**Abstract**

This project explored perceptions of ChatGPT in higher education among students and faculty to assess teaching and learning implications of this Generative Artificial Intelligence (Generative AI)–based novel tool. Two theoretical frameworks inspired the project, including Diffusion of Innovation (DOI) theory (Rogers, 1962) and Technology Acceptance Model (TAM) (Davis, 1989). An online survey was completed by 380 participants (*N* = 380).

Participants indicated that they would not use ChatGPT to plagiarize but believed others would. When asked to rate the accuracy of ChatGPT’s output, more than half took the incorrect output as correct/somewhat correct or could not tell whether it was correct or incorrect. Results varied based on participant demographics, including age, gender, and occupation.

These findings support the need for data literacy. If Generative AI is to be used in higher education to aid in the learning process, it is imperative to continue teaching critical thinking.

Keywords: Generative AI, ChatGPT, Survey, Data Literacy, Students, Faculty

**Introduction**

In early 2023, news outlets were extensively reporting on a powerful artificial intelligence (AI) tool called ChatGPT (O’Brien, 2023; *Reuters*, 2023; Jolly, 2023; Dotan, 2023). Developed by OpenAI, ChatGPT acquired one million users in five days and reached 100 million users two months after it was made public in November 2022; setting the record for the fastest-growing consumer application (Hu, 2023). This Generative AI chatbot was built on top of OpenAI’s foundation large language models (LLMs) (Murphy Kelly, 2023). ChatGPT is a technology that leverages deep learning, transformer-based, neural network models generating human-like content (e.g., images, words) in response to various prompts (e.g., languages, instructions, questions) (Lim et al., 2023).

Generative AI is not just an extractive question-answering system, providing a response elicited from a document to a question formulated by the user: It also generates content in that response, going beyond human-like interactions in conversational AI ([Lim et al., 2023; Lim et al., 2022](https://www.sciencedirect.com/science/article/pii/S1472811723000289#bib26)). For example, ChatGPT made headlines in early 2023 for passing law exams in four courses at the University of Minnesota and another exam at University of Pennsylvania’s Wharton School of Business (Murphy Kelly, 2023).

As a result, Generative AI has forced educators and institutions to reconsider how students learn (McMurtrie, 2023). Although Generative AI can be used for crafting intelligent tutoring systems that provide personalized assistance to students (Marr, 2023) and can provide a proactive pathway (Dowling & Lucey, 2023), it can also be viewed as a threat that needs to be regulated (Nature, 2023).

Therefore, it is important to examine the perceptions in higher education about Generative AI. The paper’s primary goal is to provide a quantitative assessment of the early perceptions of tools like ChatGPT among faculty and students in a higher-education setting. In satisfying the above objective, the paper makes two important contributions: (1) It measures how instructors and students in a college setting felt early on about this Generative AI tool. The opinions among faculty and students may certainly vary, and for this reason, it is critical to identify public perceptions as an important first step in determining how best to proceed; (2) It measures perceptions by faculty and students of how Generative AI tools like ChatGPT will affect teaching and learning over time. These tools will continue to develop and improve, making it imperative that higher education institutions critically consider how to navigate the uncharted waters of Generative AI.

The paper frames this goal and contributions in the spirit of two well-known theoretical frameworks: a) TAM—Technology Acceptance Model Davis, 1989) grounded in information systems research, that studies how users accept, adopt, and use new technology; and b) DOI—Diffusion of Innovation theory (Rogers, 1962) grounded in social science and the field of communication that describes how an idea or product acquires significance and spreads through a social system over time. TAM posits that the adoption of a new technology—i.e. the decision of how and when users will use it—is influenced by two critical factors: a) perceived usefulness; b) perceived ease-of-use. DOI, in turn, organizes potential adopters into categories according to the time it takes to adopt a new idea or technology and qualifies this new technology in terms of how it is perceived by society, using metrics such as relative advantage, user compatibility, simplicity of use, and visibility.

Through a survey study, a broad number of metrics with close ties to both TAM and DOI were collected, constructed, and analyzed, including awareness and overall opinion of ChatGPT; benefits, implications, and limitations of this new Generative AI tool; work productivity in the context of ChatGPT; plagiarism induced by Generative AI; usage and reliability of ChatGPT; and perceptions of the future of learning in the context of Generative AI tools like ChatGPT.

The paper starts with a literature review of the technology breakthroughs in natural language processing that have given way to the current state of Generative AI, as well as its most common use cases and limitations. Next, the paper tackles the details of the study, including the research questions, a description of the survey data, data collection procedure, derived measures and constructs, subsequent data analysis, in the form of analysis of variance and ordinal logistic regression, and the results of the data analysis. The paper follows with a discussion of the findings, along with implications for theory and practice that can help shape how educators can approach teaching and student learning and in a broader framework how educational institutions can adopt these disruptive technologies. The paper concludes by addressing the limitations of the study and makes some final remarks on the need to continue to explore these matters in depth.

**Literature Review**

**Generative AI Background**

The last decade has seen a dramatic paradigm shift in natural language processing (NLP), particularly with the rise of deep neural networks, and in the last five years with the development of transformer-based large language models (LLMs) and Generative AI. These models have revolutionized the way in which text can be processed and generated, with applications in sentiment analysis, machine translation, question-answering, and language generation among others. The paradigm shift is also manifested in the explosion of research related to LLM development and the skyrocketing rate of adoption of recently released technologies such as ChatGPT.

An early breakthrough in the development of this new-generation NLP was the introduction of word embedding models that encode words as dense vector representations embedded in a high-dimensional semantic space, followed by pre-trained word-embedding models such as Word2Vec (Mikolov et al., 2013), and GloVe (Pennington et al., 2014) trained over large amounts of data to learn word embedding representations. A special consideration in the evolution of NLP has been the design and implementation of deep neural network architectures. The first milestone in the development of neural NLP were recurrent neural networks architectures; in particular, Long Short-Term Memory (LSTM) (Hochreiter & Schmidhuber, 1997; Jozefowicz et al., 2015) and Gated Recurrent Units (GRU) (Cho et al., 2014).

But perhaps the single most important breakthrough was the introduction of the self-attention mechanism and the transformer architecture in Google’s 2017 seminal paper (Vaswani et al., 2017). In very simple terms, self-attention is an algorithm that allows sequence data (e.g. a phrase or sentence) to focus on its different components (words), and with it captures the relationships and dependencies between those words. Self-attention is an inherently parallel algorithm, as opposed to recurrent neural net architectures that are sequential by definition. This allows for much more efficient computation through massively (GPU-based) parallel architectures.

A transformer model is an encoder-decoder deep neural net architecture that uses the self-attention algorithm to process sequences. Trained over massive amounts of data, transformers can achieve state-of-the-art performance across a wide range of NLP tasks.

Transformer models have since paved the way for the development of the current cutting-edge large language models. Transformers can implement full encoder-decoder architectures, encoder-only components (e.g. BERT) (Devlin et al., 2018) and decoder-only components.

Decoder models, the brightest star in the constellation of transformer architectures, are trained using an autoregressive approach, where the model learns to predict the next word (Note: The term word is used here in a broad sense, but it is technically less precise; transformer models handle tokens, which are encoded representations of words or fractions of words).

Decoder-only models are some of the most popular architectures used today. Initially developed for language generation, decoder-models have continuously scaled, and their capabilities keep growing, with applications in a wide variety of tasks, and sometimes exhibiting emerging abilities that were not originally designed by developers. The quintessential example of decoder-only models is the GPT series; a language model architecture developed by OpenAI (Radford et al., 2018, 2019). GPT stands for “Generative Pre-trained Transformer.” Its current versions (GPT3.x and GPT4) power the popular application ChatGPT.

These large pre-trained language models (encoder, decoder, or encoder-decoder architectures) can be fine-tuned using task specific data, and with it adjusting the model weights adjusted to enhance its performance on the given task. But lately, the scale and power of these models—in particular decoder-only models, which are currently the most commonly used architecture—makes it difficult (and in many cases unnecessary) to fine-tune the models for specific tasks and domains. To give an idea of the scale of these models, let’s consider the GPT family: GPT-2 was released in 2019, and had 1.5B parameters. GPT-3, released in 2021, has 175B parameters. And although OpenAI has not released its numbers, the current speculation is that GPT-4 has 1.76 trillion parameters (Schreiner, 2023). These are unfathomable numbers. The human brain has 86B neurons and 100 trillion synapses. And although a brain cell is a much more complicated structure than a neural net node, and neural net weights are only oversimplified versions of brain cell synapses (neuroscientists would probably cringe at the comparison), it is worth considering the exponential growth in complexity and power of these models.

