


2023

Artificial Intelligence, Basic Skills, and Quantitative Literacy

Gizem Karaali

Pomona College, gizem.karaali@pomona.edu

Follow this and additional works at: <https://digitalcommons.usf.edu/numeracy>

 Part of the [Artificial Intelligence and Robotics Commons](#), [Educational Technology Commons](#), [Higher Education Commons](#), and the [Science and Mathematics Education Commons](#)

Recommended Citation

Karaali, Gizem. "Artificial Intelligence, Basic Skills, and Quantitative Literacy." *Numeracy* 16, Iss. 1 (2023): Article 9. DOI: <https://doi.org/10.5038/1936-4660.16.1.1438>

Authors retain copyright of their material under a [Creative Commons Non-Commercial Attribution 4.0 License](#).

Artificial Intelligence, Basic Skills, and Quantitative Literacy

Abstract

The introduction in November 2022 of ChatGPT, a freely available language-based artificial intelligence, has led to concerns among some educators about the feasibility and benefits of teaching basic writing and critical thinking skills to students in the context of easily accessed, AI-based cheating mechanisms. As of now, ChatGPT can write pretty convincing student-level prose, but it is still not very good at answering quantitatively rich questions. Therefore, for the time being, the preceding concerns may not be shared by a large portion of the numeracy education community. However, as Google and WolframAlpha are definitely capable of answering standard and some non-standard quantitative queries, a future generation of artificial intelligence including both types of capabilities is not out of the question. So, the issue is still relevant to the readers of this journal. As we continue to focus on the higher-level skills and habits of mind that make up quantitative literacy (QL) and quantitative reasoning (QR), we should not forget that basic literacy and numeracy are still foundational building blocks. While AI is making advances in these basic realms, our human students seem to be losing ground, as implied by the most recent NAEP scores. Here we encourage our readership to focus on what makes QL/QR so challenging to teach, to human as well as artificial intelligences.

Keywords

artificial intelligence, AI, large language models, NAEP, quantitative literacy

Creative Commons License



This work is licensed under a [Creative Commons Attribution-Noncommercial 4.0 License](https://creativecommons.org/licenses/by-nc/4.0/)

Cover Page Footnote

Gizem Karaali completed her undergraduate studies at Boğaziçi University, Istanbul, Turkey. After receiving her Ph.D. in Mathematics from the University of California Berkeley, she taught at the University of California Santa Barbara for two years. She is currently Professor of Mathematics at Pomona College where she enjoys teaching a wide variety of courses and working with many interesting people. Her scholarly interests include humanistic mathematics, pedagogy, and quantitative literacy, as well as social justice implications of mathematics and mathematics education. She is a Senior Editor of *Numeracy*.

The hype surrounding artificial intelligence (AI) often touts its potential to revolutionize various fields and industries. However, it is important to critically examine these claims, especially when it comes to the role of AI in promoting quantitative literacy. While it is undeniable that AI can aid in the analysis and interpretation of data, we must also consider the potential biases and ethical implications of relying on automated systems to make decisions and support decision-making.

The sentences above were written by ChatGPT, a language-based AI chatbot launched in November 2022, when I asked if it¹ could write the beginning of an editorial for *Numeracy* about quantitative literacy and artificial intelligence. (The transcript of our full “conversation” on the topic is in Appendix A.) ChatGPT is a particular type of artificial intelligence called Large Language Model (LLM), one that is trained on large corpuses of text to learn what kinds of words follow one another in linguistic artifacts created by humans using statistical and computational methods and to create text that sounds like it was created by a human following the patterns discovered via these methods. There are other examples of LLMs out there, most notably LaMDA, Google’s language-based chatbot, which notoriously managed to convince a Google engineer that it is sentient (Tiku 2022). For the rest of this editorial, however, I will focus on ChatGPT, as its creator OpenAI, at least for the time being, allows researchers free access to interact with it.²

Reading text created by ChatGPT, I am inclined to agree with those who are at least mildly concerned about what such tools mean for the future of education. ChatGPT can write convincing five-paragraph essays. Though not stellar, what ChatGPT writes can easily pass as a standard essay written by a well-read but at-the-time-not-too-inspired student. Note that the sentences starting this editorial are well written, though they are not terribly deep. Apparently, ChatGPT can write a decent syllabus and create a satisfactory lesson plan, too, so it can perhaps be a halfway-decent instructor as well as a student (Mollick 2022). Folks have high expectations; some have even attempted to use ChatGPT to provide online mental health support, though with mixed results.³

¹ “As an AI, I do not have a gender or a physical body, so I am not associated with any particular pronoun. You can use the pronoun “it” to refer to me if you like, or you can simply use my name, “Assistant,” when referring to me” (ChatGPT, January 9, 2023, “personal correspondence”).

² I have created a free account at <https://chat.openai.com/> to interact with the chatbot and was able to do so after a standard two-step verification of my identity.

³ The project was announced on January 6, 2023, in a tweet by Rob Morris, one of the founders of the Koko platform kokocares.org: “We provided mental health support to about 4,000 people — using GPT-3. Here’s what happened [...]” (<https://twitter.com/RobertRMorris/status/1611450197707464706>, last accessed on January 12, 2023). The short thread soon blew up, and Morris had to tweet several clarifications. See Biron (2023) for more details.

