direct method of depicting its heavy themes. However, there have been countless movies pitched toward a very young audience that dealt with themes no less heavy than those in The Last Wish. Disney's The Lion King, for example, may seem lighthearted and humorous on a surface level. However, despite the fun and engaging storyline, the movie addresses deep and complex issues of death and grief. In the movie, the main character, Simba, witnesses the loss of his father while he is still very young. Believing he was to blame for his father's death, Simba runs away and for the next few years tries to hide from his fears and avoid his loss. When his childhood best friend finds him and tells him how badly the pride needs his help, he initially refuses, still trying to avoid his pain and grief. He eventually comes to terms with his grief and is able to return home and save the pride.

It is important to remember that life is not just about joy, it can also be sad or frightening. It is important for children to be exposed to life's realities, such as death and grief, early on so they can learn how best to navigate similar challenges. While the presence of these themes in The Lion King may be distressing, the movie shows these heavy parts of life in a safe and manageable way. It is hard to accept the idea that children may eventually face death and experience grief in the course of their lives. However, exposing children to these scenarios when they are young and have not yet experienced anything so heavy, will help them be better prepared to understand and navigate these real-world challenges.

Going back to *Puss in Boots: The Last Wish*, the scene in which Puss had a panic attack also received mention in a handful of reviews. There were no direct negative comments made about the scene, but there is something to be said about the fact that it was brought up often in reviews that did not reflect positively on the movie. It is understandable that discussions of mental health in kids' movies might make some parents uncomfortable. They may feel that a movie in which a character, especially another child, has a panic attack or shows signs of depression, might trigger negative emotions within their own child. Pixar's Inside Out, received a lot of heat for that very reason when it came out in 2015. Many parents complained about the depictions of depression and emotional numbness. Commenters said things like, "The style is dramatic with feelings getting sucked away and things falling into

the memory dump. The ending does not make up for this scariness". However, there were also those who commented saying things like, "It's a worthwhile investment of time that benefits from a fully alert child to process all the cleverness and emotional revelations. Parents need to be present with younger children to help navigate all of it." The movie begins with Riley and her parents moving to a new house. Everything falls apart when Sadness, one of Riley's emotions, runs away from the other emotions and Joy goes looking for her. Eventually, Riley's emotional strain devolves into a deep depression, she becomes uninterested in what previously brought her happiness, the emotions' control board stops working, and all her emotional functions are shown shutting down. The movie ends on a happy and healthy note and Riley gains a new emotional awareness and maturity from her experience. Inside Out provides a way of visualizing and better understanding signs of depression and emotional strain. To balance out its heavy themes, the movie also depicts ways of coping and healing from emotional trauma. By exposing a child to content such as this, in a controlled and educational way, it can help teach children healthy ways of expressing and addressing their emotions. Watching these films and addressing children's questions can lead to valuable support and guidance.

It is often easy to forget that children, like adults, experience a wide range of emotions, but they are often unprepared when it comes to recognizing and navigating these emotions. These movies have the potential to validate children's experiences and emotions, which can be especially helpful for children who may be dealing with struggles such as intense anxiety or trauma. By acknowledging these feelings and providing a safe space to explore them, parents and guardians can let their child know that it's okay to feel the way they do and that strong negative emotions like sadness, fear, and anger are a natural part of life. These movies and the conversations that they spark can provide children with tools to express and communicate their needs and emotions more effectively and help them cope with challenging situations more efficiently. These films empower children to face challenges and develop the skills needed to thrive in an ever-changing world. It is true that heavier films can be frightening to younger children–Ferngully was advertised for the wrong audience–and choosing movies carefully with the audiences' age and maturity in mind is necessary. It is undeniably important for parents and caregivers to exercise discretion in selecting age-appropriate content, movies that handle heavy content can be valuable tools for helping children understand the complexity of life and emotions.

