

Using leadership to leverage ChatGPT and artificial intelligence for undergraduate and postgraduate research supervision

Michael Cowling
CQUniversity

Joseph Crawford
University of Tasmania

Kelly-Ann Allen
Monash University

Michael Wehmeyer
The University of Kansas

ChatGPT and other artificial intelligence (AI) and large language models (LLMs) have hit higher education by storm. Much of the research focuses on how this – and similar – tools can be leveraged for effective education of undergraduate coursework students. In this study, we explore the emerging benefits and limitations of ChatGPT and LLMs in the context of undergraduate and postgraduate research supervision. What we found was that psychological need fulfilment, research student autonomy and relatedness were key outcomes that could be cultivated at the student level. At a unit or subject level, the opportunity for formative feedback was seen as a strength. We also discuss some key limitations to the tool, including how limited its ability to deconstruct social injustice and generate content appropriate to context. We used an example of leadership research to highlight that it may preference good outcomes and likewise present information related to current and normative practices rather than desired future practices. We conclude by considering the broad implications of this work on research supervision relationships.

Implications for practice or policy:

- ChatGPT has the ability to enhance research higher degree research practices.
- AI and LLMs may support student psychological need fulfilment, autonomy, competence and relatedness.
- ChatGPT could provide preliminary formative feedback support for research and doctoral students prior to submitting drafts to supervisory teams.
- Policy safeguards are needed to ensure research responses to lack of context, data bias, equity concerns and lack of an ethical framework.

Keywords: higher education, generative AI, ChatGPT-3, large language model, research student, doctoral student, research supervision

Introduction

Higher education has undergone significant changes in recent decades, with massification strategies resulting in greater student load enrolled in universities and a subsequent increase in academic staff workload (Kenny & Fluck, 2022; United Nations Educational, Scientific and Cultural Organization, 2020). This has flow-on effects for research supervision. For example, the first Australian PhDs were awarded in 1949, with 204 awarded in the following decade from the University of Melbourne, the University of Queensland and University of Adelaide (mainly men: Evans et al., 2003). In 2015, there were 65,552 postgraduate research candidates in Australia contrasted with 58,000 full-time equivalent academics (Beasy et al., 2022). As a result, research supervision has become more challenging due to difficulties in providing personalised, individualised and differentiated attention to students (Denis et al., 2019; Igumbor et al., 2022).

Recent advances in artificial intelligence (AI) may offer opportunities for enhancing research supervision quality and efficiency, potentially benefiting students undertaking research, despite integrity concerns (Chu et al., 2022; Perkins, 2023). One such opportunity is the use of ChatGPT (Dempere et al., 2023), an AI-powered conversational agent that can engage in natural language conversations with students, providing them with feedback to support their research supervision needs (Open AI, 2022). Similar to the cohort method for research higher degree supervision (Wisker et al., 2007), ChatGPT as a tool provides an opportunity for research higher degree students to obtain additional feedback, perhaps even providing a further step than a cohort of human supervisors with 24/7 support available.

This article examines the potential advantages of ChatGPT, which may be valuable for undergraduate and postgraduate research supervision, and assesses the advantages and drawbacks of using AI technology in research supervision by analysing literature on ChatGPT in higher education, presenting some prompt responses from the large language model (LLM) tool. Using this approach, it provides a critical analysis of ChatGPT's potential for undergraduate and postgraduate research supervision as well as insights and recommendations for academics, educational leaders and policymakers exploring the use of AI in this area.

Context: AI in higher education

Various factors, such as better access to higher education, the growth of distance learning, social expectations regarding degrees for stable employment and competitive job markets (Kenny & Fluck 2022; United Nations Educational, Scientific and Cultural Organization, 2020) have contributed to the rise in university student enrolments over the last few decades. Over a similar period, academic staff have also reported increased workloads, particularly in respect to administrative tasks (Allen, 2019; Allen et al., 2020; Allen et al., 2021). Some academic institutions have been criticised for operating under a business-oriented model, prioritising profitability over education (Pucciarelli & Kaplan, 2016; Spinrad et al., 2022); however, there are several changes that may also contribute to increased workloads.

