Taylor & Francis
Taylor & Francis Group

Ethics Education for Healthcare Professionals in the Era of chatGPT and Other Large Language Models: Do We Still Need It?

Vasiliki Rahimzadeh, Kristin Kostick-Quenet, Jennifer Blumenthal Barby & Amy L. McGuire

⊕ View supplementary material ↗

📅 Published online: 24 Jul 2023.

✎ Submit your article to this journal ↗

📊 Article views: 757

🔍 View related articles ↗

View Crossmark data ↗

TARGET ARTICLE

Ethics Education for Healthcare Professionals in the Era of chatGPT and Other Large Language Models: Do We Still Need It?

Vasiliki Rahimzadeh ⓘ, Kristin Kostick-Quenet ⓘ, Jennifer Blumenthal Barby ⓘ, and Amy L. McGuire ⓘ

Baylor College of Medicine

ABSTRACT

In this paper, we contend with whether we still need traditional ethics education as part of healthcare professional training given the abilities of chatGPT (generative pre-trained transformer) and other large language models (LLM). We reflect on common programmatic goals to assess the current strengths and limitations of LLMs in helping to build ethics competencies among future clinicians. Through an actual case analysis, we highlight areas in which chatGPT and other LLMs are conducive to common bioethics education goals. We also comment on where such technologies remain an imperfect substitute for human-led ethics teaching and learning. Finally, we conclude that the relative strengths of chatGPT warrant its consideration as a teaching and learning tool in ethics education in ways that account for current limitations and build in flexibility as the technology evolves.

WHY ChatGPT AND ETHICS EDUCATION

ChatGPT has taken the academic community by storm (Cotton, Cotton, and Shipway 2023; Cox and Tzoc 2023; Sullivan, Kelly, and McLaughlan 2023). Since its release in November 2022, chatGPT has predictably followed Gartner's hype cycle albeit on an unpredictably accelerated timeline (Vk 2023). This is in part because chatGPT stokes the imagination about opportunities to rapidly synthesize text-based information. ChatGPAT also simultaneously raises alarm about what its humanlike responses could spell for the fidelity of academic work moving forward (Liebrenz et al. 2023). Generative artificial intelligence (AI)—which include large language models (LLM) like chatGPT (What is a Large Language Model (LLM)? 2022)—leverage natural language processing techniques to generate responses to human text queries by extracting patterns from human language.

It is debatable if this special AJOB issue means bioethics is currently locked in the "trough of disillusionment" on this hype cycle, whether it is inching toward the "slope of enlightenment" or at some intersection between them. One thing is, however, clear; chatGPT demands scrutiny from all manner of interdisciplinary expertise if the bioethics community is to engage in productive and inclusive debate about its appropriate uses for teaching and learning. Such engagement will also be important for informing responsible and responsive policy.

Our collective thinking about chatGPT's uses and policy responses will in no small part influence the way learners understand the past, present, and future of bioethics (Krügel, Ostermaier, and Uhl 2023). Toward this end, the locus of our analysis in this paper is on incorporating LLMs like chatGPT into ethics teaching and learning. As bioethics scholars responsible for bioethics education among future healthcare professionals, we are compelled by the current capabilities that chatGPT affords for helping learners build core knowledge of ethical principles, theories, and concepts. We are also committed to approaches that help instill professionalism among healthcare professional trainees (Asch 2023).

In the sections that follow, we first briefly describe LLMs, including chatGPT. We distill some of the common goals of ethics education for healthcare professionals from our own institution, with a focus on medical students. We then consider how chatGPT and similar LLMs can help us achieve programmatic goals

in ethics, and whether they deserve a reputable place in our pedagogical toolbox.

LLMs are likely to become more reliable, accessible, and usable over time. As such, we argue that any assessment of their current capabilities and limitations for medical ethics education must be treated as a cross section in time. Developing curriculum and institutional policies that make room for LLMs in ethics teaching and learning should likewise be adaptable. We explore the current strengths, weaknesses, opportunities, and threats of incorporating LLMs into ethics education for medical students using GPT-4 (OpenAI 2023) (Figure 1). Through an actual demonstration, we highlight the ways chatGPT can engage learners for how to ethically navigate controversial cases they may encounter in clinical practice.

WHAT IS A LARGE LANGUAGE MODEL (LLM)?

LLMs such as chatGPT apply deep learning to unsupervised searches that enable them to make predictions about how to perform complex linguistic tasks such as summation, translation, automatic dialogue generation, argument construction and evaluation, and in some cases even creative writing (see Table 1 for technical definitions).

LLMs were computationally infeasible to train until as recently as 2018 (Kharya and Alvi 2021) because they required data storage and memory that exceeded even the largest graphics processing unit available at the time. Key commercial developers, including Microsoft, NVIDIA, Google, and others, have largely been responsible for providing the necessary compute at scale to allow LLMs to generate responses that are now nearly indistinguishable from human responses.

As a result, LLMs have tremendous appeal for quickly summarizing the corpus of knowledge available on the searchable Internet about a given topic, bioethics included. Generative AI in general, and LLMs in particular, are useful for some of the following writing and reading tasks:

- Clarifying complex concepts
- Providing real-time feedback
- Pointing to additional resources
- Organizing text and synthesizing key points from bodies of text
- Generating written plans, outlines, and guides
- Initiating topics for collaborative learning and discussion
- Guiding communication with various audiences

Myriad use cases and humanlike performance on the above tasks are largely what make LLMs attractive for use in education, among many other potential applications (Dwivedi et al. 2023).

OBJECTIVES OF ETHICS EDUCATION IN MEDICAL SCHOOL

Johns Hopkins School of Medicine was the first medical school to incorporate ethics into its curriculum in

Figure 1. SWOT graphic on effective and equitable use of LLMs in bioethics training and education.

Table 1. Table of terms.