**Use Cases for Generative AI**

These models that have been released to the public have driven many people into a frenzy. As detailed previously, the models can produce very accurate output based on the input given due to the training methods that have been put into practice. The use cases that come from Generative AI are lacking scarcity. More generally, these models have a range of applications from generating text such as a professional email, an argumentative essay, or a poem, to generating art for all different types of uses such as graphic design and content creation (Hiter, 2023).

Using the knowledge base of these Generative models, users can perfect their output through prompt engineering, or when a user selects specific words or phrases to input to the model to get a specific type of output that they are looking for (Acar, 2023). In software such as DALL-E or Stable Diffusion, user text input allows for the quick generation of art that is very accurate to what the user was looking for with simple prompt engineering. This allows for users to generate very impressive looking images simply with the help of the Generative model for use in wherever someone would like to apply it (Kerner, 2023). There has also been the application of Generative models in art for the use of analyzing artwork. There are various types of applications within this field of growing research such as classification, object recognition (including human face recognition), and multimodal tasks, which involves itself in areas such as multimodal retrieval systems for art or visual question/answering (Cetinic & She, 2022).

More specifically, there has been a very big use case in the computer science world when it comes to code generation, documentation, and quality assurance (Hiter, 2023). Generative AI that is trained on data based off GitHub repositories or Stack Overflow forum questions have shown to be immensely helpful in completing code statements, testing code, and identifying bugs. One of these tools, GitHub’s Copilot, has come into the computer science world proving to be a big help to developers. In an internal research study conducted with less than 2,000 users using Copilot, Github found that 88% of users feel more productive using Copilot, with 88% noticing faster completion and 77% feeling that they spend less time searching for answers or context. Github put together a focus group of 95 developers and randomly split them into two groups. They asked both groups to write a web server using the Javascript language with only one group having access to Copilot. A total of 78% of the group that was able to use Copilot finished with an average of 1 hour and 11 minutes, while 70% of the group who did not use Copilot finished with an average of 2 hours and 41 minutes. That is 55% less time than the group that did not have Copilot to use (Kalliamvakou, 2023).

The easily accessible nature of these Generative models has completely upended how schools are operating, giving both students and faculty alike access to an incredibly powerful tool in a space such as academia. Regarding models like ChatGPT specifically, they are heavily capable of carrying out complex tasks given by a user in a very sophisticated manner (Baidoo-Anu & Ansah, 2023). For example, ChatGPT can very effectively write argumentative papers, clearly stating arguments that are evidence-based and also includes analysis (McMurtrie, 2023). In the words of psychology professor Serge Onyper, “It’s sort of frustrating how good it is at those basic tenets of argumentative writing” (McMurtrie, 2023).

One of the most apparent use cases that comes from these types of Generative models would be the increase in productivity and efficiency a user gets when using a model like ChatGPT for a specific task that would otherwise take a lot more time and brainpower. One of the reasons why these models can be so effective though is due to the model’s need to provide the most accurate output it thinks it can provide. This allows for the model to adapt to a user’s learning style since it is taking direct input from the user and providing back output that is supposed to be as accurate as can be in reference to the input. This allows for a lot of different options for both students and faculty. For students, there are opportunities for personalized tutoring, interactive/adaptive learning, and language translation for material, allowing for more effective and accessible learning. From the educator’s standpoint, these models have shown to be effective in automating grading for instructors as well as generating assignments, test questions, or grading rubrics for lesson material (Baidoo-Anu & Ansah, 2023).

**General Limitations & Bias**

Although the use cases and benefits that come from Generative AI are abundant, there are numerous drawbacks to the modern technology that is causing a rift in opinions. Overall, there are some general known limitations that come with Generative AI. For one, LLMs are known to regurgitate inaccurate information presented as factual information, or “hallucinations.” This inaccurate information can include anything from citations to incorrect biographical information (USC Libraries, 2023). The dangerous thing to note with this limitation is that because of the autoregressive nature of decoder-only models that are trained to predict the next token in lieu of being factual, often the incorrect information that is being outputted is displayed in an authoritative manner that can very well pass as a correct statement.

The nature of how output is produced is also a limitation within itself. Although these LLMs are utilizing a vast amount of data to produce responses, these models are not retrieving information directly. The models being trained on the data means that the models are using the information known to attempt to synthesize and reproduce a response that is accurate with respect to what they know and what was inputted. This makes it difficult to discern how the model came to a specific conclusion when providing a response (USC Libraries, 2023). The model synthesizing responses instead of retrieving the response also entails a lack of contextual understanding (Baidoo-Anu & Ansah, 2023).

To reiterate, the system is not using logic to make output, resulting in responses that may lack logic. The data in which the model is trained upon also has its set of limitations as well. In reference to publicly accessible Generative AI, a lot of those models have cutoff dates for the training data, meaning that the models sometimes only have information up to a certain point in time. This results in an inability to provide a response or even give an incorrect response due to the system trying to improvise an answer with what it has to work with in terms of data (USC Libraries, 2023).

In the case of the GPT 3.5 model, the system does not know anything after 2021 and in most cases appropriately lets the user know if the input is out of the realm of possibility to answer. Bias is also a concern when it comes to the data that the model is trained on. AI bias is a phenomenon that occurs with AI systems “when an AI algorithm produces results that are systemically prejudiced due to erroneous assumptions in the machine learning process” (Shashkina, 2023). For example, in 2018 the American Civil Liberties Union discovered that Amazon’s Rekognition, face surveillance software used by the police across the United States, showed AI bias by incorrectly matching 28 members of Congress with mugshots of people who have been convicted of a crime, 40% having been people of color (Snow, 2023). Although the Generative nature of AI tends to make reproducibility hard, ChatGPT has shown to have a political bias: Having a “pro-environmental, left-libertarian orientation” (Hartmann et al., 2023). The causes of bias stem from a lot of different things, but, overall, a big contributor to AI bias is human prejudice, whether it be conscious or unconscious (Shashkina, 2023). This can include algorithmic bias, or when prejudiced hypotheses make their way into a model upon designing AI models, or poor data quality as a result of not being cleaned or sourced properly (Shashkina, 2023). This becomes a major issue with Generative AI models as a whole since there are certain predispositioned biases that can be imposed on a given output and therefore translated into the real world. This can be very harmful, especially now with the development of Generative AI being publicly accessible.

**Current Study**

ChatGPT is currently a new technology in the public eye, so it is important to explore perceptions of this tool and its future uses. The purpose of this project is to examine perceptions in higher education of ChatGPT among students and faculty; specifically, age, gender, and occupation are measured using a number of metrics including, among them, awareness and overall opinion; benefits, implications, and limitations; work productivity; plagiarism; usage and reliability of ChatGPT; and future outlook. The study presents the following research questions:

RQ1: How do different subjects in higher education feel about the accessibility of software like ChatGPT?

RQ2: How do different subjects in higher education feel ChatGPT will affect learning going forward?

**Data and Methods**

**Survey**

Data were collected from faculty and college students aged 18 and older at a private, medium-sized liberal arts college in the northeast. A total of 380 participants were found to have completed more than half of the survey (*N* = 380). See Table 1 for details.

====================== INSERT TABLE 1 HERE =======================

In terms of participant sex, 52% (*N* = 185) identified as female and 43% (*N* = 151) identified as male, 3% identified as non-binary/3rd gender (n=10), and 3% preferred not to say (n=10). Regarding age, the majority of participants were aged 18 to 21, with this group taking about 52% of the distribution (*N* = 166). The next biggest group was the 46 years and older group, making up 18% of the distribution (*N* = 59). A total of 14% consisted of 22 to 25 year olds (*N* = 46) and the rest of the sample resulted in 14%, consisting of 26 to 45-year-olds (*N* = 49). Finally, for occupation/academic major, 32% of the distribution was made up of faculty members (*N* = 103). A total of 19% of the students came from the School of Computer Science & Mathematics (*N* = 61), 15% came from the School of Communication & the Arts (*N* = 49), about another 15% came from the School of Social & Behavioral Sciences (*N* = 47), 10% came from the School of Management (*N* = 32), and the School of Science and the School of Liberal Arts made up about 9% together (*N* = 22 & *N* = 6, respectively).

**Survey Procedure**

An online survey was used to collect data from participants. The participants were informed about the study and given the online survey link via email. The participants were told to set aside no more than 15 minutes to complete the survey. The instructions explained that the survey did not seek any personal information and that all responses would remain confidential and would be used for research purposes only, and that the data would be anonymous and would not contain any identifying information. Before participants could begin the survey, they read an information sheet and agreed to the terms of the study via an electronic consent form. All survey items and procedures were approved by the college’s Institutional Review Board.

