'Knock, knock… Who’s there?’ ChatGPT and Artificial Intelligence-Powered Large Language Models: Reflections on Potential Impacts within Health and Physical Education Teacher Education

Chad M. Killian¹, Risto Marttinen², Donal Howley³, Julia Sargent⁴, & Emily M. Jones⁵

¹Department of Kinesiology, University of New Hampshire, USA
²School of Education, George Mason University, USA
³Department of Kinesiology, Towson University, USA
⁴Institute of Educational Technology, The Open University, UK
⁵School of Kinesiology & Recreation, Illinois State University, USA

Author Note

Chad M. Killian, PhD; https://orcid.org/0000-0002-8018-7615; Twitter @chadkillian1
Risto Marttinen, EdD; 0000-0003-3807-5684; Twitter @ristomarttinen
Donal Howley, PhD; https://orcid.org/0000-0003-3534-6768; Twitter @andonrua
Julia Sargent, PhD; https://orcid.org/0000-0001-9082-8378; Twitter: @julia_sargent
Emily Jones, PhD; https://orcid.org/0000-0002-3319-1267; Twitter @dremjones28

We have no conflicts of interest to disclose.

Correspondence concerning this article should be sent to New Hampshire Hall, G31, 124 Main Street, Durham, NH, 03824 or emailed to chad.killian@unh.edu
Abstract

This research note suggests the emergence of Artificial Intelligence (AI)-powered chatbots like ChatGPT pose challenges to the future of higher education. We as a field should pay attention to issues and opportunities associated with this technology across learning, teaching, and research spaces. We propose ignoring or being indifferent to predictions about what technologies like AI-powered chatbots can do can cause us to do "dumb things." All H-PETE faculty members should make efforts to learn about these tools to facilitate informed, solution-focused decisions about whether and where to leverage them. We highlight the importance of maintaining socio-critical perspectives when considering use of digital technologies to understand and address digital (in)equity and promote equitable practices. We conclude by emphasizing the need for field-specific consensus statements to guide ethical and appropriate use of AI-powered chatbots, to ensure the value of these tools is harnessed for the good of the society. [Output by ChatGPT-3]…

Keywords: digital technology, ethics, higher education, OpenAI, Chatbots
‘Knock, knock… Who’s there? ChatGPT and Artificial Intelligence-Powered Large Language Models’: Reflections on Potential Impacts within Health and Physical Education Teacher Education

Before you read on, did you read the entire abstract? If not, please read it before proceeding. It was generated by ChatGPT, a popular Artificial Intelligence (AI) driven chatbot. We input text from this article into the chat box, along with a prompt to produce a 150-word abstract. We used the output from this interaction as our abstract, with minor edits (See Figure 1 for the original output). Did you consider that the abstract was written by a chatbot while reading it? What if the rest of this paper is also an AI chatbot output? (It is not…) We ask these questions because, while the stakes for this little meta-commentary are low, the implications are many (Gendron et al., 2022). Our intention with this was not to be misleading, but rather shed light on the challenges AI-powered chatbots like ChatGPT pose to the future of scholarship. The broader purpose of this research note is to draw attention toward some of the added issues and opportunities associated with this emergent technology within learning, teaching, and research.

We recognize there are far reaching impacts beyond areas of faculty work, however these areas are most relevant to our professional roles as academics and represent the focus of this commentary. Our aim is to invite professional peers and colleagues into broader conversation regarding how we might address the emergence of technologies such as ChatGPT and whether there is need for consensus guidelines on the appropriate and ethical use of AI-powered chatbots within our field.

ChatGPT was released for public use in November 2022. It is an AI-powered large language model (LLM), developed by San Francisco-based company, OpenAI. It is designed with the capacity to generate human-like text. There are several AI-powered LLMs in existence
ChatGPT currently represents the most well-known and widely used model. Inputting prompts to the ChatGPT model will demonstrate the considerable capacity this technology has to gather public domain information, synthesize the data and knowledge, and output sensible, technical responses based on user-generated criteria (See www.chat.openai.com).