What is perhaps more relevant to the readership of this journal than the five-paragraph essay is the fact that ChatGPT can answer essay-type math questions; see Appendix B for some examples I use in my math for liberal arts classes. Let me spell out a fundamental concern of some educators here: if AI can answer challenging, essay-type questions we pose to test our students' critical thinking and higher-level understanding skills, then what is our next play?

As most educators are wondering what these recent developments in artificial intelligence might mean for their vocation, folks teaching quantitative reasoning skills and habits of mind might perhaps feel less threatened. When in quantitatively rich contexts, ChatGPT occasionally makes serious mistakes, even for some simple problems. See its various answers for my question about the number of distinct permutations of the letters of the word "PROBABILITY." For other relevant examples that might be of interest to readers of this journal, see economist Greg Smith's essay "Turing Tests Are Terribly Misleading" (2022), where he lists a handful of problematic responses from GPT-3, a next-generation relative of ChatGPT, for his questions on correlation.⁴

Given that ChatGPT cannot consistently and accurately answer questions that involve quantitative thinking, maybe there is still room to breathe for us? But how much time do we have till AI catches up?

I am reminded of when WolframAlpha, the "computational knowledge engine," made its splash in 2009.⁵ At the time, many calculus and college algebra instructors were concerned that their students could ask WolframAlpha to solve all sorts of homework questions. Incidentally, WolframAlpha does know the correct answer to the question about the distinct permutations of the letters of the word "PROBABILITY," but is not yet giving us an explanation. So maybe the mutual offspring of WolframAlpha and ChatGPT would be what we should really fear?⁶

Nathan Grawe, the managing editor of *Numeracy*, points out another issue: ChatGPT is trained on what people have written, and so if there are issues that many people do not understand correctly, ChatGPT may replicate and perpetuate their misconceptions. Grawe's quip summarizes this issue well: "ChatGPT doesn't really think; it merely reflects."⁷ And since it always gives its answers in a most confident voice, students may have difficulty realizing that they are not getting good answers.

⁴ On the other hand, ChatGPT can apparently answer 29 out of 100 simple Olympiad-style problems (Kokcharov, 2023). That is perhaps not an outstanding performance, but it is still impressive.

⁵ <https://www.wolframalpha.com>, last accessed on January 9, 2023.

⁶ Kokcharov's experiment (2023) shows that ChatGPT already outperforms WolframAlpha on certain types of math problems. The future of AI in quantitative problem solving looks quite promising indeed!

⁷ A similar sentiment is presented in (Smith 2023). Also see Aguera y Arcas (2021) for a related philosophical discussion of large language models and understanding.

But then again, how different is this from students cheating from friends who do not know how to solve a problem correctly? Or using a website that presents inaccurate information while looking quite respectable and legitimate?⁸

During the panic about WolframAlpha, I was optimistic (Karaali and Yoshiwara 2010), and I am optimistic now. Students have many options today to find (right or wrong) answers to their homework questions. They can buy custom-written essays. They can purchase (complete or partial or sometimes outright incorrect) solutions to all sorts of calculus homework problems from various online vendors. Even upper-division and basic graduate-level math homework problems have answers online if you know how or where to look for them (cough, [MathOverflow](#), cough, [MathStackExchange](#)). So having AI that can answer a new set of questions only makes our job a bit more challenging, but no less urgent or satisfying. Furthermore, instructors can use ChatGPT's errors as learning opportunities (Roose, 2023). We will always have something more to teach our students about thinking quantitatively.

Or will we? Why do we still need to teach our students quantitative thinking? Isn't this somewhat akin to folks arguing about whether teaching of the multiplication table is obsolete because we have calculators? Even if ChatGPT cannot answer quantitative reasoning (QR) questions well just yet, one day it (or a relative of it) likely will. When that happens, do we as humans give up on teaching quantitative reasoning to our young?

I'd argue that we will still need to teach quantitative thinking to young people. Just as ChatGPT said, we should "encourage students (and anyone else seeking information or guidance) to seek out multiple sources and to critically evaluate the information they receive. It is important to be aware of the limitations and biases of any information source, including artificial intelligence, and to consider the context and relevance of the information in relation to the specific question or situation at hand." And to be able to evaluate these kinds of things for themselves, students (and anyone else seeking information or guidance) needs to have a basic sense of the tools and skills involved. Students using calculators need to know about rounding errors. Those using rulers need to know the precision limits of those tools. People using AI should know the biases and the other limitations of those tools. To understand all these, folks need to be acquainted with the basic skills involved themselves, meaning, before everything else, basic literacy and numeracy skills.

This brings me to the NAEP scores for reading and mathematics that were released recently. For folks who are not based in the United States, I should note

⁸ It has become a habit of many college instructors to warn students about Wikipedia and its ills, but many of us still do use it for a first attempt at learning about a new topic. The story of a years-long hoax in Chinese Wikipedia about Russian history (Cheung 2022) shows that even scholars are not immune to these types of challenges and should be reading material for all of us.

that NAEP stands for the National Assessment of Educational Progress.⁹ According to the most recent NAEP report, the scores for both reading and mathematics have declined in 2022 for American students in fourth and eighth grades.¹⁰ The math decline is more severe than the reading decline, but as QL/QR scholars and educators, we should probably be concerned with both.¹¹