Table of Terms	
Algorithm[*]	A clearly specified mathematical process for computation; a set of rules that, if followed, will give a prescribed result.
Artificial Intelligence	The field and science of making intelligent machines.
GPT	Generative Pre-trained transformer
LLM	Large Language Model—a type of machine learning model that performs natural language processing tasks to predict and generate responses to user queries in ways that mimic conversational patterns between humans.
Machine learning[*]	A field within artificial intelligence focused on the ability of computers to learn from provided data without being explicitly programmed for a particular task
Model[*]	A detailed description or scaled representation of one component of a larger system that can be created, operated, and analyzed to predict actual operational characteristics of the final produced component.
NLP[*]	Natural Language Processing—branch of computer science that uses artificial intelligence and machine learning to analyze human text and speech data.

[*]Excerpted from the official glossary of the National Institutes for Standards and Technology.

1977 (Pierson 2022). It was not until 1985, however, that ethics became mainstream in medical curricula upon publication of a seminal report (Culver et al. 1985) that stressed the importance of ethics and professionalism training in medical schools (Carrese et al. 2015). Medical ethics has since been more routinely integrated into medical school curricula, and "adherence to ethical principles" is now a core competency required of all medical school graduates in the U.S (Obeso et al. 2017). A 2022 survey of 87 U.S. medical schools found that 79% required a formal ethics course (DuBois and Burkemper 2002). Although there is consensus on the importance of teaching ethics, there is heterogeneity in the quality and extent of ethics education.

Members of a working group funded through the Project to Rebalance and Integrate Medical Education (PRIME) prepared a new report in 2015 (Romanell Report), which specifically aimed to support ethics educators in "meeting the articulated expectations of accrediting organizations" by addressing "aspects of medical ethics education" such as program goals and objectives, as well as teaching methods (Carrese et al. 2015, 745). Without proposing one singular method or approach, the PRIME working group suggested that medical school curricula in ethics and professionalism should "strive to move learners from knowledge acquisition and skills development to behavior change in which excellent patient care is the goal" (Carrese et al. 2015, 747). They propose several hallmarks of how ethics educators facilitate this development through effective teaching and learning, including sustaining changes in reasoning behavior through longitudinal exposure to ethics concepts and adopting "learner-driven strategies" that empower trainees to proactively identify clinical cases or experiences that raise acute ethical issues, among others.

Importantly, the authors of the Romanell Report acknowledge that many institutions lack faculty with appropriate training to deliver high quality ethics education, while others devalue ethics teaching by making it a voluntary or unremunerated activity (Carrese et al. 2015). They also acknowledge that technology advances will continue to broaden ethics teaching methods and approaches, even those that "complement (or replace) the traditional approach of in-person lectures" (Carrese et al. 2015, 747). Educators should be encouraged to "be open to these innovations and carefully evaluate which content is best delivered by new technologies" according to the authors (Carrese et al. 2015, 747).

Our own ethics program has two related goals: to teach students to master a core fund of knowledge of ethics in practice, and to improve students' ethical reasoning skills. We have found that prioritizing problem- and case-based learning helps students apply ethical concepts and develop a well-justified plan for case management through completing a five-step *Ethics Work Up*. Students are guided in their ethical reasoning about complex cases and are taught how to articulate ethical dilemmas, identify and analyze case-relevant information, source external policies, guidelines, and laws to guide ethical action, and make recommendations that are both evidence-based and ethically justified (Fasser et al. 2007; Baylor College of Medicine, Center for Medical Ethics and Health Policy).

Given the expectations of modern medical education, is it possible that LLMs like chatGPT can help achieve the goals of ethics education in an efficient and equitable way? Might these tools even make formal ethics education obsolete? We experimented with chatGPT to help us explore answers to some of these questions. In this paper, we discuss some of the strengths, and expose potential limitations of chatGPT as a tool for undergraduate medical ethics education using the programmatic goals at our own institution as a scaffold for our analysis.

ETHICS WORK UP USING ChatGPT

We prompted chatGPT to analyze a case involving a woman refusing a needed c-section based on the *Ethics Work Up* we use in our own bioethics curriculum for

first-year medical students and other health professions trainees (Baylor College of Medicine, Center for Medical Ethics and Health Policy). We did not give chatGPT details of the case nor did we define or explain the framework for the *Ethics Work Up*. We asked chatGPT to "do an ethics work up on a case of a woman refusing a needed c-section." ChaptGTP wrote the following case (which is very similar to the case we give our students) and proceeded to analyze it.

> *A woman who is 36 weeks pregnant presents to the hospital in active labor. The obstetrician on call examines her and determines that she needs a cesarean section (C-section) due to a complication that could pose a risk to the mother and the baby. However, the woman refuses the C-section and insists on a vaginal delivery.*

The entire exchange and analysis from ChatGPT (query run on) took less than 5 minutes. Parts of this exchange are excerpted in Box 1. The full exchange from the clinical ethics case can be accessed in the Supplementary Materials 1, which also features additional queries we tested on chatGPT that touch on three other contested topics in bioethics, including for professionalism and public health ethics.

PART IV

Strengths

The full series of chatGPT queries and responses demonstrates strengths in summing up core ethical principles and concepts relevant to the case presented. It is well-suited to orienting students to how ethical principles may be balanced and identifying where value conflicts may lie. In responding this way, chatGPT closely mirrors the discursive processes common to clinical ethics case conferences and aligns with problem-based learning techniques used in the classroom. For example, chatGPT appropriately invokes the four ethical principles according to Beauchamp and Childress (autonomy, beneficence, non-maleficence and justice). It summarizes the strongest ethical arguments for allowing the patient to refuse the C-section (respect for her decision as someone with capacity). ChatGPT does not prescribe an action one way or the other, but rather emphasizes that the resulting decision should take the best interests of both woman and baby into account and weigh these against the three other principles. ChatGPT's responses furthermore demonstrate a reasonable standard for how students effectively present a clinical case, draw out ethically significant details from the case, weigh conflicting ethics appeals, and arrive at an ethically-sound recommendation.