The release of ChatGPT to the general public is seen by many as a seminal event. As such, conditions are ripe for reactionary and consequential predictions among and across academic disciplines. Much of the current discourse is centered around anticipated disruption and fundamental transformations ChatGPT and AI-powered LLMs will bring to (higher) education, particularly related to how teaching, learning, and research are conducted (Gendron et al., 2022; Hutson, 2022; Nature, 2023). Discussion related to how the fields of health and physical education (should) respond to different technological developments is nothing new (Cassidy, 1965; 1971; Locke, 1997). We have been consistently wise as a field to consider the evolution and roles of digital technology with deliberation and care. So how will we respond to this new AI-powered digital dimension knocking at our door? Should we let it in, so to speak? Do we have a choice? Or is it already beyond our control? In 1965, Cassidy called for a culture of change and openness to “describe, accept, develop and incorporate” technology into our work (p. 14). This is easier said than done, particularly when the technology is so complex and the future formidable to envision, as is the case with AI-powered LLMs. As Haggerty (1997) noted, predicting the influence technology may have within kinesiology and physical education is, “a

1 In the six weeks between the initial submission of this manuscript and our final submission, ChatGPT-4 passed the Uniform Bar Exam in the 90th percentile (up from the 10th percentile by ChatGPT-3), several countries placed temporary bans on its use due to data privacy concerns and lack of age restrictions or content filters available for minors, and prominent technology entrepreneurs like Elon Musk called for a moratorium on AI advancement due to their perceived risk of its unbridled power to society.
difficult task- not difficult to predict but difficult to be right” (p. 257). Locke (1997) also warned, “believing all those predictions about what technology can do sometimes causes us to do really dumb things” (p. 272).

Correctly forecasting how ChatGPT and other AI-driven LLMs will affect our field, and the work of future professionals, is no easy task. These technologies will have an impact and likely already are. For this reason, we have decided to develop this research note. “We” are a group of faculty members working in Physical Education Teacher Education (PETE) who teach, conduct research, and serve in various roles for academic journals and books (e.g., reviewers, associate editors, editors). We are becoming increasingly cognizant of the effects AI-powered LLMs currently and will have across all facets of our work. We suggest ignoring, avoiding, or being indifferent to predictions about what technology like AI-powered LLMs can do, can also cause us (academics) to do really dumb things. Rather, we encourage all health and physical education teacher education (H-PETE) faculty members to learn about the capabilities, functions, and potential applications of these AI-powered LLM tools. Doing so can facilitate informed, solution-focused approaches to leveraging new tools and technologies, including ChatGPT, across academic workspaces. In line with Knudson (2016), we suggest we all actively develop and adapt our skills and competencies to address the ever-evolving technology landscape with appropriate wisdom. What this call is not intended to do, is to suggest we succumb to all new tools, devices, and technological trends. We hope to encourage scholars to recognize the potential implications AI-powered LLMs bring to our work and consider working together to develop field- and discipline-specific consensus statements to guide ethical and appropriate use.

**Implications of AI-Powered LLMs for Learning, Teaching and Research in H-PETE**
The following section outlines some of the possibilities and challenges associated with the emergence of AI-driven LLMs within and across our academic workspaces. We acknowledge the following considerations are incomplete but hope to stimulate thoughtful interaction with potential impacts these new technologies may have, for better and/or worse.

**Pitfalls and Potentials for Learning in H-PETE**

One concern about student use of AI-powered LLMs is a heavy reliance on them at the expense of fully engaging in learning processes. Students can now decide to skip engaging with assignments by handing in completely AI-generated artifacts. While there are clear and obvious ethical implications related to academic dishonesty using AI-powered tools, existent umbrella policies in many institutions will (hopefully) dissuade most students from submitting work that is not theirs. What becomes more disquieting is the negative impact these tools may have on student learning, along with potential downstream impacts of underprepared professionals entering the workforce upon graduation. Learning how to write has been associated with learning to think, establishing reason, logic, and sequencing of concepts. Learning how to lesson plan in H-PETE is learning to develop and apply critical thinking around effective teaching processes. If a student can prompt ChatGPT to output a lesson plan for a specific skill-theme (and ChatGPT can do this) related to specific grade-level standards, how will they develop their ability to recognize, plan, and enact appropriate instruction? Students may even reconsider why bother building deep content knowledge, when ChatGPT can serve as their resident content expert.