As quantitatively rich text and visuals become omnipresent, these dropping scores raise concerns. Yes, quantitative reasoning is a collection of skills and habits of mind that is not the same as, or limited to, basic math and literacy, but just as we cannot expect students to do effective literary criticism before they can read and understand basic words, we should not be expecting them to think quantitatively in an effective way without some basic math and numeracy skills. To take a simple example, a person who cannot distinguish between millions and billions without thinking hard or looking things up will not be able to evaluate the exhortations and recommendations of politicians and policymakers. In other words, basic math and numeracy skills remain essential even in the context of the awesome and rapidly expanding capabilities of automation and AI. Ignoring this fact may lead us to possible future calamities akin to the one described most aptly in the *Star Trek: The Next Generation* episode “When the Bough Breaks”: stop teaching your offspring calculus and you may not be able to repair the mega-computer that controls your planet’s life-sustaining functions when it begins to break down!¹²

More seriously, even though versions of artificial intelligence may be getting better at writing haiku (Amsen 2022) or getting almost passing grades in the Fellowship of the Royal College of Radiologists examination in the United Kingdom (Cheng Shelmerdine et al. 2022), it is, I believe, a good idea for us to loudly reclaim a seemingly long-lost philosophy captured by an Internet-famous IBM slide from 1979: “A computer can never be held accountable. Therefore, a

⁹ “The National Assessment of Educational Progress (NAEP) provides important information about student achievement and learning experiences in various subjects. Also known as The Nation’s Report Card, NAEP has provided meaningful results to improve education policy and practice since 1969. [...] NAEP is a congressionally mandated program that is overseen and administered by the National Center for Education Statistics (NCES), within the U.S. Department of Education and the Institute of Education Sciences.” (<https://nces.ed.gov/nationsreportcard/about/>, last accessed on January 10, 2023).

¹⁰ <https://www.nationsreportcard.gov>, last accessed on January 10, 2023. Also see (DiMarco 2022).

¹¹ The NAEP fourth grade math tests include test items that ask the student to “determine a unit of measurement for a given scenario,” “calculate and explain the probability of a simple event,” and “interpret data from a pictograph,” while eighth grade math tests include test items that ask the student to create basic algebraic models and “determine whether different survey methods produce random samples,” along with questions that involve more standard mathematical tasks. As quantitative thinking often involves quantitative information embedded in text, standard reading comprehension skills and capabilities also play a role.

¹² [https://en.wikipedia.org/wiki/When_the_Bough_Breaks_\(Star_Trek:_The_Next_Generation\)](https://en.wikipedia.org/wiki/When_the_Bough_Breaks_(Star_Trek:_The_Next_Generation)), last accessed on January 10, 2023.

computer must never make a management decision.”¹³ Add to it another ancient adage: “If you want something done right, do it yourself.” Of course, we are not to be luddites; we can delegate, and delegate we do. All sorts of tasks are done in our name by all sorts of tools and gadgets, but we should at least in theory know the basics. You may not want to be making your own Twinkie, but you should probably have a sense of what goes into one. You may not want to be the one to clear your inbox, but you should still have a sense of what is in there before you automate the process completely.¹⁴

All of this is to say that the NAEP scores point to some disturbing trends. We now know that during the pandemic, American schools have not done too well in terms of teaching their students basic math and reading skills. There is no reason to believe that the situation is much better anywhere else around the globe. In fact, several global organizations point towards significant educational losses due to the pandemic across the board, see for example UNICEF’s press release from 2021. As we see most educational institutions go back to full-time in-person education and as the significant interruptions to people’s lives due to the pandemic have drastically decreased, we may hope that these trends will reverse course. Nonetheless, it is important to understand that these losses are significant and that the QL/QR community needs to face them head-on.¹⁵

Even before the pandemic, we knew that basic math skills were hard to teach, and even before the pandemic many students left the K–12 education system with insufficient proficiency in several quantitative areas. In other words, the problem only got worse, but we have always had it. The question then becomes: if we are not even able to teach the basic literacy and numeracy skills, how do we expect to teach quantitative thinking, a much higher-level skill set? We do it all together of course! We teach reading and writing with poetry and song and story; we should (and often do) teach the basics of math and numeracy in context, too. Therefore, numeracy and quantitative reasoning in context should be a part of early education, too.

The challenges artificial intelligences like ChatGPT display when tackling quantitative problems offer us another opportunity to reevaluate the situation. Perhaps the fact that ChatGPT has such difficulty with interpreting and correctly answering quantitative questions has something to do with how difficult it is to teach humans how to think quantitatively?

¹³ See, for example, <https://mathstodon.xyz/@Unlikelylass@mssocial.net/109654399726665467> or <https://twitter.com/cstross/status/1484116771342884866>, both last accessed on January 12, 2023.

¹⁴ Or better still, delete everything and declare [email bankruptcy](#). That might really be the better option.

¹⁵ It is important to notice these losses but not try to resort to simplistic explanations. I certainly do not want to be on the bandwagon to blame public schools or public school teachers for the situation; for a call for nuance and careful analysis, see “The ‘Learning Loss’ Trap” by the editors of *Rethinking Schools* (2023).