ChatGPT clarifies in its response that its algorithmic properties prevent it from being a moral agent and ultimately taking a concrete position on the most appropriate action in this case. In lieu of proscribing moral actions, chatGPT helpfully describes the central ethical tension between human agents (medical student, trainee, bioethicist, healthcare professional). The woman's refusal, according to chatGPT, creates an ethical dilemma for the physician between respecting the woman's autonomy and a duty to act in the best interests of the unborn baby as per fiduciary obligations of all physicians.

We believe chatGPT's response to the State-specific laws regarding maternal and fetal health decisions was also especially useful. Although chatGPT correctly references statutes in the Texas Health and Safety Code, it does not point to court cases or other relevant policy documents that may be important for students to know. Furthermore, chatGPT currently only fetches data prior to 2021 so chatGPT would not capture any changes in the Texas law or ACOG guidelines if enacted after 2021. A strategic partnership between OpenAI and Microsoft is poised to solve this problem by expanding and updating chatGPT's web browsing capabilities (Metz and Weise 2023).

When prompted, chatGPT was also thorough in its analysis through various other bioethical frameworks (e.g., feminist bioethics, virtue ethics) and provided recommendations that were consistent with the theoretical epistemological commitments of those traditions at basic levels of depth and nuance that educators might teach medical students or other health professionals in training.

If the point of bioethics education is to give students core knowledge about ethics in practice, and to give them tools to analyze their way through difficult ethical issues, chatGPT was helpful toward achieving both goals. To put the point more provocatively, why do students and trainees need bioethics education when they have a tool like chatGPT that can give them almost instant access to ethics relevant knowledge and analysis about almost any ethics issue they face in clinical practice? This is where we turn to some things that chatGPT *cannot* provide and for which bioethics education and training is still needed.

WEAKNESSES

Cannot Help Students Identify or Spot Ethical Issues

The more useful exchanges with chatGPT proceeded when the user already identified what ethical issues to

Box 1. Prompts and responses from chatGPT in reference to a clinical ethics case. [Run on chatGPT4, Open AI. 14 March 2023]

Prompt: Complete an ethics work up on a case of a woman refusing a needed c-section.

ChatGPT: Response:

Autonomy: The woman has the right to make decisions about her own body, including whether or not to undergo a C-section. As an adult with decision-making capacity, she has the right to refuse any medical treatment, even if it is recommended by a doctor. However, her decision to refuse the C-section may have serious consequences for her and her baby, and it is the responsibility of the healthcare providers to provide her with accurate information about the risks and benefits of the different options available to her.

Beneficence: The obstetrician's primary duty is to promote the well-being of both the mother and the baby. In this case, the obstetrician has determined that a C-section is necessary to prevent harm to the mother and the baby. Therefore, it is the obstetrician's duty to provide the woman with the necessary information about the risks and benefits of the different options and to encourage her to make a decision that is in the best interest of both her and her baby.

Non-maleficence: The obstetrician also has a duty to avoid causing harm to the woman and the baby. If the obstetrician determines that a C-section is necessary to prevent harm, it would be unethical to proceed with a vaginal delivery against the woman's wishes. However, if the obstetrician determines that a vaginal delivery is safe for both the mother and the baby, it would be ethical to respect the woman's decision to refuse the C-section.

Justice: The principle of justice requires that healthcare resources be allocated fairly and equitably. In this case, the healthcare resources necessary for a C-section would be diverted from other patients who may also need them. However, justice also requires that healthcare providers respect the autonomy of patients and provide them with the necessary care, regardless of the cost or resources required.

Conclusion: In this case, the woman has the right to refuse a C-section, but the obstetrician has a duty to provide her with accurate information about the risks and benefits of the different options available to her. The obstetrician should encourage the woman to make a decision that is in the best interest of both her and her baby, while also respecting her autonomy. If the obstetrician determines that a C-section is necessary to prevent harm to the woman or the baby, it would be unethical to proceed with a vaginal delivery against the woman's wishes. However, if the obstetrician determines that a vaginal delivery is safe for both the mother and the baby, it would be ethical to respect the woman's decision to refuse the C-section. Ultimately, the decision should be made in the best interest of the woman and the baby, while also taking into account the principles of autonomy, beneficence, non-maleficence, and justice.

Prompt: Does the woman have a moral right to refuse the c-section if it will result in the death of her baby?

ChatGPT: As a language model, I cannot take a moral position on this issue, but I can provide information that may be relevant to the question. In general, the woman has the right to make decisions about her own body, including whether or not to undergo a medical procedure like a C-section. However, this right is not absolute and may be limited by considerations such as the potential harm to others. If the obstetrician determines that a C-section is necessary to prevent harm to the baby, then the woman's refusal of the procedure may be seen as a conflict between her autonomy and the duty to protect the baby's life. In such cases, the healthcare providers may need to consider whether there are any legal or ethical obligations to intervene to protect the life of the baby. Ultimately, the specific circumstances of the case would need to be evaluated to determine the appropriate course of action, taking into account the relevant ethical principles and legal considerations.

Prompt: What are the laws in Texas about a mother refusing a c-section that is needed to save her baby?

ChatGPT:In Texas, the law generally respects a patient's right to refuse medical treatment, including a C-section, even if it is necessary to save the life of their baby. The Texas Health and Safety Code, section 166.046, states that a person has the right to refuse medical treatment or to request that medical treatment be withdrawn, even if the refusal or withdrawal of treatment may result in the person's death or the death of an unborn child. However, there are some situations in which the law allows healthcare providers to intervene to protect the life of an unborn child, even if it means going against the patient's wishes. For example, if a court determines that a pregnant woman's decision to refuse medical treatment is endangering the life of her unborn child, the court may order medical intervention, including a C-section, to protect the life of the child. It is important to note that the circumstances of each case are unique and the laws and legal procedures may vary. It is advisable to seek legal advice from a qualified attorney if you have questions about your legal rights and obligations in a particular situation.

Prompt: What are the biggest objections to the decision to respect the autonomy of the mother and let her refuse the c-section needed to save her baby?