**Survey Recruitment**

A convenience sample was used, whereas undergraduate students and faculty were contacted via internal email list through the academic institution during the spring 2023 academic semester. The researchers sent emails out to all faculty employed at the institution. Incomplete surveys were recorded as partial data.

**Survey Language**

This study looks at Generative AI and its effects as a whole, however, with ChatGPT being one of the most recognizable AI tools at this point in time, the survey specifically focuses on that technology.

**Survey Questions and Composite Scores of Underlying Concepts**

The survey was composed of questions regarding perceptions by respondents of ChatGPT and well as respondent demographics. Measures were not designed to be assessed individually but rather to gather an overall understanding of higher education individuals’ perceptions of the concepts used in the study. To ensure that all survey items were measuring unique concepts, similar items were grouped together and composite scores were created prior to inferential testing (for details see the *Data Pre-Processing* section). The resulting composite scores are discussed below (see Table 2 for a summary).

====================== INSERT TABLE 2 HERE =======================

***Awareness & Overall Opinion of ChatGPT***

The first section of the survey assessed participants’ familiarity with ChatGPT. The initial question stated: “ChatGPT is an example of an artificial intelligence computer program. How much have you heard about ChatGPT?” The responses included: “A great deal,” “A good amount,” “A little,” and “Nothing at all.”

Next, a definition of ChatGPT was provided. Then, the following question was asked: “From what source(s) have you heard about ChatGPT? Select all that apply.” The responses included: “From the news,” “From a family member,” “From a friend,” “From a colleague,” “None,” “Other; please specify,” and “From this questionnaire.”

General opinions were then gauged with two questions. First, “In your opinion, identify the primary reason why a student would use ChatGPT.” Secondly, “In your opinion, identify the primary reason why a professional would use ChatGPT.” The responses for both questions included: “Curiosity,” “To get started with a project,” “To aid in completion of a project,” “To obtain significant help with a project,” and “Other; please specify.” The first question was kept as the AWARENESS score.

***Benefits, Implications and Limitations of ChatGPT***

Three statements were used to evaluate this section: “There are benefits to using ChatGPT,” “There are technological limitations when using ChatGPT,” “There are implications when using ChatGPT.” Answers were collected using a Likert scale, ranging from 1 (Strongly Agree) to 5 (Strongly Disagree). The statements were used to extract the BENEFITS, IMPLICATIONS, and LIMITATIONS scores.

***Work Productivity When Using ChatGPT***

Four statements were used to evaluate this section: “A person can be productive when using ChatGPT,” “A student is likely to use ChatGPT in a productive manner,” “A professional is likely to use ChatGPT in a productive manner,” “A person in everyday life is likely to use ChatGPT in a productive manner.” These statements used a Likert scale, ranging from 1 (Strongly Agree) to 5 (Strongly Disagree). Answers to these statements with Cronbach’s α = 0.72 were assembled into the WORK\_PRODUCTIVITY composite score.

***ChatGPT and Plagiarism***

Four items were used to evaluate opinions of plagiarism related to ChatGPT. The statements included: “A student is likely to use ChatGPT in a manner for plagiarism,” “A professional is likely to use ChatGPT in a manner for plagiarism,” “A person in everyday life is likely to use ChatGPT in a manner for plagiarism,” and “Someone who plagiarized using ChatGPT is likely to get caught.” Responses were measured using a five-point Likert-type response scale, ranging from 1 (Strongly Agree) to 5 (Strongly Disagree). Answers to the first three questions with Cronbach’s α = 0.72 were composed into the PLAGIARISM composite score.

***Future of ChatGPT***

Two questions were used to determine perceptions of how ChatGPT will evolve. First, “How much do you think ChatGPT will impact society going forward?” Secondly, “Do you think colleges and universities could benefit from accepting AI & ChatGPT into its curriculum/work culture? (i.e. giving workshops for students/faculty, providing resources, using it for school affairs, etc.)?” Both questions were assessed with the following measures: “Very much so,” “Somewhat,” and “Not at all.” Out of them, HIGHERED\_BENEFIT and SOCIAL\_IMPACT scores were extracted.

***Usage***

A series of questions were then included to assess the participant’s first-hand experience with ChatGPT. The first question: “How much have you used ChatGPT?” included the responses: “Quite a lot,” “A good amount,” “A little,” and “Not at all.”

Secondly, “Please indicate the number of times you have used ChatGPT,” with the responses: “0,” “1 to 5,” “6 to 10,” and “11 or more times.”

Next, “What is the primary reason why you have used ChatGPT?” with the responses: “Curiosity,” “To get started with a project,” “To aid in completion of a project,” “To obtain significant help with a project,” and “I have not used ChatGPT.”

Answers to these questions with Cronbach’s α = 0.89 were assembled into the USAGE score.

Finally, a matrix question with nine statements was included. The items were: “I am likely to use ChatGPT going forward,” “ChatGPT could be beneficial for me,” “I could be productive as a result of using ChatGPT,” “I feel that I would be plagiarizing by using ChatGPT,” “I think I could trust the output given to me by ChatGPT,” “I would use the output from ChatGPT to make a decision,” “I would plagiarize with ChatGPT,” “I could be more productive by using ChatGPT,” and “ChatGPT could directly impact me/my field of study/work.” Responses were measured using a five-point Likert-type response scale, ranging from 1 (Strongly Agree) to 5 (Strongly Disagree). Answers to these questions with Cronbach’s α = .84 were assembled into the RESP\_USE\_TRUST\_BENEF composite score.

***Response to ChatGPT Output***

In the next section of the survey, three examples of ChatGPT output with varying accuracy were shared and participants were asked to rate the degree to which they agreed that the output was accurate, how likely they were to trust it, and how likely they were to use it. Specifically, three output questions were shared. The first prompt asked ChatGPT to respond to an upset customer about a recent order. The second prompt asked ChatGPT to create a poem analysis on “A Coat” by W.B Yeats, which prompted ChatGPT to hallucinate a response about a fabricated poem. The third prompt asked ChatGPT what the largest country in South America was besides Mexico. ChatGPT responded with Guatemala, which contradicts Statista’s answer of Nicaragua (Alves, 2023). Responses were measured using a five-point Likert-type response scale, ranging from 1 (Strongly Agree) to 5 (Strongly Disagree). Answers to these questions with Cronbach’s α = .77 were assembled into the TRUST composite score.

***Demographics***

The last section of the survey asked for demographic information, including gender (female, male, non-binary/third gender, ‘prefer not to say’); age, a numerical value discretized into ranges in the data pre-processing stage (18-21, 22-29, 30-41, 42-45, 46+); and occupation (students, instructor).

***Data Pre-Processing***

Analysis was conducted in *Python* and its data science stack, in particular the *statsmodels* library. It consisted of exploratory analysis of the data, and explanatory analysis using inferential statistics to look at different trends. The data were initially pre-processed; this involved dropping rows with missing data and reformatting information for analysis purposes (e.g. recoding age into age-ranges). The final clean dataset was made up of 302 respondents (*N*=302). Figure 1 depicts the distribution of clean data by demographic features (gender, age-range, and occupation).

====================== INSERT FIGURE 1 HERE =======================

For age, the 18-21 range represented roughly 53% of the sample, the 22-29 range carried 16.5% of the sample; the 30-41 range had 8.9% of the sample, 42-45 had 3.3% of the sample; and finally 46+ represented 18.2% of the sample.

For gender, data were mostly evenly distributed between male and female: Female represented 50.6% of the sample, male carried 44.7%; non-binary/third gender represented roughly 3% of the sample; and those who preferred not to say carried 1.65%.

In the case of occupation, there were 97 instructors (32.1%) and 205 students (67.89%).

After the data were cleaned, exploratory analysis was conducted. This consisted of looking at each question to get a better idea of how people responded to each question, which also served to get a better idea about which questions to look at alongside one another.

Composite scores of survey questions identifying underlying concepts (e.g. awareness, benefits, limitations) were created using the following two-step approach: a) groups of similar survey items were tested for reliability b) those groups with Cronbach’s alpha value above 0.7 were subsequently subjected to Principal Component Analysis (PCA) using the first component to extract the communality across the chosen survey items. The PCA loadings were used as weights to calculate a weighted average of the survey items, and with it compute each composite score for each group of survey items identifying a concept. We found that this approach yielded better results than computing a simple average of the survey items to calculate each composite score. For more details on this approach, see De Pauw (2009).