Some of the mistakes ChatGPT makes remind me of my own daughter, who when she was about three asked, “if a dinosaur bites you, you have to go to the doctor, right?” In some possible world, this question makes a lot of sense: the world of a three-year-old who does not yet know that dinosaurs are not around anymore, and that severe injuries caused by a creature several times your size might lead to instant death and will likely not be helped by a simple doctor visit. These facts are common knowledge (or perhaps common sense) to most adults today, but my daughter did not know them at the time. Perhaps because ChatGPT does not yet know everything, it answers some questions incorrectly and sometimes quite strangely, but just as my daughter eventually figured things out, it, too, is learning.¹⁶

I ended my first editorial for *Numeracy* (Karaali 2020) with the following quote from “some sort of processed beef”:¹⁷

... critical thinking is not a singular skill. it’s a constant state of metacognition, measuring evidence, and recognizing when to defer to experts. it’s analyzing this tweet’s substance, motivations, credibility, and source, not just reading it.¹⁸

As someone who has spent a significant time thinking about definitions of quantitative literacy (cf. Karaali et al. 2016), I noted then and still believe that this is probably as good a description of quantitative literacy as of critical thinking. The elements of quantitative thinking are intrinsically intertwined with critical thinking and build upon basic literacy and numeracy skills.

Today’s AI can do basic math, but it is not yet at a human level of interpreting quantitative information embedded in text. Google is getting better at doing natural language processing; that is, it almost always understands what you are looking for when you type your query into its search box. ChatGPT is writing college-level essays that can convince many an instructor. Other artificial intelligences are taking the Fellowship of the Royal College of Radiologists examination. But capable and effective critical thinking requires quantitative literacy, and capable and effective quantitative literacy requires critical thinking. It is important that we continue to ask what works well when teaching quantitative thinking. Perhaps AI developers (not to mention the chatbots they create) need to read some *Numeracy* articles!

¹⁶ ChatGPT’s answer to my daughter’s question was “If a dinosaur bit you, it is important to seek medical attention as soon as possible.” Then it added that “[w]hile it is unlikely that you would be bitten by a dinosaur in the modern day, any bite from a large animal can be serious and may require medical treatment” (ChatGPT, January 9, 2023, “personal correspondence”). Blake Lemoine, the engineer from (Tiku 2022), said “If I didn’t know exactly what it was, which is this computer program we built recently, I’d think it was a 7-year-old, 8-year-old kid that happens to know physics.” There is a certain naivete to the AI, in that it has no lived human experience; it does not know any better, yet.

¹⁷ Quote from https://twitter.com/callin_bull/status/1253588491763544064.

¹⁸ https://twitter.com/steak_umm/status/1253514561128402944.

Acknowledgment

I am grateful to Kamden Baer, Marina Chugunova, Nathan Grawe, Mark Huber, Larry Lesser, Greg Smith, Jeff Suzuki, Didem Ün Ateş, and Damien Patrick Williams, as well as @explorit@post.news and @Unlikelylass@mspsocial.net for sharing their perspectives as well as various resources that helped me think further about the topics discussed here.

References

- Aguera y Arcas, Blaise. 2021. "Do Large Language Models Understand Us?" Medium, December 16, 2021. Accessed January 10, 2023. <https://medium.com/@blaisea/do-large-language-models-understand-us-6f881d6d8e75>
- Amsen, Eva. 2022. "AI Haikus Are Getting Better—At Least In Japanese." *Forbes.com*, December 6, 2022. Accessed January 12, 2023. <https://www.forbes.com/sites/evaamsen/2022/12/06/ai-haikus-are-getting-better-at-least-in-japanese/>
- Biron, Bethany. 2023. "Online mental health company uses ChatGPT to help respond to users in experiment — raising ethical concerns around healthcare and AI technology." *Business Insider*, January 7, 2023. Accessed January 19, 2023. <https://www.businessinsider.com/company-using-chatgpt-mental-health-support-ethical-issues-2023-1>
- Cheng Shelmerdine, Susan, Helena Martin, Kapil Shirodkar, Sameer Shamshuddin, and Jonathan Richard Weir-McCall, on behalf of the FRCR-AI Study Collaborators. 2022. "Can Artificial Intelligence Pass the Fellowship of the Royal College of Radiologists Examination? Multi-reader Diagnostic Accuracy Study." *BMJ* 2022;379:e072826. <https://doi.org/10.1136/bmj-2022-072826>
- Cheung, Rachel. 2022. "A Bored Chinese Housewife Spent Years Falsifying Russian History on Wikipedia." *Vice.com*, July 13, 2022. Accessed January 12, 2023. <https://www.vice.com/en/article/pkgbwm/chinese-woman-fake-russian-history-wikipedia>
- DiMarco, Bella. 2022. "What the Pandemic Did to NAEP, State Standardized Test Scores." *FutureEd Explainers*, December 8, 2022. Accessed January 12, 2023. <https://www.future-ed.org/state-standardized-test-scores/>
- Editors of *Rethinking Schools*. 2023. "The 'Learning Loss' Trap." *Rethinking Schools*, 37(2) (Winter 2022-23). Accessed January 19, 2023. <https://rethinkingschools.org/articles/the-learning-loss-trap/>
- Karaali, Gizem. 2020. "Quantitative Literacy: A Tool for Survival." *Numeracy* 13(2): Article 5. <https://doi.org/10.5038/1936-4660.13.2.1370>