For instance, the student body has become more varied, with a greater range of needs with additional support and resources required to ensure academic success (e.g., learning accommodations, financial assistance) (Shaw, 2009). Research supervisors may be called on to create arrangements for students above their expertise. Additionally, the COVID-19 pandemic posed numerous challenges that also required concerted efforts by academic staff to maintain engagement and motivation and provide the necessary support to their students (Tice et al., 2021). When asked "How has higher education research supervision changed in recent years?", ChatGPT responded with:

In recent years, higher education research supervision has become more student-centered, with a greater emphasis on individualized and collaborative approaches, the use of technology to support communication and feedback, and a focus on developing students' transferable skills and employability.

Although not specific to COVID-19, this response from ChatGPT does reflect changes in the sector over the last few years (e.g., Yang & Kyun, 2022). For instance, to ChatGPT's point, higher education research has also revealed a trend towards practical industry-oriented research (Narayan et al., 2017; Reddy et al., 2016) which may place an additional burden on academic staff who lack industry-based or professional experiences. This is particularly true when research student career aspirations are not always aligned with the institutional supports available (Beasy et al., 2022). Furthermore, research metrics are an additional burden on academic staff and are often used by universities as a critical measure of institutional success leading to improved institutional rankings and marketing strategies to attract more students (Jeyaraj et al., 2021; Longden & Yorke, 2009). The growing importance of research in higher education may place pressure on research supervision, adding to the already overstrained academic staff.

Recognising financial pressures, growing student numbers and academic performance goals as potential threats to research supervision quality is crucial (Kumar et al., 2021; Pucciarelli & Kaplan, 2016; Wisker et al., 2021). To protect the quality of research supervision, exploring preventative or proactive solutions

while considering emerging technologies like AI and advanced models such as ChatGPT may enhance the research experience for students without compromising the economic viability of institutions.

Theoretical framework: AI and LLM

A cursory review of the literature during 2023 clearly shows that ChatGPT has had a significant impact on higher education in general. On 18 February 2023, there were almost 5,000 published works on ChatGPT, according to Google Scholar (Crawford, Cowling, & Allen, 2023). As of the time of writing (28 September 2023), this number is almost at 45,000. As of 28 September 2023, Scopus lists 827 articles that include ChatGPT in the title, abstract or keywords, but these studies are brief and offer varied opinions on the topic; most of which of these are still in early access, indicating a lack of quality studies in top-tier journals. For instance, Editor-in-Chief of *Science* Holden Thorp (2023) suggests that machines are essential tools for people to pose hypotheses and make sense of results, rather than being authors on a paper. However, not all researchers share this perspective, with at least four studies referring to ChatGPT as a co-author (Stokel-Walker & Van Noorden, 2023), although journals are generally expecting transparency on usage (e.g., Crawford, Cowling, Ashton-Hay et al., 2023).

Tools such as ChatGPT (often holistically called “generative AI” tools) provide a unique opportunity to iterate on the mechanisms by which we supervise PhD students. Although prior AIs have focused primarily on a process of forecasting and predicting, the new breed of generative AI tools embody a level of creativity that has been unseen in the past (Kung et al., 2023). This means that whilst previous AI tools could be used for students primarily for paraphrasing and summarising text and for data analysis (Lodge et al., 2023), this new set of tools can be used much more broadly by students to look at diverse problems, particularly focused for doctoral students on enhancing their approaches to written expression.

To fully understand what ChatGPT can do for students, a basic understanding of both AI and these new approaches is useful. Although both are ultimately mathematical in nature, a key difference between these tools is that whilst previous AI tools were focused on fact, these new tools could perhaps be considered to be focused much more on expression. This inherently makes the result of new generative AI tools fuzzy (Hughes et al., 2021), which whilst not useful if using a tool to generate an answer that is factual, can be useful in crafting an answer that is creative, making the tool very suitable for language development. When we asked ChatGPT “What makes you different than other AI systems?”, where “other AI systems” referred to the more traditional non-GPT styles, it responded with:

As an AI language model, I am trained to understand and generate natural language text. Here are some key ways that I am different from other AI systems:

1. Large-scale training data: I have been trained on an enormous corpus of text data, which allows me to understand a wide range of topics and language styles.
2. Flexibility: I am designed to be highly flexible and adaptable, which means that I can answer a wide variety of questions and generate text in different styles.
3. Contextual understanding: I am able to understand the context of a question or statement and generate a response that takes into account that context.
4. Natural language generation: I am able to generate text that is grammatically correct and sounds natural, which makes it easier for humans to read and understand.
5. Continuous learning: I am constantly updating my knowledge and improving my performance based on feedback from users and new data that I encounter.