Yet, we ask “how is prompting an AI-powered LLM to develop a lesson plan different from when students consult slowchathealth.com or pecentral.org?” Stepping back from the initial shock of new technologies and discourse surrounding their emergence, we may begin to see potential value of the tools for knowledge construction and idea generation. In the past, students
started their lesson planning with a blank screen, whatever knowledge they had from their own health and physical education (HPE) experiences, and what they were taught in their content-knowledge courses to that point. Now, students begin with these and a resident content expert resource in the form of AI-powered LLMs. Students can now start with something on their screen, and with knowledge and support of H-PETE instructors, can conscientiously build the output into something that aligns with good practice. Such a solution-oriented perspective has the potential to encourage students to leverage the immense power of AI-powered LLMs responsibly, without sacrificing or by-passing aspects of the learning process.

Potentials and Pitfalls for Teaching in H-PETE

University teachers work to acquire specialized knowledge in their fields of interest to become known as authorities or experts who are qualified to teach others. Specialized, discipline-specific information has been produced and curated in a range of print and digital forms, many of which have converted to online/web-based resources for content sharing and dissemination. The emergence of AI-driven LLMs can potentially reposition who or what holds a position of authority on a topic and who or what are considered key stakeholders in brokering new/alternative knowledges. Provided with a well-engineered prompt, ChatGPT has the capability to scour publicly accessible digital content to produce a nuanced response for the user. Instructors could harness this AI function to augment instruction in ways that enhance or update their breadth of knowledge and understanding, which could in turn contribute to developing new (and diverse) methods of organizing, presenting, and communicating specialized information. Capitalizing on ChatGPT in this way could also introduce instructors and students to newly generated knowledge and perhaps offer contemporary perspectives to foundational knowledge. It could also save instructors time, create certain efficiencies, and facilitate easier and more regular
professional learning for re-tooling and content refreshing. These initial ideas offer value-added possibilities and positive potential for how AI-driven LLMs could enhance various teaching functions at this time.

Issues and possible pitfalls require consideration for teachers’ responsible use of this innovation. If AI-driven LLMs hold the potential to synthesize specialized knowledge into consumable, translatable, and transferable segments of information for a user, one might ask: Who and/or what will take a primary role in teaching our teachers? How long will content-driven professional training programs continue to exist in their current form? At what point does the position of subject-matter authority shift from human-generated knowledge to machine-generated knowledge? And could AI-driven LLMs allow individuals to position themselves as subject-matter experts by using ChatGPT, without proper credentials or a certification accompanied by accountability measures/mechanisms?

The use of LLMs in teaching specialized HPE/H-PETE knowledge still requires critical consideration. Not all specialized knowledge is publicly accessible, potentially biasing the knowledge base toward information available through open-access digital platforms or providing misinformation to users. Knowledge possession and consumption do not translate without applying sound pedagogical strategies designed for individual learners. Privacy and security related concerns must be raised when machine-generated information is curating summaries and accessing “for profit” databases. Could the regular use of tools such as ChatGPT foster a sense of dependence on AI-produced content and limit teachers’ natural curiosity, sense of creativity, contextual awareness, and suppress active and responsive problem-solving strategies needed within a teaching and learning environment?

**Potentials and Pitfalls for Research in H-PETE**
Overreliance on AI-powered LLMs by researchers is also a concern, with perhaps higher stakes given the nature of publication and scientific communication. Manuscripts written exclusively by AI-powered LLMs could be circulating without our knowledge or any current, public efforts to curtail by publishing houses. Yet, these tools cannot be authors on publications since it cannot be held accountable for disseminated content or generation of ideas (Nature, 2023). Alternatively, if/when researchers rely heavily on these tools at the expense of engaging fully in research processes, there is a chance for biased or misinformed results to get published. This problem speaks to a key limitation of AI-powered LLMs: Outputs are restricted by the available and/or accessible digital data. If an AI-powered LLM bases its output only on research published through open-access or if research in a particular topic area is still emerging or underdeveloped, the output will inherently be bound by the (limited) digital information available. Furthermore, if an answer is unclear or unknown the AI-powered LLM will generate a response that is based on the predictive modeling or anticipated logic, not on fact, a considerable concern for consumers of published works. The relative novelty of these technologies means there are reliability issues related to the accuracy of certain outputs, emphasizing human critical engagement with these tools is still essential (i.e., van Dis et al., 2023).