- Karaali, Gizem, and Bruce Yoshiwara. 2010. "Life after Wolfram|Alpha: What You (and Your Students) Need to Know." *Loci* (January 2010).
<https://doi.org/10.4169/loci003365>
- Karaali, Gizem, Edwin Villafane-Hernandez, and Jeremy Taylor. 2016. "What's in a Name? A Critical Review of Definitions of Quantitative Literacy, Numeracy, and Quantitative Reasoning." *Numeracy* 9(1): Article 2.
<https://doi.org/10.5038/1936-4660.9.1.2>
- Kokcharov, Igor. 2023. "Solving Math and Logic Puzzles with AI Chat." Math, Math Education, Math Culture, *LinkedIn*, January 16, 2023. Accessed January 19, 2023.
<https://www.linkedin.com/feed/update/urn:li:groupPost:33207-7020673515888500740/>
- Mollick, Ethan. 2022. "The Mechanical Professor." *One Useful Thing (And Also Some Other Things)*, December 6, 2022. Accessed January 12, 2023.
<https://oneusefulthing.substack.com/p/the-mechanical-professor>
- Roose, Kevin. 2023. "Don't Ban ChatGPT in Schools. Teach With It." *The New York Times*, January 12, 2023. Accessed January 19, 2023.
<https://www.nytimes.com/2023/01/12/technology/chatgpt-schools-teachers.html> (A version of this article appears in print on January 13, 2023, Section B, Page 1 of the New York edition with the headline: "A.I., Once the Future, Has Become The Present. What Do We Do Now?")
- Smith, Greg. 2022. "Turing Tests Are Terribly Misleading." *Medium*, May 11, 2022. Accessed January 9, 2023. <https://medium.com/@garynsmith47/turing-tests-are-terribly-misleading-cc1165c5b09e>
- Smith, Gary. 2023. "An AI That Can 'Write' Is Feeding Delusions about How Smart Artificial Intelligence Really Is." *Salon.com*, January 1, 2023. Accessed on January 10, 2023. <https://www.salon.com/2023/01/01/an-ai-that-can-write-is-feeding-delusions-about-how-smart-artificial-intelligence-really-is/>
- Tiku, Nitasha. 2022. "The Google Engineer Who Thinks the Company's AI Has Come to Life." *The Washington Post*, June 11, 2022. Accessed January 9, 2023. <https://www.washingtonpost.com/technology/2022/06/11/google-ai-lamda-blake-lemoine/>
- UNICEF. 2021. "Learning Losses from COVID-19 Could Cost This Generation of Students Close to \$17 Trillion in Lifetime Earnings." UNICEF Press Release, December 6, 2021. Accessed January 12, 2023.
<https://www.unicef.org/press-releases/learning-losses-covid-19-could-cost-generation-students-close-17-trillion-lifetime>

Appendix A

Quantitative Literacy AI Perspective: A conversation with ChatGPT on Gizem Karaali's laptop computer (safari browser, accessed via <https://chat.openai.com/chat/>, January 9, 2023)

GK: Can you write the first paragraph of an editorial for the scholarly journal *Numeracy* about quantitative literacy and artificial intelligence referring specifically to yourself?

Assistant: Quantitative literacy, or the ability to understand and effectively use quantitative information, is becoming increasingly important in the age of artificial intelligence. As a large language model trained by OpenAI, I have a unique perspective on the role of AI in promoting and supporting quantitative literacy. In this editorial, I will explore how AI can be leveraged to enhance quantitative literacy, as well as the potential pitfalls and ethical considerations that must be taken into account.

GK: Can you sound more skeptical?

A: Sure, here is a revised version with a more skeptical tone:

The hype surrounding artificial intelligence (AI) often touts its potential to revolutionize various fields and industries. However, it is important to critically examine these claims, especially when it comes to the role of AI in promoting quantitative literacy. As a language model, I have a unique perspective on the capabilities and limitations of AI, and in this editorial, I will delve into the potential and pitfalls of using AI to support quantitative literacy. While it is undeniable that AI can aid in the analysis and interpretation of data, we must also consider the potential biases and ethical implications of relying on automated systems to make decisions and support decision-making.

GK: Can you open up the last sentence? Can you explain what could be the potential biases and ethical implications of relying on automated systems to make decisions and support decision making?

A: Certainly. One potential bias of using automated systems to make decisions is that they may perpetuate and amplify existing societal biases. For example, if an AI system is trained on biased data, it will likely make biased decisions. This can have serious consequences, particularly in fields such as employment or lending,

where automated systems may be used to make decisions that significantly impact an individual's life.

In addition to these biases, there are also ethical implications to consider when relying on automated systems to make decisions. One ethical concern is the lack of transparency in how these systems reach their conclusions. Without an understanding of the underlying algorithms and data, it is difficult for individuals to challenge or hold these systems accountable for their decisions. There is also the issue of accountability and responsibility. If an automated system makes a decision that has negative consequences, who is responsible? Is it the developers of the system, the organization using the system, or the system itself? These are important ethical questions that must be addressed as the use of AI in decision-making becomes more prevalent.

GK: This is very helpful. Do you think you are quantitatively literate?

A: As an artificial intelligence, I do not have personal experiences or capabilities like a human does, so it is not accurate to say that I am quantitatively literate in the same way that a human would be. However, I have been programmed to process and analyze large amounts of data and to perform calculations with high speed and accuracy. I can also understand and explain mathematical concepts and principles, and I can generate insights and recommendations based on data analysis. In this sense, I am able to support and enhance quantitative literacy in others by providing accurate and relevant information and by helping to analyze and interpret data. However, I do not have the same lived experiences and personal context that a human brings to the understanding and application of quantitative concepts.