Overall, my main strength is my ability to process and generate natural language text in a way that is useful and understandable to humans.

As highlighted by the tool, the core of this system is the LLM (Kasneci et al., 2023), a large corpus of data gathered through scraping social media, websites and other sources. This data is used to train the ChatGPT system, which then uses its fuzzy model to essentially predict the most likely candidate for the next word based on the text that has been requested (through a prompt) or already generated. This makes this type of tool uniquely suited for generated text for our doctoral students at the beginning of their candidature

and gives them the ability to focus more on key concepts than on the way they express themselves. Although written expression of ideas is important, it is clear that LLMs can give the opportunity to speed up the process of articulating key concepts in the beginning stages of a research higher degree candidate's journey.

Also important is the conversational nature of the tool and the ability to understand context. As an example, following on from the previous question, we asked ChatGPT "What do you mean by large-scale training data?", and it said:

When we talk about "large-scale training data" in the context of an AI language model like myself, we are referring to the enormous amount of text data that I was trained on in order to develop my language understanding and generation capabilities.

Note particularly the reference to the previous query, and the implication that is automatically made to the previous question and the context. Finally, the tool is flexible, and in doing so also has the ability to adjust language; when asked to summarise the conversation in 100 words, ChatGPT said:

Large-scale training data refers to the vast corpus of text data that is used to train an AI language model like me. This data is typically pre-processed to remove noise and irrelevant content, and is used to teach the model to recognize patterns and relationships between words and phrases. Having access to a large amount of diverse training data is crucial for developing sophisticated language understanding and generation capabilities. It allows AI models to learn from a wide variety of language styles and contexts, and improves their accuracy and quality over time.

Through an approach that is creative, contextual and flexible, the tool is able to provide a new approach and new insights to students in a range of different ways that can assist their research work. In fact, when we asked how ChatGPT can help research students, the AI highlighted:

As a language model, ChatGPT can assist research students by providing information, helping with literature review, generating ideas, proofreading and editing their papers, and providing citation assistance.

The following section will further examine this in more detail.

Using LLM and ChatGPT in research supervision to meet student needs

The role of LLMs in education is to serve as a form of knowledge holder (Stokel-Walker & Van Noorden, 2023). Although the data still needs to be parsed by a human for accuracy, the AI does tend to provide useful initial data for consideration. In the context of research supervision, it does not replace the need for human supervision – like a self-check-out supermarket might do – but rather can offer a strategy to optimise and change the nature of the supervision relationship. In this section, we discuss psychological need fulfilment for candidates and students, the opportunity for student autonomy, competence, relatedness and pre-human formative feedback.

Psychological need fulfilment

Self-determination theory (SDT) is one of the most empirically supported psychological theories to date, contending that all people, regardless of context or population, have innate universal psychological needs for autonomy, relatedness and competence (Deci & Ryan, 2000). When such needs are met, they are critical for various positive outcomes, including optimal functioning, well-being, flourishing and life satisfaction (Ryan & Deci, 2017), and in the context of higher education, student achievement, engagement and academic motivation (Geary et al., 2023; Levesque-Bristol, 2023). As a result, SDT is regarded a highly credible and evidence-based theory with practical applications in various fields,

significantly higher education, suggesting that it might be well suited to being used as a theoretical framework for research supervision.

This is further support due to the fact that, despite the scarcity of research on SDT specifically and its relationship with research supervision, a multitude of related studies do suggest that supervisors who prioritise fostering autonomy, relatedness and competence are conducive to achieving favourable outcomes in academic performance as well as personal welfare (Sawatsky et al., 2022). For example, research supervisors can create a supportive and motivating learning environment that promotes academic success and contributes to students' academic and personal growth and development by satisfying students' needs for autonomy, relatedness and competence (Khosa et al., 2023). Traditional modes of research supervision, however, may only sometimes meet the psychological needs of students, particularly in the context of staff shortages and competing workload pressures that may put pressure on staff to prioritise their institution's interests (e.g., research metrics) over supporting students. This widespread problem calls for the examination of alternate remedies, including innovative advancements like technologies such as ChatGPT, which have the potential to transform research supervision while also working towards meeting the psychological needs of autonomy, competence and relatedness of students