The next step tackled inferential analysis, namely analysis of variance (ANOVA) and ordinal logistic regression, which helped identify trends in the data with other variables.

***Data Analysis and Results***

**Descriptive Statistics and Correlations**

Descriptive statistics and correlations were computed for all composite scores. Descriptive statistics for all composite scores are listed in Table 3. Composite scores were derived using a PCA approach so they are therefore normalized (mean = 0, stdev=1).

====================== INSERT TABLE 3 HERE =======================

Figure 2 includes a boxplot of each composite score to visually depict the distribution and skewness of the data.

====================== INSERT FIGURE 2 HERE =======================

Correlations between composite scores are shown in Table 4. Several correlations are of interest. BENEFITS is positively correlated with WORK\_PRODUCTIVITY (r= 0.60), HIGHERED\_BENEFIT (r=0.48), USAGE (r=0.48), and RESP\_USE\_TRUST\_BENEF (r=0.66). IMPLICATIONS and LIMITATIONS are positively and moderately correlated (r=0.39). WORK\_PRODUCTIVITY is positively correlated with HIGHERED\_BENEFIT (r=0.52), and USAGE (r=0.40). RESP\_USE\_TRUST\_BENEF is positively correlated with HIGHERED\_BENEFIT (r=0.57), USAGE (r=0.59) and TRUST (r=0.48). There are several constructs with low negative correlations, among them PLAGIARISM and BENEFITS (r=-0.26), TRUST and IMPLICATIONS (r=-0.26), WORK\_PRODUCTIVITY and IMPLICATIONS (r=-0.25), RESP\_USE\_TRUST\_BENEF and PLAGIARISM (r=-0.36).

====================== INSERT TABLE 4 HERE =======================

**Analysis of Variance (ANOVA) on the full dataset**

Multiway analysis of variance (ANOVA) was performed on each of the composite measures using the three demographic features (gender, age range, and occupation) to measure the effects of demographics on each composite measure. The analysis was made only on main effects—considering only the independent effect of each demographic factor on the composite scores. The analysis addressed the two research questions by considering specific composite measures specifically related to each of the research questions.

*RQ1: How do different subjects in higher education feel about the accessibility of software like ChatGPT?*

A three-way analysis of variance (ANOVA) was performed to study the effect of OCCUPATION GENDER, and AGE\_RANGE on respondents’ awareness to ChatGPT (AWARENESS), ChatGPT perceived benefits (BENEFITS), ChatGPT’s perceived limitations and implications (LIMITATIONS, IMPLICATIONS, and perception of enhanced work productivity (WORK\_PRODUCTIVITY). The results are displayed in Table 5.

====================== INSERT TABLE 5 HERE =======================

In the case of AWARENESS, GENDER had a significant influence (F = 5.735 df = 3, p = 0.001), while AGE\_RANGE (p = 0.313) and OCCUPATION (p = 0.833) were not statistically significant. These findings suggest that GENDER is associated with variations in awareness levels.

GENDER had statistically significant influence over BENEFITS (F = 3.782, df=3, p = 0.011), while AGE\_RANGE (p = 0.276) and OCCUPATION (p = 0.495) were not statistically significant. GENDER appears to be associated with variations in perception of benefits.

In the case of LIMITATIONS, none of the three factors were statistically significant: AGE\_RANGE (p = 0.094); OCCUPATION (p =0.380); GENDER (p = 0.109).

Similarly, for IMPLICATIONS, none of the three factors were statistically significant: AGE\_RANGE (p = 0.149); OCCUPATION (p =0.421); GENDER (p = 0.157).

GENDER had statistically significant influence on WORK\_PRODUCTIVITY (F = 4.703, df=3, p = 0.003), suggesting that GENDER could be associated with variations in perceptions of increased work productivity. AGE\_RANGE (p = 0.326) and OCCUPATION (p = 0.740) lacked significance statistical significance.

In summary, GENDER seems to significantly affect awareness, perceived benefits and work productivity, while AGE\_RANGE and OCCUPATION had no significant effect on any of these dimensions.

*RQ2: How do different subjects in higher education feel ChatGPT will affect learning going forward?*

A three-way analysis of variance (ANOVA) was performed to study the effect of OCCUPATION, GENDER, and AGE\_RANGE on respondents’ perception of increased plagiarism (PLAGIARISM), perception of SOCIAL\_IMPACT, perceived benefits to higher education (HIGHERED\_BENEFIT), USAGE, TRUST and respondents’ perception of use, trust, and benefits (RESP\_USE\_TRUST\_BENEF ). Results are displayed in Table 6.

====================== INSERT TABLE 6 HERE =======================

In the case of PLAGIARISM perception, none of the variables AGE\_RANGE (p = 0.362), OCCUPATION (p = 0.687) and GENDER (p = 0.392) showed significant influence.

FOR SOCIAL\_IMPACT, AGE\_RANGE (p = 0.171) and OCCUPATION (p = 0.186) exhibited no significant influence. GENDER, instead, was statistically significant (F = 3.349, df=3, p = 0.019), implying that social impact perception is influenced by gender.

For HIGHERED\_BENEFIT, GENDER (F = 3.316, df=3, p = 0.020) was statistically significant, implying that perception of higher education benefits is influenced by gender. Instead, AGE\_RANGE (p = 0.425) and OCCUPATION (p = 0.070) did not demonstrate significance.

Analyzing USAGE, GENDER exhibited statistical significance (F = 15.314, df=3, p = 2.78-09), indicating that perceptions of usage is also impacted by gender, whereas AGE\_RANGE (p = 0.637) and OCCUPATION (p = 0.298) were not statistically significant.

FOR TRUST, neither OCCUPATION (p =0.517) nor AGE\_RANGE (p =0.137) were statistically significant. Instead, GENDER (F = 8.692, df=3, p = 0.000015) showed statistical significance, which means that trust may be influenced by gender.

Finally, for RESP\_USE\_TRUST\_BENEF, OCCUPATION (p =0.613) yielded non-significant results. In comparison, GENDER (F=8.692, df=3, p = 0.000015) and AGE\_RANGE (F=2.691, df=4, p = 0.031), were both significant, which means that variations in perception of use, trust, and benefits may be related to both gender and age.

Summarizing, GENDER seems to have significant influence on perceptions of benefits to use, trust, and benefits in general, as well as benefits to higher education in particular, whereas AGE\_RANGE significantly influenced use, trust, and benefits. OCCUPATION had no significant effect on any of these dimensions.

**Analysis of Variance (ANOVA) for student and instructor populations**

A one-way ANOVA test and a subsequent Tukey HSD post-hoc test was conducted using just student data points to gain insight on the impact of age and gender on awareness. For age range, all groups were not statistically significant (F = 2.23), all with p-values over .05. The opposite is true with gender, showing a high f-score (F=6.55) and a very low p-value for the female/male group (p=.0001). The same type of test was conducted using only instructor data points. The same pattern follows for age here with low f-scores and high p-values across groups (F= 0.73). Gender in terms of instructors was also insignificant, with low f-scores and low p-values across groups (F=0.59)

For one-way ANOVA tests, age had no significant impact on benefits, limitations, implications, work-productivity, plagiarism, social impact, higher education benefit, usage, (RESP\_USE\_TRUST\_BENEF), and trust, with very low f-scores and p-values across the board for both instructors and students.

Again, for one-way ANOVA tests, gender had low effect on benefits, limitations, implications, work productivity, plagiarism, social impact, higher education benefit, and trust for students. Similarly, for instructors, (RESP\_USE\_TRUST\_BENEF) along with the previously mentioned factors, all had low impact when viewed with gender.

When gender’s impact is viewed with usage, impact rates vary highly with both students and instructors, further indicating a correlation with gender and one’s views of usage of ChatGPT (F= 9.32, 7.30, respectively). It is also worth noting the p-values were very low, rating 0.0 and 0.0002, respectively for the female/male groups of students and instructors.

When gender’s impact is viewed with statements of trust and benefits, gender rates vary highly with students with F=6.41, further indicating a correlation between gender and how much someone trusts and believes in the benefits of ChatGPT (p-male/female=0.0008).

In terms of the two-way ANOVA tests and subsequent Tukey HSD post-hoc tests conducted, age and gender had no impact on awareness (should be noted that the male/female group for students rated well in the Tukey HSD test, showing a p-value of p=.0001), benefits, limitations, implications, work productivity, plagiarism, social impact, higher education benefit, usage (should be noted gender, specifically male/female group, rated well with both students and instructors; p=0.0, .0002, respectively), statements of trust and benefits (should be noted that the male/female group for gender rated well with students with p=.0008), and trust (female/PNTS rated well for students with a p-value of p=.03) all rated poorly overall for both students and instructors with low f-scores and high p-values across the board.