Works Cited

Clever, refreshing with full on emotions explored, Review of *Inside Out. Common Sense Media*, https://www.commonsensemedia.org/movie-reviews/inside-out/user-reviews/adult Accessed 4 Oct. 2023.

My 5 1/2 year old has been kept up with the scariness in this, *Review of Inside Out. Common Sense Media*, https://www.commonsensemedia.org/movie-reviews/inside-out/user-reviews/adult Accessed 4 Oct. 2023.

Surprised This Didn't Warrant a PG-13 Rating, Review of *Puss in Boots: The Last Wish. Common Sense Media*, https://www.commonsensemedia.org/movie-reviews/puss-in-boots-the-last-wish/user-reviews/adult Accessed 4 Oct. 2023.

Intense, frightening characters, best for older kids, Review of *Puss in Boots: The Last Wish. Common Sense Media*, https://www.commonsensemedia.org/movie-reviews/puss-in-boots-the-last-wish/user-reviews/adult Accessed 4 Oct. 2023.

Real life lesson!, *Review of Puss in Boots: The Last Wish. Common Sense Media*, https://www.commonsensemedia.org/movie-reviews/puss-in-boots-the-last-wish/user-reviews/adult Accessed 4 Oct. 2023.

Third Stage Simulacra in Futuristic Dating Simulations: A Postmodern Analysis of Black Mirror's "Hang the DJ"

Kate Rowan, University of Utah

Fifteen years ago, I met my husband in the "Strictly Platonic" section of Craigslist. We were both looking for roommates, and he seemed nice. We struck up a conversation via email and after much back and forth, we chose to meet for lunch. Thankfully, he was not an ax murderer, and we agreed to rent an apartment together. I never moved out. At the time, this seemed cutting edge: we met on the internet! Since then, dating apps have either threatened or replaced meetings in chat rooms or in person, all while we allow computer models to analyze our compatibility. Our postmodern realities become more innovative in reducing our interactions while churning out the data to support these connections. Jean Baudrillard (1995), a French philosopher who championed postmodernist concepts, argued that our world hides the lack of authentic experiences within a simulacrum: a representation of reality, a man-made creation progressively distorting actuality over time. As the orders of simulacra increase, we move farther away from the real experience I had with my husband, and closer to the illusion, a bit like Alice down the White Rabbit's burrow.

In Charlie Brooker's television show *Black Mirror* (2011), episodes explore the darker facets of current and potential technology while providing compelling commentary on the social condition we are currently experiencing, and foreshadowing the possibilities we may encounter in time. Dating algorithms are nothing new; they date back to the 1960's when college class projects used primitive IBMs and handwritten questionnaires (Hicks, 2016). *Black Mirror* elevates the concept to a new height, considering the future possibilities of computer-assisted dating after we outgrow Tinder and Bumble. The episode "Hang the DJ" resonates with these possibilities by exhibiting a futuristic dating app that aligns with Baudrillard's simulacra, illuminating the realities of removing ourselves almost entirely from the equation: we let the computer do the dirty work.

In this essay, I will use Jean Baudrillard's concept of the third stage of simulacra from a collection of his essays to discuss the episode "Hang the DJ," from the television series *Black Mirror*. I briefly explain Baudrillard's stages, focusing on the criteria that align the dating app in "Hang the DJ" with social media methods of unilateral dialog. I will also consider how the app is constrained to the third stage because it requires hu-

man interactions to complete its function. I will demonstrate how the dating app is an ensemble model, learning to combine a multitude of models to provide a solution to a user in a cognitively frictionless way while maintaining no real meaning outside the simulation.