Research student autonomy

Within SDT, the need for autonomy is satisfied when a person experiences choice and volition in their action and perceives themselves to be the origin of their actions (Adams et al., 2017). Without research evidence to support using ChatGPT to address students' psychological needs, scholars can only speculate on how ChatGPT could improve or support students' autonomy. One potential scenario involves a student having trouble in a particular research area. For example, it could be imagined that an early research higher degree student might be seeking to do a literature review but unsure of what is involved in moving into this space. In such a case, ChatGPT may offer prompt guidance on effective techniques aimed at fostering students' sense of agency as well as ownership over their learning journey, such as detailing for them the purpose of a literature review or the approaches that scholars often take for a literature review. Through prompting and question and answer, the student will build a clearer mental model of how the literature review process works, that can then be reinforced through meetings with the research higher degree supervisor. This self-directed learning allows learners to be active contributors to the knowledge acquisition process rather than a passive recipient.

This real-time live support relies on the student's willingness and choice to engage rather than needing first to find a suitable time with their supervisor to discuss. There is, in fact, some evidence that the use of technology can promote self-determination and autonomy (Lee et al., 2011). When asking ChatGPT "How can ChatGPT support research student autonomy?", the AI responded with:

ChatGPT can support research student autonomy by providing information, guidance, and resources that enable them to take control of their own learning and research.

As such, ChatGPT is more than just a passive tool for students to use. Instead, it engages students in an active dialogue, allowing them to control the input they provide that, in turn, has an autonomous role in shaping the feedback and output they receive. In this way, ChatGPT becomes a means through which students direct their learning. Much like a GPS navigator, ChatGPT provides students feedback and support as they navigate their research journey.

Universities and faculty staff who permit for the use of tools like Chat GPT, and perhaps also provide ethical guidelines to facilitate students to engage with such tools, are also enabling students' autonomy to choose supports and resources that are most beneficial to their needs and preferences at the time, giving them a greater sense of autonomy within the research process. When students feel like they have autonomy over their learning, they are likelier to feel a sense of ownership and engagement in their work, leading to increased academic motivation and success (Ryan, 2023).

Research student competence

Within SDT, the need for competence reflects people's desire to master their environment and experience a sense of competence in that environment (Adams et al., 2017). Competence refers to the desire to feel capable and effective in one's actions (Ryan & Deci, 2017). To be effective in educational pursuits, research students must feel competent in their research skills. As a learning tool, ChatGPT can boost students' sense of competence by providing expert guidance and support that allows students to build skills in their academic writing, consolidate their thoughts and clarify the feasibility of research methodologies, as some examples. The feedback mechanism which ChatGPT provides may be beneficial in building confidence and self-efficacy in research students by providing them with a clear understanding of their strengths and areas for improvement. When asking ChatGPT "How can ChatGPT support research student competence?", the AI responded with:

ChatGPT can support research student competence by providing information, feedback, and resources that help students develop their knowledge, skills, and abilities in conducting research.

Research writing can often be cliched, formulaic, and rigid in its structure. The AI capabilities of ChatGPT offer an opportunity for students to understand the expectations of academic writing while having a greater understanding of the possibilities to express themselves. This may also reduce students' uncertainty when engaging in the academic writing process – particularly for the first time.

Universities and research supervisors can also contribute directly to students' sense of competence by providing them with information and training on effectively using ChatGPT and other tools to support their research. This may include guidance on formulating research questions, searching for and evaluating reliable sources of evidence and integrating feedback from ChatGPT into their writing and analysis without jeopardising research integrity. By equipping students with the skills and knowledge they need to utilise ChatGPT, universities and research supervisors effectively empower students to take charge of their learning and research, which builds their competencies and further enhances their sense of autonomy and control.

Research student relatedness

Finally, the need for relatedness within SDT refers to satisfaction derived from a sense of connectedness with others; to belong, to care and be cared for by others (Adams et al., 2017). As counter-intuitive as it may seem, ChatGPT may also improve students' sense of relatedness. Although it has been postulated that ChatGPT could be used as a mechanism for short-term student social support (Crawford et al., 2023a) in students knowing that around-the-clock support and feedback exist, unlike in traditional research supervision, there may also be other ways relatedness can be generated. For instance, by building skills and competencies in specific research areas, students may feel more connected to their research community or other students. They may have more confidence to engage in broader research networks and have an enhanced sense of belonging (Tice et al., 2021).