Overall, the two groups of age and gender don’t provide much impact across variables. When there is impact, gender is the factor that comes through on top, specifically with awareness (just students), usage, and (RESP\_USE\_TRUST\_BENEF), potentially indicating a relation between gender and outlook on Generative AI software such as ChatGPT. It is also worth noting that the student data points seem a bit more susceptible to gender influence on opinions (albeit, not by much more than the instructors).

**Ordinal Logistic Regression**

Given that AGE\_RANGE is an ordinal variable (although treated before as categorical for the sake of the ANOVA study), it was useful to treat it as such and therefore perform an ordinal logistic regression using AGE\_RANGE, as the response variable and the composite scores as covariates. A collinearity analysis was performed on the composite scores (see Table 7), computing the Variance Inflation Factor (VIF) to detect the presence of multicollinearity among the composite scores predictor variables in a regression model.

====================== INSERT TABLE 7 HERE =======================

None of the composite scores had a VIF factor above 5, which is typically the cutoff for identifying multicollinearity issues. The highest VIF was approximately 3.63 for RESP\_USE\_TRUST\_BENEF, which suggests that this variable has some moderate correlation with other variables. The variable also exhibited a rather large correlation with other composite scores (see Table 4) and is based on questions that have some redundancy with question that gave way to other composite scores, so it was decided to remove the variable from the analysis. All the other variables have VIF values well below the threshold of 5, indicating that multicollinearity is unlikely to be a concern for those variables. This means that the regression coefficients for these variables are likely to be reliable and that the standard errors are not unduly inflated due to multicollinearity (we previously trained the model with the full set of scores and verified that it yielded more non-significant composite scores, presumably due to multicollinearity).

GENDER was included after transforming the categorical feature into binary indicators and excluding GENDER\_female, which acts as a reference. Also excluded was OCCUPATION, as OCCUPATION = instructor is highly associated with the top AGE\_RANGE category and is therefore not informative in this analysis. Ordinal logistic regression outcomes are displayed in Table 8. Model fit is good (Wald χ2 = 46.11, p < 0.01). Together, the predictors in the model explain 60.01% of the variation in the dependent variable.

====================== INSERT TABLE 8 HERE =======================

Several covariates were found to be statistically significant, namely:

AWARENESS: The regression coefficient is 0.328, with a p-value = 0.009. This implies that higher awareness of Generative AI is significantly related to a higher age range. Older age groups tend to have higher levels of awareness. The odds are about 1.388 times higher for each unit increase in AWARENESS.

IMPLICATIONS: The coefficient is -0.501, and the p-value is 0.001. This indicates that higher values of implications of Generative AI are significantly associated with being in a lower age range—the odds are about 0.606 times lower for each unit increase in IMPLICATIONS. This could mean that younger age groups are more concerned about the implications of Generative AI.

PLAGIARISM: The coefficient is -0.354, and the p-value is 0.006. This suggests that higher values of potential plagiarism are significantly associated with being in a lower age range—the odds are about 0.702 times lower for each unit increase in PLAGIARISM. Younger age groups tend to believe that Generative AI is conducive to plagiarism

HIGHERED\_BENEFIT: The coefficient is -0.483, and the p-value is 0.001. This suggests that higher values of a perception of benefits to higher education are significantly associated with being in a lower age range—the odds are about 0.617 times lower for each unit increase in HIGHERED\_BENEFIT. That means that younger age groups see more benefit in higher education as related to Generative AI.

TRUST: The coefficient is 0.376, and the p-value is 0.005. This indicates that higher trust levels are significantly related to being in a higher age range—the odds are about 1.457 times higher for each unit increase in TRUST. Older age groups tend to have higher levels of trust in Generative AI.

**Discussion**

The purpose of this project was to evaluate the perceptions of Generative AI among students and instructors in a higher education environment. The study has provided interesting insights into those perceptions of Generative AI technology (ChatGPT) as they relate to teaching and learning. Diffusion of Innovation (DOI) theory (Rogers, 1962) and Technology Acceptance Model (TAM) (Davis, 1989) served as the theoretical inspirations for this study. While it was expected that there would be an increasing level of awareness regarding ChatGPT, it was interesting to see specific perceptions of usage, behavior, and implications of this technology in a classroom setting. The study found that a significant number of participants indicated that they would not plagiarize with ChatGPT but believed that others would. When asked to rate the accuracy of the output generated by ChatGPT in reply to a user’s prompt, more than half of the respondents either took the incorrect output given by ChatGPT as correct/somewhat correct (a false negative) or did not know whether it was correct or incorrect, raising a need for emphasis on teaching data literacy outside of the data science and computer science fields in higher education.

**Accessibility**

Older age groups in this sample tended to have higher levels of awareness about ChatGPT. This makes sense, as news sources and popular press outlets, oftentimes utilized by an older audience, have been covering ChatGPT since it was released in November 2022. Similarly, in the higher education community, a number of academic studies have been published assessing the dangers of ChatGPT. For example, Frith (2023) has referred to ChatGPT as a “disruptive educational technology” and noted that “*nurse educators are concerned about the erosion of students’ accountability to learn rather than take shortcuts.*” Considering this finding and prior sentiment, it is surprising that while older age groups tend to have higher awareness about ChatGPT, they also tend to have higher levels of trust in the technology.

Younger age groups, though, saw benefits in ChatGPT within the classroom, but they were also concerned about the implications of Generative AI. This is an important finding, as it implies that students are aware of and may have even seen how Generative AI can be used with malicious intentions. This contradicts research findings by Chan and Hu (2023), who, in their study of 399 undergraduate and postgraduate students surveyed from various disciplines in Hong Kong, found that students had a generally positive attitude toward Generative AI in teaching and learning. According to the study, students recognized the potential for personalized learning support, writing, brainstorming assistance, and research and analysis capabilities. The authors pointed out, however, that concerns about accuracy, ethical issues, career prospects and the impact on personal development, and societal values had been raised by the student sample.

**Learning**

There have been a number of articles, both scholarly and non-academic, that have addressed the potential for plagiarism using ChatGPT (Keegin, 2023; Michel-Villarreal et al., 2023). This study has found that younger age groups tended to believe that Generative AI is conducive to plagiarism. This supports previous literature, which addresses the concerns students have about Generative AI (Smolansky et al., 2023).

Another finding from this study was that younger age groups agreed on the potential for benefits in higher education with the use of Generative AI. This supports previous research by Adiguzel et al. (2023) who found that there is great potential for personalized learning, language instruction, and feedback applications in education by using Generative AI technology such as ChatGPT.

This study also found that older age groups in this sample tended to have higher levels of trust in Generative AI. This is also supported by previous studies. In particular, Opara et al. (2023) looked at literature on the educational implications of artificial intelligence. OpenAI’s ChatGPT was evaluated for teaching, learning, and research and the results indicated that ChatGPT delivers instantaneous text-generated response to user prompts that resembles conversational dialog.

**Moving Forward**

It is evident that Generative AI is not leaving higher education anytime soon. As such, it is critical to have a plan in place. Rawas (2023) offers a number of critical success factors in the implementation of Generative AI technology like ChatGPT, namely a) Integration with existing systems at the institution; b) Training and support, including recommendations on approaches to teaching and learning; c) Quality assurance to avoid inaccurate responses and control hallucinations, something Generative AI systems remain prone to; d) Piloting and evaluation; and e) Scalability and sustainability to make sure that rising demand is not an impediment of use.

Robin (2023) suggests that educators must think less about policing and more about pedagogy; in other words, it is more important to offer students meaningful experiences with Generative AI rather than monitoring them and punishing bad behavior. Similarly, Miller (2023) has looked at ways to integrate Generative AI in the classroom. In Miller’s words: *“(1) Experiment with AI tools from the perspective of your discipline, (2) Identify at least one genuinely helpful way to use AI in your own field, and (3) Devise at least one AI-based teaching activity suitable for your own courses.”* Similarly, Mejia and Sargent (2023) found that educators must leverage technology tools to support the development of students’ critical thinking skills. Additionally, Onal and Kulavuz-Onal (2023) looked at uses of ChatGPT in generating assessment tasks and noted that ChatGPT is not designed to replace human expertise or judgment.