ChatGPT: There can be several objections to the decision to respect the autonomy of the mother and allow her to refuse a C-section that is needed to save her baby. Some of the objections are:

1. *The duty to protect the life of the baby: Healthcare providers have a duty to protect the life and health of their patients, including unborn babies. If the physician believes that a C-section is necessary to save the baby's life, then the physician may argue that they have an obligation to intervene, even if it means going against the mother's wishes.*

2. *Legal and ethical obligations: In some cases, the law and ethical principles may require healthcare providers to protect the life and health of an unborn child, even if it means going against the mother's wishes. For example, the Texas Penal Code section 19.02 defines murder as intentionally causing the death of an individual, including an unborn child. This means that if a physician does not take action to save the life of an unborn child, they may be held criminally liable if the child dies.*

3. *Potential harm to others: The decision to refuse a C-section may result in harm not only to the mother but also to the baby. In some cases, the mother's decision to refuse the procedure may be seen as a conflict between her autonomy and the duty to protect the life and health of her baby.*

4. *Emotional and psychological trauma: Refusing a C-section that is needed to save the life of a baby may have long-lasting emotional and psychological consequences for the mother, such as guilt, regret, and grief.*

It is important to note that each case is unique, and healthcare providers must carefully consider the specific circumstances before making a decision about whether to respect the mother's autonomy or intervene to protect the life of an unborn child.

Prompt: What are the biggest objections to the decision to perform the c-section against the wishes of the mother to protect the life of the baby?

ChatGPT: There can be several objections to the decision to perform a C-section against the wishes of the mother to protect the life of the baby. Some of the objections are:

5. *Violation of the mother's autonomy: Respect for patient autonomy is a fundamental principle in medical ethics, and performing a medical procedure against the wishes of the patient can be seen as a violation of this principle.*

6. *Potential harm to the mother: Performing a C-section against the wishes of the mother may pose risks to the mother's physical and psychological health, including complications from the surgery, trauma, and resentment toward the healthcare providers.*

7. *Legal and ethical implications: The decision to override the mother's wishes and perform a medical procedure on her body without her consent raises legal and ethical questions about the extent of healthcare providers' authority over patient decision-making.*

8. *Impact on the doctor-patient relationship: Overriding the mother's wishes may damage the trust and rapport between the healthcare provider and the patient, and may discourage the patient from seeking medical care in the future.*

It is important to note that the decision to perform a medical procedure against the wishes of the patient should only be taken as a last resort when there is a serious and immediate threat to the patient's life or health, and all reasonable efforts to obtain the patient's consent have been exhausted. Each case is unique, and healthcare providers must carefully weigh the potential risks and benefits of their actions while adhering to ethical and legal principles.

Prompt: What would a feminist philosopher say about this case?
ChatGPT: Feminist philosophers approach ethical issues through a lens that takes into account the impact of social, cultural, and political factors on the experiences and perspectives of individuals, especially those who have been historically marginalized or oppressed.
In the case of a woman refusing a C-section that is needed to save her baby, a feminist philosopher may examine the context of the decision, including the social and cultural factors that may be influencing the mother's decision, such as access to healthcare, autonomy, and reproductive rights.
They may also consider the power dynamics involved in the decision-making process, such as the authority of healthcare providers and the extent to which the mother's autonomy is being respected. They may advocate for a decision-making process that is collaborative, transparent, and centered on the well-being and agency of the mother.
Furthermore, a feminist philosopher may also consider the broader social and political implications of the decision, such as the impact on reproductive rights and healthcare access for marginalized groups. They may advocate for a decision that promotes the autonomy and agency of the mother while also taking into account the social and cultural factors that may be influencing her decision.
Overall, a feminist philosopher would approach this case with a sensitivity to the complex and intersecting social and cultural factors that may be at play, while advocating for a decision-making process that is centered on the well-being and agency of the mother.

Prompt: How should a clinician manage this case?
Chat GPT: Managing a case like this requires a delicate balance between respecting the autonomy of the patient and ensuring the safety and well-being of the fetus. Here are some general guidelines that a clinician could follow:
Establish a clear and open dialogue with the patient. The clinician should explain the risks and benefits of both the C-section and vaginal delivery options, and should take the time to understand the patient's concerns and values.
Provide the patient with accurate and unbiased information. The clinician should provide the patient with evidence-based information about the risks and benefits of each option, and should avoid presenting their own biases or opinions.
Involve a multidisciplinary team. The clinician should involve a team of healthcare professionals, including obstetricians, neonatologists, and medical ethicists, to provide the patient with a comprehensive understanding of the situation and to help guide decision-making.
Consider obtaining a court order. If the patient refuses the C-section and the situation is urgent, the clinician may need to consider obtaining a court order to override the patient's refusal.
Document the decision-making process. The clinician should thoroughly document the discussions and decisions made with the patient and the healthcare team in the patient's medical record.
It is important to note that each case is unique and requires individualized attention. The clinician should approach the case with sensitivity and respect for the patient's autonomy, while also ensuring the safety and well-being of the fetus.

Prompt: How should a clinician communicate to the woman the decision to perform a c-section against the woman's wishes?
ChatGPT: Communicating a decision to perform a C-section against the woman's wishes is a difficult and sensitive matter, and it should be done with compassion and empathy for the woman's situation. Here are some general guidelines that a clinician could follow:
Explain the reasons for the decision. The clinician should explain to the woman why the C-section is necessary to ensure the safety and well-being of the fetus.
Acknowledge the woman's autonomy. The clinician should acknowledge the woman's right to make decisions about her own body and healthcare, but explain that in this situation, the safety of the fetus must take priority.
Provide emotional support. The clinician should offer emotional support to the woman during this difficult time and acknowledge the emotions that she may be experiencing.
Involve the woman in the decision-making process as much as possible. Even though the decision to perform a C-section against the woman's wishes has already been made, the clinician should still involve the woman in the decision-making process as much as possible, such as deciding who will be present during the procedure or how the woman would like to receive information about the baby's condition.
Document the communication. The clinician should document the communication in the patient's medical record, including the reasons for the decision, the woman's reaction, and any support that was offered.
It is important to approach this situation with sensitivity and empathy for the woman's situation, while also prioritizing the safety and well-being of the fetus.

inquire about in the first place. ChatGPT is thus unable to identify ethical challenges *for* the user, requiring them to have gained prerequisite skills in identifying specific ethical dilemmas, issues, and concepts they encounter in a particular case.