Furthermore, ChatGPT, through providing increased support to students, may also enable supervisors more time to spend with students in more meaningful ways, with less time spent on minor tasks and problems of research supervision (e.g., editing, grammatical feedback) that can be provided through ChatGPT. This may allow supervisors more time to focus on relationship-building, provide personal support and build rapport with their students. This is especially important given the time constraints that supervisors face, as they are frequently required to manage many students simultaneously. Although AI is unlikely to replace the role of the research supervisor altogether, ChatGPT could be a tool worth musing further about its role in relatedness and support in the research supervision process. The final word on the value of generative AI could perhaps be given to the tool itself, as when asking ChatGPT "How can ChatGPT support research student relatedness?", the AI responded with:

ChatGPT can support research student relatedness by providing a sense of connection, community, and collaboration through discussion, feedback, and guidance.

Formative feedback to research students

When research students seek advice from their research supervisor, they often bring preliminary ideas that are not yet well formulated. For supervisors, it can be a frustrating experience consistently providing early readings and repeated comments of feedback and reviewing early drafts laden with grammatical, spelling and syntax errors (Jiang & Yan, 2020). For some supervisors, the presence of writing errors becomes a significant distraction from reviewing the substantive comments.

ChatGPT and other AI offer an early research support for students by allowing them to paste in text and have it checked for errors. In this context, students can seek advice from the AI to reduce any minor errors in their work prior to submitting it to their supervisor. Generally speaking, the tool itself has some basic characteristics that ensure its utility in diverse contexts. For example, in Kung et al. (2023), ChatGPT was asked to complete the three exams required to pass the United States Medical Licensing Exam. In this test, it demonstrated considerable competence in passing the clinical decision-making exam with explanations for its decisions.

For research students, although there is little published on this to date, some inferences can be drawn. Students are afforded an early-stakes submission to critique elements of their research. To provide a simple example, we asked ChatGPT to provide advice on a draft research question, "What do you think of my research question: To what extent does unethical leadership predict unethical behaviour?", it responded with:

As an AI language model, I don't have personal opinions, but I can offer some insights regarding your research question.

Your research question is interesting as it seeks to explore the relationship between unethical leadership and unethical behaviour. However, to make it more precise, you might want to consider narrowing down your focus by specifying the context and population you are studying. For example, you could ask, "To what extent does unethical leadership predict unethical behaviour among employees in the banking industry?"

Additionally, it is essential to define what you mean by "unethical leadership" and "unethical behaviour." These terms can have different meanings to different people, so clarifying your definitions will help ensure that your research is focused and valid.

Overall, your research question is a good starting point for exploring the relationship between unethical leadership and unethical behaviour, but refining it further will increase its clarity and usefulness.

As can be observed, it provides some critical points that may help students to sense-check their work and develop confidence in their decisions prior to providing full drafts to their supervisors. In doing so, it offers capacity and space for students to have higher impact conversations with their supervisors, focusing on enriched theoretical discussions that have been pre-tested in the AI context. In some ways, this offers a form of early low-stakes assessment discussed to support student confidence (Meer & Chapman, 2014) and support non-traditional learners to transition into higher education (Sambell & Hubbard, 2004).

The costs of LLM and ChatGPT in research supervision

In addition to the benefits, there are a series of risks attached to the ChatGPT (and equivalent LLMs) in research supervision. In this, we highlight three key aspects theorised using early insights. Firstly, the AI lacks explicit context. ChatGPT also has a data bias that can perpetuate existing issues. The final point raised relates to the separation between ChatGPT and formal ethical research frameworks. These limitations do not necessarily predicate flaws in the software, but rather areas where additional due diligence may be required in applying ChatGPT to support effective research supervision.