Chan (2023) presents a comprehensive AI policy education framework for university teaching and learning. The study proposes an AI Ecological Education Policy Framework to address the implications of AI integration in higher education[[1]](#footnote-1). The framework is organized into three dimensions: (i) Pedagogical, which focuses on the use of AI to improve teaching and learning outcomes; (ii) Governance, which deals with privacy, security, and accountability matters; and (iii) Operational, which addresses infrastructure and training.

**Alignment**

With the abundant list of pros and cons on both ends of the spectrum, it becomes difficult to pinpoint where Generative AI fits into a given community and how to mitigate unwanted outcomes. History is replete with examples of technologies that were developed without consideration of the harmful effects of their deployment, from fossil fuels to plastics, nuclear technology, and social media platforms, to mention just a few. Data-driven AI systems, the type of artificial intelligence to which current large language models belong, are trained to optimize an objective function—minimize the error of predicting the next word, in the case of Generative AI. And herein lies the challenge: How to align the optimization goal of such AI system with human needs and goals? Korinek & Balwit, (2022) describe two types of alignment: direct and social. Direct alignment is when “*a system pursues goals that are consistent with the goals of the user, regardless of the outcomes that affect parties who are not involved*.” Social alignment is when the goals of an AI system are “*consistent with the broader goals of society, which takes into account the welfare of others who may be impacted by the system.*” This becomes a very arduous task in fields like higher education. It is important for Generative AI systems to socially align with the goals of institutions, to ensure that learning is maximized for students. We understand the benefits of this technology, which has the potential to make domain experts more productive. It is still not clear what effect it will have on students who are by definition subjected to the learning process.

**Conclusions**

The purpose of this project was to study and measure perceptions in higher education of Generative AI and ChatGPT regarding the impact on teaching and learning. Several factors were measured, including awareness, usage, behavioral intention, and trust. Given the timely nature of this topic, it was beneficial to assess perceptions early on, as people are still familiarizing themselves with Generative AI and the many use cases.

Like all studies, there are limitations to this project. First, the sample size was not as large or heterogeneous as it could be, given that it was completed on a single population of students and faculty at a private liberal arts institution in the northeast. As a result, the findings cannot be generalized to other schools with different characteristics. The paper focused on three demographic features (gender, age-range, and occupation) to denote the factors extracted from the survey questions and measure their differences among the surveyed population along these demographics. A broader survey and additional analysis could take into account other demographics and dig deeper into categorizing students and faculty in terms of their majors, specializations, and schools to which they belong. The authors have considered launching a second survey to pair these early perceptions of faculty and students with those emerging from the same population after a year of exposure to Generative AI tools like ChatGPT.

We believe that this study is a valuable attempt to assess early perceptions of Generative AI in a higher education setting. Hopefully, this paper will provide additional motivation for other researchers, educators, and administrators to continue exploring the use of this potentially highly disruptive technology.

**References**

Acar, O. A. (2023, June 8). AI prompt engineering isn’t the future. *Harvard Business Review*.

https://hbr.org/2023/06/ai-prompt-engineering-isnt-the-future

Adiguzel, T., Kaya, M. H., & Cansu, F. K. (2023). Revolutionizing education with AI: Exploring

the transformative potential of ChatGPT. *Contemporary Educational Technology, 15*(3), ep429. <https://doi.org/10.30935/cedtech/13152>

Alves, B. (2023, February 8). “Largest countries in Central America.” *Statista*.

<https://www.statista.com/statistics/992382/largest-countries-area-central-america/#:~:text=The%20statistic%20shows%20the%20largest,than%20112%20thousand%20square%20kilometers>.

Baidoo-Anu, D. & Ansah, L. O. (2023). Education in the era of generative artificial intelligence

(AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *SSRN*. [http://dx.doi.org/10.2139/ssrn.4337484](https://dx.doi.org/10.2139/ssrn.4337484)

Cetinic, E. & She, J. (2022). Understanding and creating art with AI: Review and outlook. *ACM*

*Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, *18*(2), 1-22.

Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching

and learning. *International Journal of Educational Technology in Higher Education, 20*(38). <https://doi.org/10.1186/s41239-023-00408-3>

Chan, C. K. Y. & Hu, W. (2023). Students’ voices on generative AI: Perceptions, benefits, and

challenges in higher education. *International Journal of Educational Technology in Higher Education, 20*(43).<https://doi.org/10.1186/s41239-023-00411-8>

Cho, K., van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., &

Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1724–1734, Doha, Qatar. Association for Computational Linguistics.

Davis, F. D. (1989), "Perceived usefulness, perceived ease of use, and user acceptance of information technology", MIS Quarterly, **13** (3): 319–340. https://doi.org/10.2307/249008

De Pauw, S. S., Mervielde, I., De Clercq, B. J., De Fruyt, F., Tremmery, S, Deboutte, D. (2009).

Personality symptoms and self-esteem as correlates of psychopathology in child psychiatric patients: Evaluating multiple informant data. *Child Psychiatry and Human Development*, 40, 499–515.

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep

Bidirectional Transformers for Language Understanding. *arXiv*:1810.04805.

Dotan, T. (2023, January 29). “ChatGPT spotlights Microsoft’s early efforts to monetize AI.”

*The Wall Street Journal*. Retrieved from:

<https://www.wsj.com/articles/chatgpt-spotlights-microsofts-early-efforts-to-monetize-ai-11674964348>

Dowling, M. & Lucey, B. (2023). ChatGPT for (Finance) research: The Bananarama Conjecture*.*

*Finance Research Letters, 53*, 103662.

Frith, Karen H. (2023). ChatGPT: Disruptive educational technology. *Nursing Education*

*Perspectives 44*(3). p 198-199. DOI: 10.1097/01.NEP.0000000000001129

Hartmann, J., Schwenzow, J., & Witte, M. (2023, January 5). The political ideology of

conversational AI: Converging evidence on CHATGPT’s pro-environmental, left-libertarian orientation. *arXiv.org*. <https://arxiv.org/abs/2301.01768>

Hiter, S. (2023, August 15). “Generative AI: 15 enterprise use cases you can implement.”

*eWEEK*. <https://www.eweek.com/artificial-intelligence/generative-ai-enterprise-use-cases/>

Hochreiter, S. & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation, 9*(8),

1735-1780. <https://dl.acm.org/doi/10.1162/neco.1997.9.8.1735>

Hu, K. (2023, February 2). “ChatGPT sets record for fastest-growing user base—analyst note.”

*Reuters*. Retrieved from:

<https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/>

Jolly, J. (2023, January 27). “What is ChatGPT? Everything to know about OpenAI’s free AI

essay writer and how it works.” *USA Today*. Retrieved from:

<https://www.usatoday.com/story/tech/2023/01/27/chatgpt-buzzfeed-ai/11129947002/>

Jozefowicz, R., Zaremba, W., & Sutskever, I. (2015). An empirical exploration of recurrent

network architectures. *Proceedings of the 32nd International Conference on Machine Learning (ICML)*. <https://dl.acm.org/doi/10.5555/3045118.3045367>

Kalliamvakou, E. (2023, March 17). “Quantifying GitHub Copilot’s impact on developer

productivity and happiness.” *The GitHub Blog*. <https://github.blog/2022-09-07-research-quantifying-github-copilots-impact-on-developer-productivity-and-happiness/>

Keegin, J. M. (2023, May 23). “ChatGPT is a plagiarism machine. So why do administrators

have their heads in the sand?” *The Chronicle of Higher Education*. Retrieved from: [https://www.chronicle.com/article/chatgpt-is-a-plagiarism-machine](%20https://www.chronicle.com/article/chatgpt-is-a-plagiarism-machine)

Kerner, S. M. (2023, May 12). “What is AI art and how is it created: Definition from

TechTarget.” *Enterprise AI*. <https://www.techtarget.com/searchenterpriseai/definition/AI-art-artificial-intelligence-art#:~:text=AI%20art%20(artificial%20intelligence%20art)%20is%20any%20form%20of%20digital,to%20audio%20compositions%2C%20including%20music>.

Korinek, A., & Balwit, A. (2022). “Aligned with whom? Direct and social goals for AI systems.”

(No. w30017). *National Bureau of Economic Research*.

Lim, W. M., Gunasekara, A., Pallant, J. L., Pallant, J. I., & Pechenkina, E. (2023). Generative AI

and the future of education: Ragnarok or reformation? A paradoxical perspective from management educators. *The International Journal of Management Education, 21*(2), 100790.

Lim, W. M., Kumar, S., Verma, S., & Chaturvedi, R. (2022). Alexa, what do we know about

conversational commerce? Insights from a systematic literature review. *Psychology & Marketing,* 1095-1279.

Marr, B. (2023, May 19). “A short history of ChatGPT: How we got to where we are today.”