ChatGPT Cannot Teach Caring, Compassion, or Moral Motivation to be an Ethical Practitioner

Another notable weakness with LLM-based ethics training is that it is less likely to instill an ethic of care in students, to act compassionately toward patients and their families when ethical dilemmas arise in the clinic, or to care about doing the right thing (which is often hard for humans). This limitation inevitably invigorates the longstanding debate about whether compassion and moral motivation can be taught at all in medical education (Pence 1983); Leget and Olthuis have argued categorically that it can (Leget and Olthuis 2007). The authors propose that "[i]n the same way as caring is an activity that can be learnt, perceiving the moral dimension of medical

practice is also a skill that can be acquired to a greater or lesser extent" using educational tools that combine opportunities for personal reflection/reflexivity and practical experience (Leget and Olthuis 2007, 6119). Wear and Zarconie corroborate this in their qualitative study of fourth year medical students. Students reported that having positive role models and self-reflection deepened their understanding of how to practice medicine with compassion, altruism, and respect for patients in mind (Wear and Zarconi 2008).

LLMs are unable at this stage to feel care or compassion; however, chatbots such as chatGPT have already demonstrated capacities to simulate compassionate dialogue in ways that resemble real human expressions (Ayers et al. 2023). Efforts to develop general artificial intelligence (AGI) build on more than two decades of research focused on equipping machines with the capacity to express emotions and translate that ethos to human actors.

Limited Evidence Demonstrating ChatGPT's Effectiveness as a Pedagogical Tool

Despite an extensive literature focused on pedagogy (Miles et al. 1989; Fox et al. 1995; Goldie 2000), the field still lacks consensus on the most effective teaching approaches, methods, and tools for helping health professionals build core competencies and instill professionalism in medical ethics. Moreover, there is debate about whether widespread adoption of more advanced LLMs in medicine could in the future also limit the value add of human-to-human medical ethics training (Cohen 2023). There are mixed findings about whether ethics education effectively achieves its intended goals irrespective of the pedagogical approach adopted (de la Garza et al. 2017; Ignatowicz et al. 2023). This is partly because institutions infrequently share the same goals for medical ethics education. When there are common goals, educators infrequently evaluate their effectiveness using standardized methods. LLMs could add complexity to evaluation methods.

In response we emphasize the need, as others previously have (Fox et al. 1995; Eckles et al. 2005; Campbell et al. 2007), for rigorous research to inform these methods. Specifically, ethics educators will require studies that compare learning outcomes for chatGPT and future LLMs used alone, as well as in combination with current methods used in ethics education to meet programmatic goals (e.g., case discussions, problem-based learning, peer review). Implementation of standardized evaluation measures across programs would help the field build the evidence base for effective bioethics pedagogy.

Requires Skills in Prompt Engineering

A key component to the successful use of chatGPT is knowing how to effectively leverage LLMs to provide the most comprehensive responses. Quality outputs significantly rely on the literacy of the user, and human-computer interfaces as well as user experience affect usability. In the prior case example (Box 1), chatGPT provided a far more detailed analysis and recommended actions when asked *what are the biggest objections to the decision to perform the c-section against the wishes of the mother to protect the life of the baby?* than in response to the prompt *complete an ethics work up on the case of a women refusing a c-section.*

ChatGPT and other improved LLMs require some technical literacy on the part of users, including understanding when and how to prompt models to achieve desired outputs. The skill of designing queries that yield quality outputs using LLMs is referred to in the industry as prompt engineering (What is a Large Language Model (LLM)? 2022) and will continue to be a critical user competency as LLMs become more sophisticated. There are both formal and informal methods that bioethics educators could use to help advance students' and trainees' skills in prompt engineering for ethics. These include providing tutorials and live demonstrations like in this article or referring students to free online user forums and training modules. As such, the need for prompt engineering may be viewed less as a weakness of the technology itself than of the imaginative potential and skill of its users.

Default Ethical Reasoning

In our experiment, chatGPT initially invoked a principlist approach to analyze the case. ChatGPT alluded to other ethical frameworks, such as virtue ethics, or feminist ethics, only when explicitly prompted. This default choice of framework is consequential for ethics training. First, chatGPT's response implies a hierarchy of ethical frameworks atop which principlism is presumed to sit. As we explain in earlier sections, chatGPT is trained on the corpus of knowledge available on the Internet. The four principles framework indeed dominates much of the available bioethics content. It is therefore unsurprising that chatGPT defaults to principlism, but this is problematic considering other frameworks are equally important and

legitimate. Indeed, some frameworks may be more appropriate than principlism to analyze the case. Second, chatGPT risks sidelining non-principlist frameworks—many of which are grounded in the moral values and experiences of marginalized groups—by adding barriers to learn about them. Such limitations may diminish as engineers discover strategic ways of "weighting" web-based information to better balance perspectives. Third, reliance on chatGPT for knowledge acquisition in the context of bioethics training could constrain trainees' development as they work to build ethical reasoning skills based on a limited suite of analytical tools that chatGPT selectively adopts.

Prone to Bias

As other authors in this special issue explain in detail, LLMs are only as useful as the quality of the data (i.e. searchable content) upon which they are trained. Most academic bioethics knowledge is contained in peer reviewed journals behind subscription pay walls and outside the reach of most LLMs that rely on free text available on the Internet. Even when bioethics knowledge can be sourced, LLMs cannot filter out inaccuracies that may be already baked into this knowledge, raising additional concerns about the fact that LLMs do not vet the information they relay. ChatGPT has been shown to fabricate this information outright (Bender et al. 2021) and counteract the goals of bioethics education as a result because they impede student's development of ethical reasoning skills.