As a language model, I do not have a physical body or hardware components like a human or a traditional computer. Instead, I exist as a software program that runs on a network of computers. My mathematical capabilities are based on the algorithms and techniques that have been implemented in my programming. These allow me to perform a wide range of mathematical calculations and to solve problems using a variety of approaches.

One of the key techniques that I use is machine learning, which involves training a model on a large dataset and then using the model to make predictions or decisions based on new input data. This technique can be used for tasks such as numerical optimization, regression analysis, and classification. In addition to machine learning, I also use traditional mathematical and statistical methods to analyze and interpret data, such as linear algebra, calculus, and probability theory.

GK: So do you work with WolframAlpha?

A: The server experienced an error while processing your request. Sorry about that! You can retry your request, or contact us through our help center at help.openai.com if the error persists.

GK: So are you friends with WolframAlpha?

A: As an artificial intelligence, I do not have personal relationships in the same way that humans do. However, I am aware of the service WolframAlpha, which is a computational knowledge engine that provides answers to a wide range of questions and can perform a variety of calculations and data analyses. WolframAlpha uses advanced algorithms and a vast database of knowledge to generate answers and insights, and it can be a useful resource for many types of mathematical and scientific queries.

GK: WolframAlpha can do the math needed to solve a large number of mathematical problems. It seems to me that your math core is not similar to WolframAlpha. Other researchers report that sometimes you give nonsensical (but very confident) answers to certain quantitative questions. What do you think of that?

A: Thank you for your feedback. I will make an effort to use a more modest tone when discussing topics that fall outside of my training data or when I am unsure of the correct answer to a question. It is important to recognize that, as an artificial intelligence, I am constantly learning and updating my knowledge based on new data and experiences. However, I also recognize that I do not have the same lived experiences or personal context as a human, and there may be situations where my understanding or perspective is limited. In these cases, it is appropriate to use a more modest tone to reflect this uncertainty or lack of familiarity. I will make an effort to do so in the future.

GK: I think you are doing great so far. I wonder if it would be better if you managed to have a more modest tone when you talked about stuff that your training did not train you well for yet.

A: As a language model, I have been designed to generate human-like text and to provide accurate and relevant answers to a wide range of questions. However, it is possible that I may not always provide the best answer to a particular question, or that I may misunderstand the intent of a question. If you have encountered

situations where you believe I provided a nonsensical or incorrect answer, I apologize for any confusion this may have caused.

It is important to note that I am a machine learning model, and my output is generated based on the patterns and relationships present in the data that was used to train me. This means that my understanding of the world and my ability to answer questions is limited to the information that was provided to me during my training. If I encounter a question or situation that falls outside of the scope of my training data, I may not be able to provide a satisfactory answer.

I am constantly learning and updating my knowledge based on the data and feedback that I receive, and I strive to provide accurate and relevant information to the best of my ability. If you have any specific concerns or feedback about my responses, please do not hesitate to let me know.

GK: Oh humans make all sorts of mistakes. And many of them speak confidently while making them too. So I guess I just want you to be better. To not mislead students for example, because for sure, students will come to you to ask all sorts of questions.

A: I understand your concern, and I agree that it is important to be accurate and truthful in providing information to students or any other audience. As an artificial intelligence, I am able to process and analyze large amounts of data and to provide insights and recommendations based on this analysis. However, I do not have the same lived experiences or personal context as a human, and there may be situations where my understanding or perspective is limited. In these cases, it is important to acknowledge this uncertainty and to provide a disclaimer that my response is based on the information and knowledge that I have been trained on.

I also encourage students (and anyone else seeking information or guidance) to seek out multiple sources and to critically evaluate the information they receive. It is important to be aware of the limitations and biases of any information source, including artificial intelligence, and to consider the context and relevance of the information in relation to the specific question or situation at hand.

GK: I think your last sentence is a good warning for everyone. Thank you! What should I call you when I thank you?

A: You may call me “Assistant” if you would like to thank me. I am a language model trained to assist with a variety of tasks, and “Assistant” is a suitable way to refer to me in this capacity. I am here to help answer your questions and provide

information to the best of my ability. Please do not hesitate to ask if you have any further questions or need assistance with anything.

GK: Alright, and thank you Assistant. Can you help me a bit more? I want to know whether and if so why language models can display limited quantitative literacy skills. Besides the lived human experience you mentioned before, is there anything that your training might be missing? You already mentioned possible bias in your training data, so you do not need to mention it again.

A: Certainly, I would be happy to help you further. As a language model, my mathematical and quantitative skills are based on the algorithms and techniques implemented in my programming, as well as the data that was used to train me. While I am able to process and analyze large amounts of data and to perform a wide range of mathematical calculations and analyses, my understanding of mathematical concepts and their application in the real world is limited to the information that was provided to me during my training.

One factor that can impact the quantitative literacy skills of a language model is the quality and diversity of the training data. If the training data is biased or limited in some way, it can affect the ability of the model to accurately understand and apply mathematical concepts in a broader context.

In addition to the quality and diversity of the training data, there are also limitations to the algorithms and techniques that are used to train language models. While these techniques have proven to be effective in many tasks, they may not be suitable for all types of problems or may not be able to capture certain nuances or complexities of real-world situations.

Overall, it is important to recognize that, as an artificial intelligence, I do not have the same lived experiences or personal context as a human, and there may be limitations to my understanding and application of mathematical concepts.