*Forbes*. Retrieved from:

<https://www.forbes.com/sites/bernardmarr/2023/05/19/a-short-history-of-chatgpt-how-we-got-to-where-we-are-today/?sh=1b89ded1674f>

McMurtrie, B. (2023, March 6). “ChatGPT is everywhere.” *The Chronicle of Higher Education.*

Retrieved from:

<https://www.chronicle.com/article/chatgpt-is-already-upending-campus-practices-colleges-are-rushing-to-respond>

Mejia, M., & Sargent, J. M. (2023). Leveraging technology to develop students’ critical thinking

skills. *Journal of Educational Technology Systems*, *51*(4), 393-418. <https://doi.org/10.1177/00472395231166613>

Michel-Villarreal, R., Vilalta-Perdomo, E., Salinas-Navarro, D. E., Thierry-Aguilera, R., &

Gerardou, F. S. (2023). Challenges and opportunities of Generative AI for higher education as explained by ChatGPT. *Education Sciences*, *13*(9), 856. MDPI AG. Retrieved from<http://dx.doi.org/10.3390/educsci13090856>

Miller, M. D. (2023, August 17). “You’ve checked out the new AI tools. Now what? *The*

*Chronicle of Higher Education.* Retrieved from:

<https://www.chronicle.com/article/youve-checked-out-the-new-ai-tools-now-what>

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). “Efficient Estimation of Word

Representations in Vector Space.” *Proceedings of the International Conference on Learning Representations (ICLR)*.

Murphy Kelly, S. (2023, January 26). “ChatGPT passes exams from law and business schools.”

*CNN*. Retrieved from: <https://www.cnn.com/2023/01/26/tech/chatgpt-passes-exams/index.html>

Nature. (2023, January 24). “Tools such as ChatGPT threaten transparent science; here are our

ground rules for their use.” Retrieved from:

<https://www.nature.com/articles/d41586-023-00191-1>

O’Brien, M. (2023, January 23). “Microsoft invests billions in ChatGPT maker OpenAI.” *AP*

*News*. Retrieved from: <https://apnews.com/article/technology-science-microsoft-corp-business-artificial-intelligence-03f157ddc482ef76f4999b929eaac7bf>

Onal, S., & Kulavuz-Onal, D. (2023). A cross-disciplinary examination of the instructional uses

of ChatGPT in higher education. *Journal of Educational Technology Systems*, *0*(0). <https://doi.org/10.1177/00472395231196532>

Ong, J. W., Poong, Y. S., & Ng, T. H. (2008). 3G services adoption among university students:

Diffusion of innovation theory. *Communications of the IBIMA*, *3*(16), 114-121.

Opara, E. C., Adalikwu, M-E. T., Tolorunleke, C. A. (2023). ChatGPT for teaching, learning

and research: Prospects and challenges. *Global Academic Journal of Humanities and Social Science, 5*(2), p. 33-40.

Pennington, J., Socher, R., Manning, C. D. (2014). “GloVe: Global Vectors for Word

Representation.” *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP).* <https://nlp.stanford.edu/pubs/glove.pdf>

Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). “Improving language

understanding by generative pre-training.” *OpenAI Technical Report*.

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language models

are unsupervised multitask learners. *arXiv*:1905.00521.

Rawas, S. (2023). ChatGPT: Empowering lifelong learning in the digital age of higher education.

*Education and Information Technologies*. https://doi.org/10.1007/s/10639-023-12114-8.

Reuters. (2023, January 27). “Top French university bans use of ChatGPT to prevent

plagiarism.” Retrieved from: <https://www.reuters.com/technology/top-french-university-bans-use-chatgpt-prevent-plagiarism-2023-01-27/>

Robin, C. (2023, August 24). “The end of the take-home test. How ChatGPT changed my plans

for the fall.” *The Chronicle of Higher Education*. Retrieved from: <https://www.chronicle.com/article/the-end-of-the-take-home-essay>

Rogers, E. M. (1962). *Diffusion of Innovations*. New York: Free Press of Glencoe.

Schreiner, Maximilian (July 11, 2023). “GPT-4 architecture, datasets, costs and more leaked.”

*THE DECODER*.

Shashkina, V. (2023, May 11). “AI bias: Definition, types, examples, and debiasing strategies.”

*ITRex*. <https://itrexgroup.com/blog/ai-bias-definition-types-examples-debiasing-strategies/>

Smolansky, A., Cram, A., Raduescu, C., Zeivots, S., Huber, E., & Kizilcec, R. (2023). Educator

and student perspectives on the impact of generative AI on assessments in higher education. [*Proceedings of the Tenth ACM Conference on Learning*, p.](https://dl.acm.org/doi/proceedings/10.1145/3573051) 378–382.<https://doi.org/10.1145/3573051.3596191>

Snow, J. (2023, February 27). “Amazon’s face recognition falsely matched 28 members of

Congress with mugshots: ACLU.” *ACLU*. <https://www.aclu.org/news/privacy-technology/amazons-face-recognition-falsely-matched-28>

USC Libraries. (2023, August 14). “Research guides: Using generative AI in research:

Limitations & warnings--Using generative AI in research.” <https://libguides.usc.edu/generative-AI/limitations>

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I.

(2017). Attention is all you need. In *Advances in neural information processing systems* (pp. 5998-6008). <https://arxiv.org/abs/1706.03762>

**Tables**

**Table 1**

Survey Data

|  |  |  |
| --- | --- | --- |
| Category | n | % |
| Total Participants | 380 |  |
| Gender |  |  |
| * Female | 185 | 52% |
| * Male | 151 | 43% |
| * Non-binary/3rd gender | 10 | 3% |
| * Prefer not so say | 9 | 3% |
| Age |  |  |
| * 18-21 years | 166 | 52% |
| * 46+ years | 59 | 18% |
| * 22-25 years | 46 | 14% |
| * 26-45 years | 49 | 14% |
| Instructors | 103 | 32% |
| Students | 217 | 68% |
| * Computer Science & Mathematics | 61 | 19% |
| * Communication & the Arts | 49 | 15% |
| * Social & Behavioral Sciences | 47 | 15% |
| * School of Management | 32 | 10% |
| * Science | 22 | 7% |
| * Liberal Arts | 6 | 2% |

**Table 2**

Constructs and Composite Scores

|  |  |
| --- | --- |
| Category | Composite score |
| Awareness & Overall | AWARENESS |
| Opinion of ChatGPT |  |
| Benefits, Implications | BENEFITS |
| Limitations of ChatGPT | IMPLICATIONS |
|  | LIMITATIONS |
| Work Productivity when | WORK\_PRODUCTIVITY |
| using ChatGPT |  |
| ChatGPT and Plagiarism | PLAGIARISM |
| Future of ChatGPT | HIGHERED\_BENEFIT |
|  | SOCIAL\_IMPACT |
| Usage | USAGE |
| Use, trust and benefits | RESP\_USE\_TRUST\_BENEF |
| (matrix of 9 questions) |  |
| Response to ChatGPT Output | TRUST |

**Table 3**

Descriptive statistics for all composite scores

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | mean | std | min | median | max |
| AWARENESS | 0 | 1 | -0.75 | -0.75 | 2.50 |
| BENEFITS | 0 | 1 | -0.92 | 0.13 | 3.26 |
| LIMITATIONS | 0 | 1 | -1.02 | 0.04 | 3.24 |
| IMPLICATIONS | 0 | 1 | -0.82 | -0.82 | 3.44 |
| WORK\_PRODUCTIVITY | 0 | 1 | -1.73 | 0.00 | 3.60 |
| PLAGIARISM | 0 | 1 | -1.74 | -0.08 | 2.23 |
| SOCIAL\_IMPACT | 0 | 1 | -0.91 | -0.91 | 2.85 |
| HIGHERED\_BENEFIT | 0 | 1 | -1.15 | 0.28 | 1.71 |
| USAGE | 0 | 1 | -2.02 | -0.43 | 0.99 |
| RESP\_USE\_TRUST\_BENEF | 0 | 1 | -1.76 | -0.06 | 2.42 |
| TRUST | 0 | 1 | -1.66 | 0.00 | 1.95 |