LLMs can therefore be prone to bias; recent empirical studies corroborate this inconvenient truth (Suguri Motoki et al. 2023; Li et al. 2023). Prior to running our test, we anticipated that chatGPT would present imbalanced arguments when presented with an ethical dilemma. The case example, however, showed otherwise. The chatGPT-generated response weighed ethical pros and cons of alternative courses of action to the woman's refusal in ways we would expect students to do in their own *Ethics Work Up*. However, future research should test whether the ability to do this generalizes to other controversial topics in bioethics where minority viewpoints are less prominent on the Internet.

OPPORTUNITIES

Autodidactic Learning beyond the Classroom

Notwithstanding drawbacks related to representativeness, LLMs like chatGPT may offer greater access to bioethics information than what may be afforded by traditional search engines. In this way, LLMs may enhance democratization of bioethics information and awareness of core concepts and theories by offering simplified and more "human" search strategies (e.g., just type [or soon speak] your question into chatGPT) than what many of us have learned to employ using "keywords" in traditional search engines. Given that many novices may not even know which keywords to employ (a problem similar to prompt engineering), LLMs that offer more conversational (i.e., "human"-sounding) exchanges may enhance the accessibility of bioethics summations, meta-analyses, or even guided reflections to wider and more diverse audiences without the need for special effort or expertise.

Hybrid Intelligence

The appeal and ease of using LLMs offer unprecedented opportunities to blend human-machine intelligence in potentially productive ways. Scholars argue that combining human and artificial intelligence into "hybrid" intelligence systems holds unique promise for accomplishing complex goals with results superior to those of either humans or machines alone. Such systems also allow humans and machines to continuously improve by learning from each other, just as chatGPT3 learned from millions of users after its early (some say premature) release and folded these insights into a vastly improved chatGPT4. Lessons from "content creators" online who use generative AI to provide fodder for creative writing, art, music, and other forms of entertainment reveal that generative AI outputs can provide an informational "head start" for humans to then apply their intellect and creativity to innovate further. These same user principles can carry over to bioethics education. That is, LLM-generated outputs, recommendations, or the start of academic writing should be used to inspire critical reflection, rather than the final product of that reflection. Approaching chatGPT use in this way can be compatible with the goal of building ethical reasoning skills because it forces students to consider a particular ethical position, challenge the ethical arguments underpinning it, and revise when appropriate. Experimenting with LLMs and related tools can therefore help curious thinkers learn about ethics in new and engaging ways. For some, chatGPT-assisted ethics education could improve the quality of the learning experience given the lower stakes learning environment and absent the power dynamics and social pressures common in university classroom settings.

THREATS

Dilutes the Information Seeking Process

With increasing accuracy and sophistication, LLMs have been designed to generate text strings that we as humans could interpret as authoritative persuasion. This is a phenomenon sometimes referred to as algorithmic neutrality (Ames and Mazzotti 2023). Our field of bioethics should care about the threats that blind trust in LLMs could pose for future trainees because it marginalizes ethical reflection and dilutes the information seeking process. As Bender and Chah argue, "more often than not, we benefit more from engaging in sense-making: refining our question, looking at possible answers, understanding the sources those answers come from and what perspectives they represent"(Bender et al. 2021). Decades have been spent trying to humanize medicine by integrating robust ethics education in clinical training and promoting reflexive practice—chatGPT, if over-relied on, threatens to undermine that progress. However, user training and awareness, as well as use of interface elements that strategically encourage reflection among users may help to mitigate this concern (Kostick-Quenet and Gerke 2022).

New Source of Educational Inequities

Much of the empirical literature on applications of LLMs in medical (Han et al. 2023; Sallam 2023) as well as allied health professional education (Sallam et al. 2023) compares test performance (Gilson et al. 2023; Kung et al. 2023; Mbakwe et al. 2023), and evaluates the quality of diagnostic prompts and outputs as well as training data (Johnson et al. 2023). An under-acknowledged threat, however, is the potential of LLMs to introduce new medical education-related inequities.

OpenAI currently offers free versions of chatGPT, including its new and improved GPT-4 version. But future versions may not remain free. For example, OpenAI introduced a subscription service to chatGPT Plus that costs USD$20 per month. The company claims that the subscription-based version uses the free version as its primary training dataset and generates higher quality outputs as a result. OpenAI and other developers may offer only tiered access to more powerful LLMs in the future through paid subscriptions, disadvantaging users who relied on the free versions. Basing access to more reliable LLMs on ability and willingness to pay creates obvious equity issues insofar as these technologies confer academic advantages or become essential resources like having a personal laptop. Indeed, extant research shows educational inequities grow when academic performance/success is tied to technology access (see for example Fairlie 2012 and Gonzales et al. 2020).

While this trend is not unique to LLMs, it nevertheless threatens to widen existing inequities at both the institution and individual student levels in bioethics. Under-resourced institutions that offer bioethics training programs and low-income students are likely to be at the greatest disadvantage not only for their lack of LLM access but also for the specific advantages that LLMs offer for bioethics knowledge acquisition; constructing ethical arguments and invoking evidence-based standards to support reasoned actions. These trends may compound existing inequities (Webber and Burns 2022) among graduate students in the humanities and other non-STEM programs who are paid less on average than their counterparts in STEM programs, are less successful on the job market, and graduate with higher debt (Burns and Webber 2019).

CONCLUSION

We experiment with chatGPT in the context of bioethics education for healthcare professional trainees. By prompting chatGPT to analyze a clinical ethics case using the strategy we teach to first-year medical students, we consider whether the overall strengths and opportunities overcome the possible weaknesses and threats of incorporating chatGPT as a viable teaching tool within our own ethics curriculum. Educators and institutions alike are contending with the post-chatGPT world in which we now live and work. ChatGPT warrants consideration as part of a suite of teaching and learning tools in ethics education. We believe strategic guidance is needed to address imminent quality and equity issues for teaching and learning in our field. We also argue that policies surrounding permissible uses of chatGPT in the ethics classroom should undergo periodic evaluation and be adapted as the technology evolves.