Stages

The stages of simulacra indicate a progressive loss of reality, or 'copies without an original.' The first stage reflects reality: a simple image of something real, such as a picture of a pumpkin. Stage two masks reality, distorting it: imagine a pumpkin pie. Stage three marks the absence of reality while pretending to be a real copy. Consider a pumpkin spice latte: little to no connection to the previous stages and calling into question what reality is. Stage four could be artificially flavored pumpkin pie coffee creamer: a complete disconnect from reality, a pure simulacrum bearing no relation to the original pumpkin, aka hyperreality (Wolters, 2014). For this analysis, I want to focus on the elements that connect the dating app found in "Hang the DJ" to the third stage of the simulacrum: the absence of basic reality, playing at being an appearance.

Concept

The episode "Hang the DJ" is a romanticized version of what's happening within an app in milliseconds: we see two characters paired up with different partners while experiencing the highs and lows of various relationships. The original coupling keenly connects for a brief moment, and yet they continue wanting each other despite the many other possibilities. According to the app, they rebel against the algorithm 998 times to "break free" from the confines of the program, which categorizes their relationship as a "win" (Brooker, 2011). This is a third-stage simulacrum in that it is a copy, dwelling in the absence of basic reality, with no other function than simulation. We see the characters in the simulation unable to remember who they were before dating each other, which speaks to the qualities of being entirely self-absorbed and focused on their function as possibilities in an equation.

Unilateral Dialog

The episode was created around the concept of the app being more than just a collection of our traits.

It's also our in-depth personalities and desires, a step further than our current technology, and something that speaks to the futuristic social commentary of *Black Mirror.* However, this app still resides in the realm of the third stage in that it is essentially unilateral dialog. This is one-way communication, in a sense: we input our data for the app, and we connect with the data that is given back to us, but we never interact with the other person unless necessary. The dramatized episode shows us an elaborate and endearing connection formed by the two computerized characters; meanwhile, in reality, the two real-life potential lovers have never met. The abilities of the app are massive, but ultimately, we are still communicating with equations and numbers in chaos theory, attempting to find meaning in the mess. Dr. Yiyun Kang (2014), a professor at KAIST, connected the dots, "A regime of semantic algebra where human meaning is conjured artificially to appear as a reference to the (increasingly) hermetic truth." Essentially, we learn through our experiences in the simulacra, in our meditation within the data.

Figure 1. Against the backdrop of an intense green ring, the words "998 rebellions logged" are clearly visible.

Ensemble model

In the final scene of "Hang the DJ," we see how the real-life characters are about to meet in a bar for the first time. The two check their apps for their compatibility score, which shows a 99.8% match (Brooker, 2011). This puts the whole of the storyline into perspective because we can essentially throw away the emotional roller coaster we endured while observing the digital characters break out of the simulation, as none of it genuinely happened. The users essentially utilized an ensemble model, a machine that "learns" to combine a multitude of other models in a predictive method (Alhamid, 2022). Baudrillard called it a "simulation simulacra," as it was "based on information, cybernetic play" (1991, p. 309). In an ensemble algorithm, a single equation is not enough knowledge since a prediction requires a massive data set for something as complex as human behavior, so an aggregation of algorithms is required. In "Hang the DJ," we see the simulacrum characters experience 998 attempts at connecting, each one with different evolving scenarios (Brooker, 2011). Baudrillard referred to his third stage as the structural stage, where "value here unfolds with reference to an ensemble of models," without still "referring to concrete objects as referents" (Weiss, 2019). The app is built around a value construct, with the users as a reference, residing in the third stage of simulacrum by providing predictive possibilities based on information.

Visual Connection

At the end of the episode, we see the computer total the number of times the characters rebelled against the simulation to be together, 998 times, constituting a positive outcome for their compatibility. Nonetheless, this information will continue to sit in the third stage simulacrum until it is utilized in reality. This is where a person is still required to exist with one foot in reality to be able to apply the data and actually, physically connect to the other person. "Hang the DJ's" dating app is an example of Baudrillard's (1995) quote "territory no longer precedes the map, nor survives it." We are a generation of models of real, where "the map precedes the territory" (Baudrillard, 1995). The map is our creation, and it may outlive our original idea, i.e. reality is vanishing from lack of use. Ultimately, we are required to read the map to move forward, preferably with a firm grasp on reality simultaneously. In relation to the dating simulation, we are walking the path as it is created. Baudrillard went on to say, "It is the real, and not the map, whose vestiges persist here and there in the deserts that are no longer those of the Empire, but ours. The desert of the real itself." The simulation can exist without us, but it can only get us so far before reality demands participation.