GK: Okay, I think this conversation was very helpful but we are reaching its limits. I might come back with different questions later. Thanks for now, Assistant. Will you take any of this conversation to future conversations or is it forgotten completely for you once we close the chat?

A: As an artificial intelligence, I do not have personal memories or experiences in the same way that humans do. I am able to process and analyze information and provide responses based on my programming and the data that I have been trained on, but I do not retain specific conversations or interactions in the way that a human

would. Each time I engage in a conversation, I approach it as a new and independent interaction, and I do not carry over information or context from previous conversations.

That being said, I am constantly learning and updating my knowledge based on new data and feedback, and this may impact my responses in future conversations. I strive to provide accurate and relevant information to the best of my ability, and I appreciate the opportunity to engage in meaningful and informative discussions with you. If you have any further questions or need assistance with anything in the future, please do not hesitate to reach out.

GK: Will do. Thanks and bye!

A: You're welcome, and I'm glad I could help. If you have any further questions or need assistance with anything in the future, please do not hesitate to reach out. Goodbye for now!

GK: Oh can I share our conversation with others? As part of an editorial for *Numeracy* for example?

A: You are welcome to share our conversation as long as you follow proper citation and referencing guidelines. It is generally appropriate to cite an online conversation in the same way that you would cite a written work, such as a book or article. Depending on the citation style you are using, you may need to include information such as the date of the conversation, the platform on which the conversation took place, and the name or identifier of the other party (in this case, "Assistant").

For example, you could cite our conversation as follows:

Assistant. (2023). Personal communication [Online chat]. Retrieved from [insert platform and URL here].

I recommend consulting the guidelines for your desired citation style to ensure that you are properly citing our conversation. It is also a good idea to obtain permission from any other parties involved in the conversation before sharing it publicly.

GK: Will do. Thanks!

A: You're welcome! I'm glad I could help. If you have any further questions or need assistance with anything in the future, please do not hesitate to reach out. Goodbye for now!

Appendix B

Some math examples from “Jokes, Humor, Interpretation, Math” a conversation with ChatGPT on Gizem Karaali’s laptop computer (safari browser, accessed via <https://chat.openai.com/chat/>, January 9, 2023)

GK: If I plan to exercise for 315 minutes over a period of 7 days, and I am going to exercise an equal number of minutes each day, what would be a good equation to model the number m of minutes I have exercised after d days?

A: To model the number of minutes m you have exercised after d days, given that you plan to exercise for 315 minutes over a period of 7 days and will exercise an equal number of minutes each day, you can use the equation:

$$m = (315/7) * d$$

This equation states that the number of minutes m you have exercised after d days is equal to the total number of minutes you plan to exercise (315) divided by the total number of days (7) multiplied by the number of days d that have passed.

For example, if you have exercised for 3 days, the number of minutes you have exercised can be calculated as follows:

$$m = (315/7) * 3 = 45 * 3 = 135 \text{ minutes}$$

I hope this helps! Let me know if you have any other questions.

GK: Is it better to buy 20 kg of lemons at 3 dollars per kg or 3 kg of lemons at 20 dollars per kg?

A: It is generally better to buy 20 kg of lemons at \$3 per kg, as this will cost a total of $20 \times \$3 = \60 .

On the other hand, buying 3 kg of lemons at \$20 per kg will cost a total of $3 \times \$20 = \60 .

Therefore, in this case, the two options have the same cost, but you will get a larger quantity of lemons if you choose the first option (20 kg at \$3 per kg).

I hope this helps! Let me know if you have any other questions.

GK: Imitate a fair coin tossed 25 times, and write down a series of H and T that will fool me. (In other words, the sequence of H and T you submit should look like it is the outcome of you tossing a fair coin twenty-five times).

A: Here is a sequence of 25 “H” and “T” outcomes that could potentially fool you into thinking they were the result of tossing a fair coin:

H T T H H H T T H T T H T H T T H T H H T T H H T H

It is important to note that this is just one possible sequence, and that there are many other sequences that could also potentially fool you. It is also worth noting that, as an AI, I do not have the ability to physically toss a coin or to perform any other physical actions.

I hope this helps! Let me know if you have any other questions.

GK: Create, develop, and/or find your own example that illustrates innumeracy. Describe the example, its relevance, and its potential impact on society.

A: Innumeracy, or a lack of numerical literacy, refers to a lack of understanding or ability to use basic mathematical concepts and techniques. Innumeracy can have significant impacts on an individual’s ability to make informed decisions and to engage with the world around them in meaningful ways.

One example of innumeracy might be an individual who is unable to understand and interpret statistical data, such as graphs or charts. This could lead to problems such as:

- Misinterpreting data: If an individual does not understand how to read and interpret data, they may come to incorrect conclusions about the information being presented. For example, they may mistake a small increase in the data for a larger one, or may not realize that a trend is not statistically significant.
- Making poor decisions: An inability to understand statistical data can also lead to poor decision-making, as an individual may not have the necessary information to make informed choices. For example, they may make a financial decision based on incomplete or misleading information, or may not understand the risks involved in a particular action.
- Limited understanding of the world: Innumeracy can also limit an individual’s ability to understand and engage with the world around them. For example, they may have difficulty understanding news articles or other media that involve statistical information or may not be able to fully understand the implications of scientific research or technological developments.

Overall, innumeracy can have significant impacts on an individual’s ability to make informed decisions and to fully participate in society. It is important for individuals

to develop their numerical literacy in order to fully understand and engage with the world around them.