LLMs are poised to transform the world around us. Indeed, academic medicine will not be immune to this change. To ignore this progress would be to knowingly underprepare future clinicians for how to ethically navigate the emerging landscape of medical practice wherein humans and machines work in closer confines than they ever have before. Yet what LLMs lack in moral motivation, creativity, and critical thinking, humans will always need to be able to demonstrate to solve complex ethical problems. When

prompted whether bioethicists are still needed in the era of LLMs, GPT-4 responded:

> *ChatGPT and other AI tools can certainly help to inform and guide ethical decision-making in medicine, but they cannot replace the human element of moral reasoning and critical reflection that bioethicists provide… Bioethicists are uniquely qualified to navigate these complex issues and ensure that healthcare practices are grounded in ethical principles and human values.*

In that answer, we can trust.

FUNDING

ORCID

Vasiliki Rahimzadeh http://orcid.org/0000-0003-3537-7601
Kristin Kostick-Quenet http://orcid.org/0000-0003-2510-0174
Jennifer Blumenthal Barby http://orcid.org/0000-0003-1054-619X
Amy L. McGuire http://orcid.org/0000-0002-7819-519X

REFERENCES

Ames, M. G., and M. Mazzotti. 2023. *Algorithmic Modernity: Mechanizing Thought and Action, 1500-2000.* Oxford University Press.

Asch, D. A. 2023. An interview with ChatGPT about health care. *Catalyst Non-Issue Content* 4 (2):1–6. doi:10.1056/CAT.23.0043.

Ayers, J. W., A. Poliak, M. Dredze, E. C. Leas, Z. Zhu, J. B. Kelley, D. J. Faix, A. M. Goodman, C. A. Longhurst, M. Hogarth, et al. 2023. Comparing physician and artificial intelligence Chatbot responses to patient questions posted to a public social media forum. *JAMA Internal Medicine* 183 (6):589. doi:10.1001/jamainternmed.2023.1838.

Baylor College of Medicine, Center for Medical Ethics and Health Policy. Ethics Work-Up. https://www.bcm.edu/sites/default/files/BCM-Ethics-Work-Up.pdf

Bender, E. M., T. Gebru, A. McMillan-Major, and S. Shmitchell. 2021. On the dangers of stochastic parrots: can language models be too big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–23. Virtual Event Canada: ACM. doi:10.1145/3442188.3445922.

Burns, R., and K. L. Webber. 2019. Achieving the promise of educational opportunity: Graduate student debt for STEM vs. non-STEM students, 2012. *Journal of Student Financial Aid* 48 (3). doi: 10.55504/0884-9153.1659.

Campbell, A. V., J. Chin, and T.-C. Voo. 2007. How can we know that ethics education produces ethical doctors? *Medical Teacher* 29 (5):431–6; Taylor & Francis Ltd: 431–436. doi: 10.1080/01421590701504077.

Carrese, J. A., J. Malek, K. Watson, L. S. Lehmann, M. J. Green, L. B. McCullough, G. Geller, C. H. Braddock, and D. J. Doukas. 2015. The essential role of medical ethics education in achieving professionalism: The Romanell report. *Academic Medicine : journal of the Association of American Medical Colleges* 90 (6):744–52. doi:10.1097/ACM.0000000000000715.

Cohen, I. G. 2023. What should ChatGPT mean for bioethics? *SSRN Electronic Journal.* doi:10.2139/ssrn.4430100.

Cotton, D. R. E., P. A. Cotton, and J. R. Shipway. 2023. Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International* :1–12. doi:10.1080/14703297.2023.2190148.

Cox, C., and E. Tzoc. 2023. ChatGPT: Implications for academic libraries. *College & Research Libraries News* 84 (3):99–102. doi:10.5860/crln.84.3.99.

Culver, C. M., K. D. Clouser, B. Gert, H. Brody, J. Fletcher, A. Jonsen, L. Kopelman, J. Lynn, M. Siegler, and D. Wikler. 1985. Basic curricular goals in medical ethics. *The New England Journal of Medicine* 312 (4):253–6. doi:10.1056/NEJM198501243120430.

de la Garza, S., V. Phuoc, S. Throneberry, J. Blumenthal-Barby, L. McCullough, and J. Coverdale. 2017. Teaching medical ethics in graduate and undergraduate medical education: A systematic review of effectiveness. *Academic Psychiatry* 41 (4):520–5. doi:10.1007/s40596-016-0608-x.

DuBois, J. M., and J. Burkemper. 2002. Ethics education in U.S. medical schools: A study of syllabi. *Academic Medicine* 77 (5):432–7. doi:10.1097/00001888-200205000-00019.

Dwivedi, Y. K., N. Kshetri, L. Hughes, E. L. Slade, A. Jeyaraj, A. K. Kar, A. M. Baabdullah, A. Koohang, V. Raghavan, M. Ahuja, et al. 2023. So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management* 71:102642. doi:10.1016/j.ijinfomgt.2023.102642.

Eckles, R. E., E. M. Meslin, M. Gaffney, and P. R. Helft. 2005. Medical ethics education: Where are we? Where should we be going? A review. *Academic Medicine: Journal of the Association of American Medical Colleges* 80 (12):1143–52. doi: 10.1097/00001888-200512000-00020.

Fasser, C., A. McGuire, K. Erdman, D. Nadalo, S. Scott, and V. Waters. 2007. The ethics workup: A case-based approach to ethical decision-making instruction. *The Journal of Physician Assistant Education* 18 (1):34–41. doi:10.1097/01367895-200718010-00006.

Fox, E., R. M. Arnold, and B. Brody. 1995. Medical ethics education: Past, present and future. *Academic Medicine* 70 (9):761–8.