AI-powered LLMs may provide powerful support to aspects of the research process. In fact, our writing is already influenced by them, often without notice. Spellcheck, grammar-checks, and auto-complete are all LLMs we already use. When does a researcher cross the ethical AI-powered LLM use line? At the time of writing, these tools have the capacity to generate fully developed, technical manuscripts. Despite the existence of programs designed to distinguish AI-generated text from human-developed writing, they too, are unreliable for use in peer-review publishing process at this time. Claiming authorship on a paper generated by an AI-
powered LLM is unethical, fraudulent, and potentially dangerous to the public. But what about using it to develop a manuscript introduction or to output a 200-word abstract for a book chapter? And if we can agree that some use is helpful, is it necessary to notify readers? And if we deem that it is, what is the appropriate way of acknowledging its use?

These questions only address tasks AI-driven LLMs can conduct. Other research-related tasks general AI can perform are well beyond generating human-like text, like creating art, sequencing genes, and predicting social outcomes, which have broader implication for learning, teaching, and research enterprises. van Dis and colleagues (2023) developed a helpful list of questions to facilitate discussion and refinement of how scholars might approach the use of AI-powered LLMs, and AI more generally, within the research enterprise. Some of these questions include: Which research tasks should or should not be outsourced to LLMs? Which academic skills and characteristics remain essential to researchers? How should research integrity policies be changed to address LLMs? How should LLMs be incorporated into the education and training of researchers?

**Ethical considerations for using AI-powered LLMs in research and teaching**

Earlier we discussed some potentials and pitfalls of using AI-driven LLMs in research and teaching. We want to also highlight potentially consequential ethical issues surrounding the use of such technologies.

One ethical concern is how these tools access data without consent, which (at the very least) poses the risk that researchers do not receiving adequate credit for their work. There is risk that scholars using these tools could present what they consider “original” thoughts (generated from an AI output) and pass them off as their own in subsequent publications, given the relevance and accuracy of output references are currently unreliable. In this scenario, not only
would the ideas of original authors be disseminated without due credit, but there would also be
no citation to document the meaningful work of the original scholar. This in turn would affect
important metrics considered for tenure and promotion at research intensive institutions (i.e.,
number of citations and H-Index).

Most literature published in major journals is currently behind paywalls. If an AI-driven
LLM is trained only on freely available data (and we will note that OpenAI has not been
transparent about what ChatGPT is trained on), it will skew output contents exclusively toward
open access articles. This favors information produced by scholars and institutions that have
funds paying publishing fees. What will the landscape of research in our field look like when the
AI-generated outputs are biased toward a small sliver of the actual research in existence? Might
this put pressure on more journals to become open access and/or at what cost?

**Taking a Socio-Critical Perspective**

Considering the place of digital technologies within the broad spectrum of (health and
physical) education, there is a tendency to focus on the closest environments within our social
and work ecology- the university or K-12 school and the potential for different technologies to
effectively support research and the teaching and learning of young people. Our focus on such
aspects can result in a blinkered view of how digital technologies can create complex social
inequalities and inequity in educational environments which also mirror those experienced in
human communications in our society (Mascheroni & Ólafsson, 2016; Chen et al., 2023). As
digital technologies develop and evolve into profit-driven platforms, free use of high-quality
programs will be more restricted (as evidenced by the move of ChatGPT-4 to a pay-per-use
model). This may result in disparities in student access, support, use, and proficiencies, along
with any number of downstream effects (Centeio, 2017). We emphasize maintaining a socio-
critical perspective when considering if and the extent to which these tools might be useful in our work and within society more broadly. This involves being mindful of the social phenomenon inherent in digital technology use and critically investigating social environments in which they are used (i.e., the home, or online; Selwyn, 2016). Taking a socio-critical perspective can provide a lens through which digital (in)equality issues can be understood and linked to broader social inequities, while also making it possible to consider alternatives to ensure the use of such tools contribute to promoting equitable practices (e.g., supporting students who are neurodivergent).

In addition, the assumption that all socio-economic groups of educators and learners can access AI-driven LLMs in the first place is unwarranted. (In)accessibility becomes problematic when these types of digital resources are more readily available in certain privileged spaces. Just as we lean on our current technology, knowledge, and expertise to understand new technological developments in our own institutions, ensuring every HPE/H-PETE instructor and teacher and their programs and schools are appropriately equipped and resourced to utilize this approach cannot be taken for granted. There are already some people being left behind in this regard.