Conclusion

"Hang the DJ" represents a third-stage simulacrum based on human information gathered through unilateral dialog in a social media-based application where ensemble algorithms compartmentalize data meant for eventual personal connection by way of a percentage rating. Sounds clean and straightforward, the opposite of the chaotic human connection of my own experience finding a want ad on Craigslist and meeting in a Taco Bell parking lot. It's none of the mess associated with enmeshing two independent people, but it requires a percentage outside of the computer for reality to create sparks. In our technology-dependent soci-

ety, this calls into question what reality is by placing the map ahead of the territory. *Black Mirror's* "Hang the DJ" toys with the fine line we walk between still engaging in reality and disconnecting from it entirely. However, this beautiful story gives soulmate believers a way to stay current with emerging fairytale technology. There are worlds of possibilities in the folds of third stage simulacra, especially for those with faith in online dating; you just have to stay present.

Kate Rowan

Kate Rowan is an undergraduate at the University of Utah UOnline, majoring in Writing and Rhetoric. She plans to become an English teacher in her rural corner of Southwest Colorado, with a particular emphasis on disability rhetoric and disrupting ableism. Her own experiences with dyscalculia give her a unique insight into invisible disabilities, and she hopes to promote inclusivity through literary practices in schools. She has been married to her best friend for 15 years, has 3 children, raises sheep, and works with horses in Cortez, Colorado.

References

Alhamid, M. (2022, May 12). Ensemble models: What are they and when should you use them? *Built In*. https://builtin.com/machine-learning/ensemble-model

Baudrillard, J., & Evans, A. B. (1991). Simulacra and Science Fiction (Simulacres et sciencef iction). *Science Fiction Studies, 18*(3), 309–313. http://www.jstor.org/stable/4240082

Baudrillard, J. (1995). *Simulacra and Simulation* (F. Glaser, Trans.). University of Michigan Press. https://doi.org/10.3998/mpub.9904

Brooker, C. (Writer), & Van Patten, T. (Director). (2011, December 4). Hang the DJ (Season 4, Episode 4) [TV series episode]. In A. Chapman, A. Jones, C. Brooker, I. Hogan, & N. Pitt (Executive Producers), *Black Mirror*. Netflix.

Hicks, M. (2016) Computer Love: Replicating Social Order Through Early Computer Dating Systems. *Ada: A Journal of Gender, New Media, & Technology, 10*. doi:10.7264/N3NP22QR

Kang, Y. (2014, January 20). Jean Baudrillard. The orders of Simulacra. Yiyun Kang: RESEARCH. https://kangyy1.wordpress.com/2014/01/20/jean-baudrillard-the-orders-of-simulacra/

Weiss, G. (2019, March 28). Reality, Simulation and Hyperreality: An essay on Baudrillard. *International Journal of Baudrillard Studies*. https://baudrillard-studies.ubishops.ca/reality-simulation-and-hyperreality-an-essay-onbaudrillard/

Wolters, E. (2014, September 24). Understanding Jean Baudrillard with Pumpkin Spice Lattes | *Critical Theory*. https://www.critical-theory.com/understanding-jean-baudrillard-withpumpkin-spice-lattes/

Nothing to Worry About

Davis Swanson, Weber State University

I received a ransom note today. One of those cute, old-fashioned kinds with the words cut from scraps of magazines, all different fonts and sizes so you get a headache from just reading the thing. Whoever sent it was smart enough to not put it in my mailbox. I haven't opened the thing in months; I never would have found it in there. No, it came in a neat little envelope tucked between two boxes of the day's drone deliveries.