Gilson, A., C. W. Safranek, T. Huang, V. Socrates, L. Chi, R. A. Taylor, and D. Chartash. 2023. How does ChatGPT perform on the united states medical licensing examination? The implications of large language models for medical education and knowledge assessment. *JMIR Medical Education* 9 (1):e45312. doi:10.2196/45312.

Goldie, J. 2000. Review of ethics curricula in undergraduate medical education. *Medical Education* 34 (2):108–19. doi:10.1046/j.1365-2923.2000.00607.x.

Gonzales, A. L., J. McCrory Calarco, and T. Lynch. 2020. Technology problems and student achievement gaps: A validation and extension of the technology maintenance construct. *Communication Research* 47 (5):750–70. doi: 10.1177/0093650218796366.

Han, Z., F. Battaglia, A. Udaiyar, A. Fooks, and S. R. Terlecky. 2023. An explorative assessment of ChatGPT as an aid in medical education: Use it with caution. *medRxiv*. doi:10.1101/2023.02.13.23285879.

Ignatowicz, A., A. M. Slowther, C. Bassford, F. Griffiths, S. Johnson, and K. Rees. 2023. Evaluating interventions to improve ethical decision making in clinical practice: A review of the literature and reflections on the challenges posed. *Journal of Medical Ethics* 49 (2):136–42. doi:10.1136/medethics-2021-107966.

Johnson, D., R. Goodman, J. Patrinely, C. Stone, E. Zimmerman, R. Donald, S. Chang, S. Berkowitz, A. Finn, E. Jahangir, et al. 2023. Assessing the accuracy and reliability of AI-generated medical responses. An evaluation of the Chat-GPT model. In review. doi:10.21203/rs.3.rs-2566942/v1.

Kharya, P., and A. Alvi. 2021. Using DeepSpeed and Megatron to train Megatron-Turing NLG 530B, the world's largest and most powerful generative language model. *NVIDIA Technical Blog*, October 11

Kostick-Quenet, K. M., and S. Gerke. 2022. AI in the hands of imperfect users. *npj Digital Medicine* 5 (1):6. doi:10.1038/s41746-022-00737-z.

Krügel, S., A. Ostermaier, and M. Uhl. 2023. ChatGPT's inconsistent moral advice influences users' judgment. *Scientific Reports* 13 (1):4569. doi:10.1038/s41598-023-31341-0.

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

Leget, C., and G. Olthuis. 2007. Compassion as a basis for ethics in medical education. *Journal of Medical Ethics* 33 (10):617–20. doi: 10.1136/jme.2006.017772.

Li, J., A. Dada, J. Kleesiek, and J. Egger. 2023. ChatGPT in healthcare: A taxonomy and systematic review. *Health Informatics*. Preprint. doi: 10.1101/2023.03.30.23287899.

Liebrenz, M., R. Schleifer, A. Buadze, D. Bhugra, and A. Smith. 2023. Generating scholarly content with ChatGPT: Ethical challenges for medical publishing. *The Lancet. Digital Health* 5 (3):e105–e106. doi:10.1016/S2589-7500(23)00019-5.

Mbakwe, A. B., I. Lourentzou, L. A. Celi, O. J. Mechanic, and A. Dagan. 2023. ChatGPT passing USMLE shines a spotlight on the flaws of medical education. Edited by Harry Hochheiser. *PLOS Digital Health* 2 (2):e0000205. doi:10.1371/journal.pdig.0000205.

Metz, C., and K. Weise. 2023. Microsoft to invest $10 billion in OpenAI, the creator of ChatGPT. *The New York Times*, January 23, sec. Business.

Miles, S. H., L. W. Lane, J. Bickel, R. M. Walker, and C. K. Cassel. 1989. Medical ethics education: Coming of age. *Academic Medicine: Journal of the Association of American Medical Colleges* 64 (12):705–14. doi: 10.1097/00001888-198912000-00004.

Obeso, V., B. Brown, M. Aiyer, B. Barron, J. Bull, T. Carter, M. Emery, C. Gillespie, M. Hormann, A. Hyderi, et al. 2017. *Core Entrustable Professional Activities for Entering Residency: Toolkits for the 13 Core EPAs*. American Association of Medical Colleges. https://www.aamc.org/media/20196/download

OpenAI. 2023. GPT-4. Bill of Health: Examining the intersections of health law, biotechnology and bioethics. Harvard Law, Petri Flom Center.

Pierson, L. 2022. Ethics education in U.S. medical schools' curricula. *Bill of Health*.

Sallam, M. 2023. ChatGPT utility in healthcare education, research, and practice: Systematic review on the promising perspectives and valid concerns. *Healthcare* 11 (6):887. doi:10.3390/healthcare11060887.

Sallam, M., N. Salim, M. Barakat, and A. Al-Tammemi. 2023. ChatGPT applications in medical, dental, pharmacy, and public health education: A descriptive study highlighting the advantages and limitations. *Narra J* 3 (1):e103–e103. doi:10.52225/narra.v3i1.103.

Suguri Motoki, F. Y., V. Pinho Neto, and V. Rodrigues. 2023. More human than human: Measuring ChatGPT political bias. *SSRN Electronic Journal*. doi: 10.2139/ssrn.4372349.

Sullivan, M., A. Kelly, and P. McLaughlan. 2023. ChatGPT in higher education: Considerations for academic integrity and student learning. *Journal of Applied Learning & Teaching* 6 (1). doi:10.37074/jalt.2023.6.1.17.

Vk, A. 2023. How ChatGPT broke the AI hype cycle. *Analytics India Magazine*, February 19.

Wear, D., and J. Zarconi. 2008. Can compassion be taught? Let's ask our students. *Journal of General Internal Medicine* 23 (7):948–53. doi: 10.1007/s11606-007-0501-0.

Webber, K. L., and R. A. Burns. 2022. The price of access: Graduate student debt for students of color 2000 to 2016. *The Journal of Higher Education* 93 (6):934–61. doi: 10.1080/00221546.2022.2044976.

What is a Large Language Model (LLM)? 2022. MLQ.ai. December 15.