A Call for Consensus on the Appropriate and Ethical Use of AI-Powered LLMs

Over 25-years ago Locke (1997) cautioned us to “remember that a technological train may never leave the station and, even if it does, it may go somewhere you would really not want to be” (p. 273). A horse-drawn wagon might also have left the same station long before the train ever arrived. Ranging caution around AI-Powered LLMs is warranted and encouraged. This is not the first-time technological development has knocked at the door of our field, and it certainly will not be the last. Just as Cassidy (1965; 1971) wrote about the advances of technology among other changes to the world during her era - this is our new dimension. These and other issues (many of which are not yet foreseen), will be critical for H-PETE faculty to attend to as we
navigate a new world that now includes AI-powered tools. There are currently more questions about appropriate and ethical AI-powered LLM use than answers. While scholars across various disciplines have made general recommendations, appropriate and ethical application of AI in learning, teaching, and research spaces has not been articulated by any field.

Further, consideration is needed regarding how teachers and scholars are trained to approach technology advancements and AI-driven tools like ChatGPT. As we consider if, how, and what that approach may be within HPE/H-PETE, we acknowledge critical, empathetic digital citizens must examine the “important ethical questions at the intersection of technology and humanity” (iste.org, para.1). Conversations about potentials and pitfalls of AI-driven LLMs would be incomplete without acknowledging the individual users of these tools. We find ourselves at such an impasse: at the intersection of AI-driven LLMs and humanity. We are compelled to prompt this intellectual discourse around intentional consequences, and integrity of use, protection of vulnerable populations from inappropriate use, equitable use, and access. Considering the user reminds us of the importance and value of diversity of thought, philosophy, perspective, experience, pedagogical approaches, and instructional styles. We acknowledge and embrace thoughtful rationales for use and non-use of AI-driven LLMs in H-PETE, with an important caveat: That we, collectively commit not to allow these differences to divide but to unify and embrace this challenge as an opportunity to understand one another better as we engage in thoughtful discourse related to opportunities and risks on the horizon.

This research note represents a call to scholars in health, physical education, and kinesiology more broadly to consider the current technological landscape and ask how we can be intentional about addressing the proliferation of AI-powered LLMs. We ask whether a collective consensus statement or set of guidelines for appropriate and ethical applications across academic
workspaces would be a beneficial or necessary endeavor at this stage? We believe the security of
our integrity into the future, as well as the maintenance of public trust in the teaching, learning
and research we and our students conduct, depends on our ability to confront the present and
ever-changing technological landscape with collective good intention and action. However, that
statement reflects the voice and beliefs of this author group. What is now needed is your voice,
your perspective, and your input. All HPE/H-PETE stakeholder voices are valuable in
broadening the understanding of AI-powered LLMs in our teaching, learning, and research
spaces. Active consensus building requires a clear definition of the issue. We hope this research
note is only a launching point and is followed by open forums, critical formal and informal
conversations, conference papers, technical reports, and more. These efforts will shape a clearer
picture of the dimensions and nuisances of LLMs in HPE/H-PETE and be significant in shaping
our collective understanding of the issue. We believe that a proactive approach to consensus
building around an initial set of discipline-specific guidelines for appropriate and ethical
applications within our discipline-specific workspaces will help current and future HPE/H-PETE
professionals responsibly navigate this new dimension we find ourselves entering.
318 References
328 Nature. (2023). Tools such as ChatGPT threaten transparent science; Here are our ground rules for their use. [Editorial]. Nature, 613, 612. https://doi.org/10.1038/d41586-023-00191-1
and what might be done about it”, Learning, Media and Technology, 41 (3), pp. 437-441.

This research note by a group of Physical Education Teacher Education (PETE) faculty members argues that the emergence of AI-powered chatbots like ChatGPT pose challenges to the future of scholarship, and that we need to pay attention to the added issues and opportunities associated with this technology within learning, teaching, and research spaces. The note suggests that ignoring or being indifferent to predictions about what technology like AI-powered chatbots can do can cause academics to do "dumb things," and that all H-PETE faculty members should learn about the capabilities, functions, and potential applications of AI-powered LLM tools to facilitate informed, solution-focused strategies and approaches to leveraging new tools and technologies across academic workspaces. The authors emphasize the need for field- and discipline-specific consensus statements to guide ethical and appropriate use of AI-powered chatbots in scholarship, to ensure that the value of these tools is harnessed for the good of the society. The authors also highlight the importance of maintaining a socio-critical perspective when considering the use of digital technologies, to understand and address potential digital (in)equity issues and to promote equitable practices.