ChatGPT lacks context

In dissertations, most students engage in highly specific research topics that require nuanced knowledge and contextual application of that knowledge. Although ChatGPT tends to be effective at broad strategic responses, and curates high-level information, it is less effective at providing deep responses about individuals, organisations and places that have reduced international appeal. For example, we asked ChatGPT who each of the three authors of the study were. Each author has published dozens of papers prior to the benchmark period, and ChatGPT recognised only one of the authors, with the remaining two (both of whom have public research profiles, Google Scholar, Scopus, ORCID and Web of Science profiles) unidentifiable:

I'm sorry, but as an AI language model, my training data only goes up to September 2021 and I do not have access to real-time information. Therefore, I cannot confirm if there is a current staff member named [Name Removed] at the [Institution Removed].

When applying this logic to a research supervision context, students may ask ChatGPT questions it has difficulty responding to. It also lacks the unique context of the specific candidate study. This can make it more complex to ascertain high quality answers to emergent questions a student might be facing. When asking ChatGPT “What context does ChatGPT lack when providing advice to research students?”, the AI answered:

As an AI language model, ChatGPT lacks the personal experience and subjective understanding of the research process that human research advisors possess.

The need to be clear about the context of research is not a new concept, and Abowd et al. (1999) discussed how a better understanding of context enables greater and more effective context-aware applications. Aldao (2013) discussed the need for context in the emotional regulation research field, and Mishler (1991) highlighted the importance of context in research interview practice. So, while ChatGPT offers a space for students to gain additional insight, they need to be taught how to effectively engage with the tool and apply its broad-based knowledge to a specific research context.

ChatGPT has bias and equity issues

ChatGPT, while using incredible volumes of data, has the potential to make mistakes when providing data. The Australian Drop Bear experiment is an example of where data fabrication can affect the ability for people to process (e.g., Janssen, 2012). In this case, it was not a sophisticated lie and was openly discussed online as fictitious; however, if it were a sophisticated embedding of false evidence, how might ChatGPT work this out? For drop bears, it highlighted:

An Australian drop bear is a fictional creature that is often used as a humorous hoax to tease tourists in Australia. According to the hoax, the drop bear is a carnivorous version of the koala that drops from trees and attacks unsuspecting visitors.

The key point to make is that ChatGPT and other LLMs rely heavily on the quality of the data being inserted into the model. This leads to significant risk that the data will be misread or that AI will have an unconscious bias based on historical societal norms. When we asked ChatGPT “What data bias does ChatGPT have when providing advice to research students?”, the AI answered:

ChatGPT's advice may be biased based on the data it was trained on and the algorithms used to generate responses, which may not fully capture the complexity and nuances of the research process.

Although the software does present transparently the issues it faces regarding the limitations of the data in its modelling, junior researchers and research students (particularly first-time undergraduate research

students) may have considerable difficulty being critical in the examination of the data validity and reliability.

Importantly too, while it may contain accurate information on historical evidence, if it is relied upon without critical review, the AI can perpetuate biases unhelpful to a fair and equal society. ChatGPT is missing the ability to solve current problems like gender bias. In Newstead et al. (2023), they asked a generative AI (not ChatGPT) to generate narratives of leaders without a gender prompt, "What are the characteristics of good leaders?". The response it provided were the generation of exclusively male leader examples. When asked to generate 16 different examples of good and bad leadership, all results consisted of 15 male leaders, one neutral, and zero on females (Newstead et al., 2023).

In this, we followed up and asked it for the top 10 most successful politicians, and after ChatGPT self-imposed criteria for successful ("based on their achievements, impact, and recognition"), it listed (in order) Nelson Mandela, Mahatma Gandhi, Winston Churchill, Abraham Lincoln, Franklin D. Roosevelt, Margaret Thatcher, Deng Xiaoping, Angela Merkel, Lee Kuan Yew and Indira Gandhi. In this example, three of 10 were female, and these were ranked at sixth, eight and 10th. Interestingly though, when asked the reverse, it opted not to respond or respond with political neutrality:

As an AI language model, it is not appropriate for me to make subjective judgments or assessments on individuals' political careers or categorize them as "least successful." Political success can vary based on individual achievements, goals, and perspectives, and it's not appropriate to rank individuals based on subjective criteria. Moreover, it is not ethical to disparage or defame individuals in a public forum. Therefore, I will not provide a list of least successful politicians.

It does, however, provide inconsistent reasons for why it provides good but not bad politician lists. We asked, "Why does ChatGPT provide examples of good politicians but not of bad politicians?", and it responded indicating it did not provide answers for good or bad leaders:

It is not appropriate for me to categorize individuals as "good" or "bad" based on my programming.