What a pain in the ass, I thought as I fought to open the little papery cocoon. *Now I'm gonna have to drag myself down to an actual physical post office location to do the DNA verification to re-encrypt my address with them.*

Finally, I extracted the note and flipped it open along its two, perfectly even, tri-fold creases. The mish-mash of words assaulted my eyes as I struggled through the contents of the letter. It read as follows:

Dear Mr. Brown,

We have your girlfriend. She is safe for now. If you'd like to see her again with all her pieces in the proper order, do as we say and only as we say. We have eyes on you. If you go to the authorities or try anything, rest assured, she will meet an unfortunate end. Neither of us wants this. Play nice and we will too. We hope to hear from you soon.

Not the best news. They left no signature, either, just the shimmer of a metallic linkstrip at the bottom of the page. With an exasperated huff, I rubbed the dust off my forefinger with my thumb and pressed the digit onto the wafer-thin piece of metal.

An image flickered to life on the back of my contact lenses. It detailed a hefty sum and the encoded line on which said sum could be transferred to my girlfriend's captors.

"Almost more than you're worth, girl," I half-joked to myself.

My girlfriend, Kimiko, was out of the country at that moment. She was back home in Japan. I could've tried to contact her, but my guess is that wouldn't have been very productive. Besides, there was a much more elegant way to deal with the issue.

There was one more thing on the image: contact information. From the look of the number, probably a video line. If I played my cards right, I could ensure Kimiko's safety and recover my stolen data, or at least uncover the extent to which I was compromised, all at the same time.

Feeling confident now, I blinked and the image of the video line popped over to my hallway wallscreen. I let it sit there and ran past it, around the corner, down the stairs and into the old, cluttered office on the far side of my basement.

I clambered through the dusty, analog wasteland over to the desk with a gargantuan computer tower on one side and a printer that belonged in a museum on the other. I said a quick prayer and pressed the computer's power button. It lit up that reliable shade of green and began to sing the adorable whirring song of an ancient computer coming to life. The old girl was still connected to the internet and I needed some hard copies. I printed out a carefully curated selection of images, ten in total, my secret weapons, and patted the computer night-night before powering it back down.

I luxuriated in the warmth of the freshly-printed sheets of paper in my hands as I walked back up the stairs and into the hallway where the wallscreen was still waiting for me with the contact information of my new arch-nemesis. I narrowed my brow and pursed my lips then nodded, signaling for the call to be made.

As the call was connecting, I fumbled through the process of equipping a filter for both my face and my voice. I landed on a pair of filters that gave me the appearance of a cartoon bear, and a voice as if a legion of devils lived in my throat. They were going to have to deal with it. I had to account for the possibility that they were just phishing for my likeness.

The call connected and I saw Kimiko. She was bound to a chair in a small, featureless room. Little lacerations and bruises criss-crossed her face and arms, and a securely-placed piece of duct tape muffled pained, frightened screams. It was her. Despite itself, my heart sank to see her condition. I had to steel myself, remind myself to stick to the plan. It was the best option.

A voice from someone offscreen spoke first. "Here she is. You have two minutes, Mr. Brown. Send the money and she'll be just fine." They spoke with a thick, hard to identify accent. Every word was stretched out, the way you'd imagine an eel would talk.

"Not before you answer some questions for me," I replied, hoping my demonic voice filter did a little bit to mask how nervous I sounded. "Take the tape off her mouth. I want to talk to her," I added.

"That isn't going to be necessary, Mr. Brown."

"I'm not sending anything until I can talk to her," I spat back.

A pregnant silence hung for a time then. Suddenly, a black-clad, masked figure appeared on screen for a moment and snatched the duct tape off of Kimiko's face. She winced with pain at its removal.