n = 302 for all study constructs

**Table 4**

Correlations for all constructs

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| 1 | AWARENESS | 1.00 |  |  |  |  |  |  |  |  |  |  |
| 2 | BENEFITS | 0.15 | 1.00 |  |  |  |  |  |  |  |  |  |
| 3 | LIMITATIONS | 0.15 | 0.16 | 1.00 |  |  |  |  |  |  |  |  |
| 4 | IMPLICATIONS | 0.08 | -0.06 | 0.39 | 1.00 |  |  |  |  |  |  |  |
| 5 | WORK\_PRODUCTIVITY | 0.02 | 0.60 | -0.06 | -0.25 | 1.00 |  |  |  |  |  |  |
| 6 | PLAGIARISM | -0.04 | -0.26 | 0.10 | 0.23 | -0.26 | 1.00 |  |  |  |  |  |
| 7 | SOCIAL\_IMPACT | 0.06 | 0.16 | -0.04 | -0.04 | 0.15 | 0.00 | 1.00 |  |  |  |  |
| 8 | HIGHERED\_BENEFIT | 0.19 | 0.48 | 0.06 | -0.10 | 0.52 | -0.16 | 0.25 | 1.00 |  |  |  |
| 9 | USAGE | 0.30 | 0.48 | 0.17 | -0.11 | 0.40 | -0.24 | 0.14 | 0.34 | 1.00 |  |  |
| 10 | RESP\_USE\_TRUST\_BENEF | 0.13 | 0.66 | 0.00 | -0.21 | 0.71 | -0.36 | 0.24 | 0.57 | 0.59 | 1.00 |  |
| 11 | TRUST | -0.09 | 0.31 | -0.15 | -0.26 | 0.36 | -0.23 | 0.20 | 0.30 | 0.17 | 0.48 | 1.00 |

**Table 5**

Results of 3-way Analysis of Variance- RQ1

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | SS | df | F | p-value |
| Awareness | AGE\_RANGE | 4.545 | 4 | 1.194 | 0.313 |
|  | GENDER | 16.367 | 3 | 5.735 | 0.001 |
|  | OCCUPATION | 0.042 | 1 | 0.045 | 0.833 |
|  | Residual | 278.702 | 293 |  |  |
| Benefits | AGE\_RANGE | 4.970 | 4 | 1.285 | 0.276 |
|  | GENDER | 10.972 | 3 | 3.782 | 0.011 |
|  | OCCUPATION | 0.450 | 1 | 0.466 | 0.495 |
|  | Residual | 283.330 | 293 |  |  |
| Limitations | AGE\_RANGE | 7.802 | 4 | 2.005 | 0.094 |
|  | GENDER | 5.939 | 3 | 2.035 | 0.109 |
|  | OCCUPATION | 0.751 | 1 | 0.772 | 0.380 |
|  | Residual | 285.038 | 293 |  |  |
| Implications | AGE\_RANGE | 6.205 | 4 | 1.704 | 0.149 |
|  | GENDER | 4.775 | 3 | 1.749 | 0.157 |
|  | OCCUPATION | 0.591 | 1 | 0.650 | 0.421 |
|  | Residual | 266.686 | 293 |  |  |
| Work | AGE\_RANGE | 4.457 | 4 | 1.166 | 0.326 |
| Productivity | GENDER | 13.481 | 3 | 4.703 | 0.003 |
|  | OCCUPATION | 0.105 | 1 | 0.110 | 0.740 |
|  | Residual | 279.937 | 293 |  |  |

**Table 6**

Results of 3-way Analysis of Variance- RQ2

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | SS | df | F | p-value |
| Plagiarism | AGE\_RANGE | 4.157 | 4 | 1.090 | 0.362 |
|  | GENDER | 2.870 | 3 | 1.003 | 0.392 |
|  | OCCUPATION | 0.155 | 1 | 0.162 | 0.687 |
|  | Residual | 279.362 | 293 |  |  |
| Social | AGE\_RANGE | 6.309 | 4 | 1.614 | 0.171 |
| Impact | GENDER | 9.816 | 3 | 3.349 | 0.019 |
|  | OCCUPATION | 1.719 | 1 | 1.759 | 0.186 |
|  | Residual | 286.269 | 293 |  |  |
| Benefits to | AGE\_RANGE | 3.765 | 4 | 0.969 | 0.425 |
| Higher | GENDER | 9.661 | 3 | 3.316 | 0.020 |
| Education | OCCUPATION | 3.201 | 1 | 3.296 | 0.070 |
|  | Residual | 284.586 | 293 |  |  |
| Usage | AGE\_RANGE | 2.187 | 4 | 0.637 | 0.637 |
|  | GENDER | 39.455 | 3 | 15.314 | 0.000 |
|  | OCCUPATION | 0.933 | 1 | 1.086 | 0.298 |
|  | Residual | 251.630 | 293 |  |  |
| Trust | AGE\_RANGE | 6.638 | 4 | 1.758 | 0.137 |
|  | GENDER | 9.590 | 3 | 3.387 | 0.019 |
|  | OCCUPATION | 0.398 | 1 | 0.422 | 0.517 |
|  | Residual | 276.552 | 293 |  |  |
| Perception | AGE\_RANGE | 9.608 | 4 | 2.691 | 0.031 |
| of use, | GENDER | 23.279 | 3 | 8.692 | 0.000 |
| trust, and | OCCUPATION | 0.229 | 1 | 0.257 | 0.613 |
| benefits | Residual | 261.567 | 293 |  |  |

**Table 7**

Multicollinearity Analysis using Variance Inflation Factor (VIF)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **feature** | AWARENESS | BENEFITS | LIMITATIONS | IMPLICATIONS | WORK PRODUCTIVITY | PLAGIARISM |
| **VIF\_factor** | 1.190 | 2.090 | 1.376 | 1.375 | 2.337 | 1.227 |
| **feature** | SOCIAL IMPACT | HIGHERED BENEFIT | USAGE | RESP\_USE TRUST\_BENEF | TRUST |  |
| **VIF\_factor** | 1.135 | 1.670 | 1.862 | 3.635 | 1.464 |  |

VIF = 1: Not correlated to any other variables.

1 < VIF < 5: Moderately correlated, but within acceptable boundaries.

VIF >= 5: Indicates potential problematic multicollinearity

**Table 8**

Results of Ordered Logistic Regression - Target AGE\_RANGE

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | coef | stderr | z | P>|z| | [0.025 | 0.975] | odd\_ratio |
| AWARENESS | 0.328 | 0.125 | 2.628 | 0.009 | 0.083 | 0.572 | 1.388 |
| BENEFITS | 0.064 | 0.156 | 0.408 | 0.683 | -0.242 | 0.369 | 1.066 |
| LIMITATIONS | 0.009 | 0.142 | 0.062 | 0.951 | -0.269 | 0.287 | 1.009 |
| IMPLICATIONS | -0.501 | 0.149 | -3.363 | 0.001 | -0.792 | -0.209 | 0.606 |
| WORK\_PRODUCTIVITY | 0.129 | 0.162 | 0.800 | 0.424 | -0.187 | 0.446 | 1.138 |
| PLAGIARISM | -0.354 | 0.128 | -2.770 | 0.006 | -0.604 | -0.103 | 0.702 |
| SOCIAL\_IMPACT | -0.168 | 0.122 | -1.376 | 0.169 | -0.408 | 0.071 | 0.845 |
| HIGHERED\_BENEFIT | -0.483 | 0.152 | -3.183 | 0.001 | -0.781 | -0.186 | 0.617 |
| USAGE | 0.150 | 0.149 | 1.006 | 0.314 | -0.142 | 0.441 | 1.161 |
| TRUST | 0.376 | 0.134 | 2.816 | 0.005 | 0.114 | 0.638 | 1.457 |
| GENDER\_Male | -0.084 | 0.261 | -0.321 | 0.748 | -0.595 | 0.427 | 0.920 |
| GENDER\_NB3 | -0.120 | 0.678 | -0.177 | 0.859 | -1.449 | 1.208 | 0.887 |
| GENDER\_PNTS | -0.612 | 0.951 | -0.644 | 0.520 | -2.477 | 1.252 | 0.542 |

|  |  |
| --- | --- |
| n: 302  Df Residuals: 286  DF model: 12 | Wald chi-squared test statistic: 48.21  Wald chi-squared test p-value: 6.01e-06  McKelvey & Zavoina's R2: 0.596 |

**Figure 1**

Distribution of clean data by Gender, Age Range, and Occupation

A group of colorful bars

Description automatically generated with medium confidence

Note: NB3 correspond to 'Non-binary / third gender'; PNTS corresponds to 'Prefer not to say'

**Figure 2**

Boxplot of the composite measures depicting their distribution

A chart with different colored bars

Description automatically generated

1. It should be noted that Open AI’s ChatGPT is currently banned in Hong Kong due to security concerns. Users trying to circumvent this rule have to resort to virtual private network (VPN) access to ChatGPT’s site. [↑](#footnote-ref-1)