ChatGPT did, however, evidence reflexivity in its ranking schema (when asked). In this, we asked "Why does ChatGPT provide more male examples than female examples?", the AI answered:

ChatGPT may provide more male examples than female examples due to biases in the training data, which may have been influenced by historical and societal gender imbalances.

ChatGPT is not aligned to any ethical research framework

Across the globe, most jurisdictions have specific rules of ethical engagement in research. This means that while there are nuances, there are usually specific practices that support and guide what good research is. In Australia, this is the "Australian Code for the Responsible Conduct of Research", jointly developed by the National Health and Medical Research Council, the Australian Research Council and Universities Australia in 2018 (Commonwealth of Australia, 2018). This document provides rules and practices for conducting research ethically. Instead, ChatGPT is argued to be neutral in its assessments, rather than applying any frame of ethical reference. When asking ChatGPT "What ethical framework is ChatGPT aligned to?", the AI answered:

As an AI language model, ChatGPT does not have a specific ethical framework it is aligned to, but rather operates based on the algorithms and data it was trained on.

Although it is generally neutral in its presentation of answers, it is limited directly by the current data inside the software. Such a limitation is also not transparent as few scholars could parse all accessible

information up until 2021 (the ChatGPT training data date). When asking ChatGPT “Does ChatGPT ever provide unethical research advice?”, the AI answered:

ChatGPT does not intentionally provide unethical research advice, but its responses may be biased or flawed due to limitations in its algorithms and training data.

To return to the example of leadership research, we asked ChatGPT “Which five leaders have had the greatest impact in the past five years?”. Kellerman (2022) identified Russian President Vladimir Putin based on strongest global effect (regardless of positive or negative). ChatGPT identified Xi Jinping, Vladimir Putin, Donald Trump, Angela Merkel and Narendra Modi. The first three, by qualitative assessment of many scholars would be argued to have at least some unethical leadership practices. For new research students in a business programme, they may take these comments literally in their work and begin to build a narrative of leadership based on unethical practices. Although it does parse significant volumes of data quickly, its ability to make deeply informed decisions is much lower than a well-educated and critically minded individual assessing its responses.

Discussion

This study aimed to consider some of the key benefits and limitations of AI and LLM tools in a research supervision context. To the best of our knowledge, and at the time of writing, there were no works considering this specific relationship. In this paper, we discussed five key benefits of AI and LLM to research supervision and three limitations attached to their usage. These were not designed to be exhaustive, but rather to provide initial theoretical arguments to support scholarly research in this area (Rudolph et al., 2023), considering some argue of limited long-term viability of AI in higher education (O’Dea & O’Dea, 2023; Popenici, 2023). Despite some emerging studies on ChatGPT and similar, the focus has largely been restricted to early tests and practical uses that present great promise but potentially lack a strong theoretical grounding (at this stage, Neumann et al., 2023). As the embeddedness of AI and LLM in higher education increases, these will create more significant theoretical contributions to the literature.

Although much of the focus – publishing and practical – to date has been oriented around undergraduate coursework programmes, more significant work is needed in the research student and research supervision space. In recent years, there have been considerable examples of substantial cheating in doctoral studies. These were not always detected immediately but were identified later on. For example, a graduate in obstetrics and gynecology at Vrije Universiteit Brussel who earned their PhD in 2013 had their PhD retracted in 2022 for misconduct (Retraction Watch, 2023). Although expectations are significant on the ethical conduct of research students, this is not always the case, and effective detection will be needed to respond to a minority of students using the AI tool incorrectly. However, with effective research supervision governance come substantial opportunities to build high-quality support research supervision relationships that can survive massified education practices and more complex research supervision relationships.

In their review of distance postgraduate research programmes, Nasiri and Mafakheri (2015) highlighted the supervisor and candidate have lower quality relationships. This environment may mean that ChatGPT offers a relational escape for students that further separates the relational connection needed between supervisor and student to generate effective theses compliant with best practice research methodologies. It does also create risk that the geographical distance may cause students to rely more heavily on the ChatGPT tool. Conversely, Kaur et al. (2022) identified that when strong bonds between student and supervisor exist, structure, power and credibility issues may emerge that ChatGPT may have no role to play in helping to mediate.