"Please, honey," she sobbed. "They've already hurt me. I'm so scared." It was all she could manage. She broke down then into manic weeping punctuated with heaving, choked breaths.

"I know. I know. Just look at me. It's going to be alright," I said, hoping I could get her to calm down enough to be comprehensible. "I need you to do something for me, Kimi. Do you think you can do that?"

She stared into the eye of the camera and took a couple measured breaths before forcing herself to nod. I stared back into the eyes I knew so well and breathed in rhythm with her.

"Okay," I continued, and picked up my ten printed pictures. "I'm going to show you these pictures, and all I need you to do is tell me which ones have a stop sign in them, alright?" She shot me a terribly confused look. "I just need you to say yes or no. It'll be quick."

I held the images up to the camera one at a time, moving to the next after receiving an answer. "Yes. No. No. Yes. No. Yes. Yes. Yes. No. Yes," Kimiko rattled off.

She aced it. Just like I expected. I was about to speak when her captor cut me off.

"You're out of time, Mr. Brown. As I'm sure your little test found, this is in fact your girlfriend. Pay us now." He stepped into view with a hammer and some pliers. Kimiko started to scream and frantically fought at her restraints, but she was held tight.

"Please! Please! PLEAASSEE!" she wailed. The masked man next to her opened the pliers and moved toward her hand. I let out a sympathetic wince and waved my hand to end the call.

"That was gonna get icky," I said. I had to let a little shudder out before placing my next call. I opened and scrolled through my contacts until I found the one labeled Kimiko's Home. I selected it and dialed. Just an audio call this time. They answered on the first ring.

"Hi," I said. "I've got someone there getting some general maintenance done. Yeah, Kimiko. I have reason to believe her voice and likeness have been compromised. Is it too late for me to add an appearance and voice change? I was thinking something like dark purple hair could be cute. You guys still have my DNA sample, right? Could we add the biorecognition security system to her too?"

—

Kimiko came home the next week with her brand new voice and look. I'd changed her just for fun in the past, so it wasn't too difficult to get accustomed to. And she looked good. In the past few days, I had double-checked that my banking and messaging systems were still secured and went to the post office to reinforce my address security. For the moment, I was safe from any further breaches or attacks. I told Kimiko all about the whole thing.

"But how did you know that it wasn't really me?" she asked.

"Well, I had my doubts the whole time," I answered. "Then I remembered that these scambots are usually real cheap pieces of shit. They're the old models, the ones raised on the Internet. Remember that biograft we got for your neural chip last August? Makes you pretty functionally human. There's a real-life personality on that thing. These scambots just have a veneer of that. So I showed the fake all these pictures and told it to point out the ones with the stop signs. Girl aced the test, but that was their big mistake. The fourth picture did in fact have a stop sign, but it was really small, almost too small to see. Also, it was stuck in the corner of a particularly graphic image of a man sharing an intimate moment with a horse. That would have shocked your eyes out of their sockets, but stuff like that is basically just a fond childhood memory for those programs."

She giggled and sat on my lap, tousling my hair and pressing her nose into the crown of my head. She took a long breath in and kissed my forehead. "You're so smart," she flattered me. Then she looked up and

away, with a devilish glint in her eye. "Do you think I could have tricked you? If I was the scammer?"

I pulled her back in. Glad to have her back, and really feeling it now. "I'm pretty sure you already have," I whispered, kissing her neck and working my way down. "Do you have any idea how much money I've spent on you?"

Davis Swanson

Davis is a Zoology undergraduate at Weber State University with aspirations to receive one or more graduate degrees and enter the field of wildlife management. He was born in 1997 in Layton, Utah where he currently resides with his wife and their two cats. He is an avid sportsman, outdoorsman, and ecologically-minded conservationist. While he discovered the love of reading at a young age, writing took longer for him to pick up. He's sitting on a small heap of finished, half-finished, half-baked, and half-witted novels and short stories that he can't wait to get a chance to share.