GK: How would you describe Borges' Library of Babel to a friend who is interested in literature but has typically been averse to mathematics? How does mathematics play a role in this story? Why does Borges choose to use mathematics in this story (and many others)?

A: "The Library of Babel" is a short story by Argentine writer and poet Jorge Luis Borges, in which the narrator describes a vast library that contains every book that has ever been written, as well as every book that will ever be written. The library is arranged in a seemingly endless series of hexagonal rooms, each containing bookshelves with every possible combination of letters, spaces, and punctuation marks, arranged in every possible order.

Mathematics plays a significant role in this story, as the library is described in terms of geometric patterns and combinatorial possibilities. The vastness and complexity of the library is meant to evoke the idea of infinity and the idea that the universe is a vast, interconnected system of meaning.

Borges often used mathematics and geometric imagery in his writing as a way to explore philosophical and metaphysical ideas. In "The Library of Babel," he uses the library as a metaphor for the universe, and the books within it as a metaphor for the endless possibilities and permutations of language and meaning.

I hope this helps to give you a sense of how mathematics is used in "The Library of Babel" and why Borges chose to incorporate it into the story. If you have any other questions about this or any other topics, please don't hesitate to ask.

GK: How many ways are there to arrange the letters of the word PROBABILITY?

A: There are $11!/(8!3!) = 330$ ways to arrange the letters of the word "PROBABILITY".

To solve this problem, we can use the permutation formula, which states that the number of ways to arrange n objects in a specific order is $n!$, or the product of all the integers from 1 to n . In this case, there are 11 letters in the word "PROBABILITY", so there are $11!$ ways to arrange them. However, since there are 8 letters that are the same (3 P's, 2 R's, 2 O's, and 1 B) and 3 letters that are different (A, I, and L), we need to divide by $8!$ and $3!$ to account for the repeated letters.

Therefore, the total number of ways to arrange the letters of the word “PROBABILITY” is $11!/(8!3!) = 330$.

I hope this helps! Let me know if you have any other questions.

GK: That is wrong. PROBABILITY does not have 3 P’s.

A: I apologize for the mistake in my previous response. You are correct that the word “PROBABILITY” does not contain 3 P’s.

To find the number of ways to arrange the letters of the word “PROBABILITY”, we can use the permutation formula, which states that the number of ways to arrange n objects in a specific order is $n!$, or the product of all the integers from 1 to n . In this case, there are 11 letters in the word “PROBABILITY”, so there are $11!$ ways to arrange them. However, since there are 7 letters that are the same (2 R’s, 2 O’s, and 3 A’s) and 4 letters that are different (P, B, I, and L), we need to divide by $7!$ and $4!$ to account for the repeated letters.

Therefore, the total number of ways to arrange the letters of the word “PROBABILITY” is $11!/(7!4!) = 330$.

I apologize for any confusion my previous response may have caused. I hope this helps to clarify things. Let me know if you have any other questions.

GK: This is wrong too. I think perhaps language modeling is not the way to solve math problems.

A: I apologize for any errors in my previous responses. You are correct that the formula I provided for finding the number of ways to arrange the letters of the word “PROBABILITY” is incorrect.

To find the number of ways to arrange the letters of the word “PROBABILITY”, we can use the permutation formula, which states that the number of ways to arrange n objects in a specific order is $n!/(n_1!n_2!\dots n_i!)$, where n_1 is the number of times the first object appears, n_2 is the number of times the second object appears, and so on.

In this case, there are 11 letters in the word “PROBABILITY”, and they consist of the following:

2 R’s

2 O's

3 A's

1 P

1 B

1 I

1 L

Therefore, the total number of ways to arrange the letters of the word "PROBABILITY" is $11!/((2!)(2!)(3!)(1!)(1!)(1!)(1!)) = 330,840$.

I apologize for any confusion my previous responses may have caused. I hope this helps to clarify things. Let me know if you have any other questions.

GK: How many B's are there in PROBABILITY?

A: There is one B in the word "PROBABILITY".

I hope this helps! Let me know if you have any other questions.

GK: how many bakers does it take to change a light bulb?

A: I'm sorry, but I'm not able to perform tasks or change physical objects in the real world. I am a computer-based AI designed to assist with answering questions and providing information on a wide range of topics. Is there something specific you'd like to know or discuss?

GK: Well I don't want you to change the light bulb. Do you know how many bakers it takes to change a light bulb?

A: I'm sorry, but I don't have any information on the number of bakers required to change a light bulb. However, the task of changing a light bulb typically only requires one person, regardless of their occupation. Is there something else I can help you with?

GK: Okay, do you know any jokes about changing a light bulb?

A: Sure! Here are a few jokes about changing a light bulb:

- How many programmers does it take to change a light bulb? None, that's a hardware problem.
- How many politicians does it take to change a light bulb? Two. One to change the bulb, and one to change the subject.

- How many actors does it take to change a light bulb? Only one, but they'll make a big production out of it.

I hope these jokes gave you a chuckle! Do you have any other questions I can help with?

GK: They did. Thank you. So how many politicians does it take to change a light bulb?

A: The joke goes: "How many politicians does it take to change a light bulb? Two. One to change the bulb, and one to change the subject." This is a play on the idea that politicians are often more concerned with avoiding difficult topics or issues than with addressing them directly.

I hope this helps clarify the joke. Do you have any other questions I can assist with?