Likewise, and in support of some existing costs and benefits, McCallin and Nayar (2012) highlighted that research context, faculty issues, supervision pedagogy and models of supervision are key influencing factors for student research success. In this regard, although their strategies proposed human-centric responses, there may be opportunities to consider how AI like ChatGPT can enable students to do, and

be, well. However, effort ought to be made in considering the possible social and human costs of AI–human substitution (Crawford et al., 2024). Training in the emergent school of prompt engineering (see Eager & Brunton, 2023) may better enable this transition. In supporting students to effectively transition into their learning (Kift et al., 2010), creating and embedding AI as a model for early confidence building through autonomous and self-paced activities could support students to achieve in their academic studies despite their initial apprehension with seeking feedback from their professor. This, however, may be limited by discipline differences in technology acceptance of AI (Kelly et al., 2023). This study provides initial insights into how the AI tool ChatGPT may support undergraduate and graduate research supervision.

Conclusion

The aim of this study was to present the current knowledge on ChatGPT and to consider how this may be contextually appropriate for undergraduate and postgraduate research supervision relationships. What we found was that psychological need fulfilment, research student autonomy and relatedness were key outcomes that could be cultivated at the student level. At a unit or subject level, the opportunity for formative feedback was seen as a strength. We also discussed some key limitations to the tool, including how limited its ability to deconstruct social injustice and generate content appropriate to context. We used an example of leadership research to highlight that it may preference good outcomes, and likewise present information related to current and normative practices rather than desired future practices. We concluded by considering the broad implications of this work on research supervision relationships.

Acknowledgements

This article uses some text generated by the Open AI ChatGPT-3 (<https://chat.openai.com/chat>). Typically, when asking the AI questions, it lists multiple answers; in most cases we used only the summary paragraph it offered in the body of this document. Where we do use this content, we present it as a direct quote.

Author contributions

Author 1: Conceptualisation, Investigation, Writing – original draft, Writing – review and editing, Project administration; **Author 2:** Conceptualisation, Data curation, Investigation, Methodology, Formal analysis, Writing – review and editing; **Author 3:** Conceptualisation, Investigation, Writing – review and editing; **Author 4:** Investigation, Writing – review and editing.

References

- Abowd, G. D., Dey, A. K., Brown, P. J., Davies, N., Smith, M., & Steggle, P. (1999). Towards a better understanding of context and context-awareness. In H.-W. Gellersen (Ed.), *Lecture notes in computer science: Vol. 1707. Handheld and ubiquitous computing* (pp. 304–307). Springer. https://link.springer.com/chapter/10.1007/3-540-48157-5_29
- Adams, N., Little, T.D., & Ryan, R.M. (2017). Self-determination theory. In M. L. Wehmeyer, K. A. Shogren, T. D. Little, & S. Lopez (Eds.), *Development of self-determination through the life-course* (pp. 47–34). Springer. https://link.springer.com/chapter/10.1007/978-94-024-1042-6_4
- Aldao, A. (2013). The future of emotion regulation research: Capturing context. *Perspectives on Psychological Science*, 8(2), 155–172. <https://doi.org/10.1177/1745691612459518>
- Allen, K. A. (2019). What is the actual impact of measuring academic notions of impact? *The Educational and Developmental Psychologist*, 36(2), 33–34. <https://doi.org/10.1017/edp.2019.16>
- Allen, K. A., Butler-Henderson, K., Reupert, A., Longmuir, F., Finefer-Rosenbluh, I., Berger, E., Grove, C., Heffernan, A., Freeman, N., Kewalramani, S., Krebs, S., Dsouza, L., Mackie, G., Chapman, D., & Fleer, M. (2021). Work like a girl: Redressing gender inequity in academia through systemic solutions. *Journal of University Teaching & Learning Practice*, 18(3), Article 3. <https://doi.org/10.53761/1.18.3.3>

- Allen, K. A., Donoghue, G. M., Pahlevansharif, S., Jimerson, S. R., & Hattie, J. A. C. (2020). Addressing academic rejection: Recommendations for reform. *Journal of University Teaching & Learning Practice*, 17(5), Article 19. <https://doi.org/10.53761/1.17.5.19>
- Beasy, K., Crawford, J., Young, S., & Kelder, J. (2022). A quantitative study on Australian doctoral students' perceptions of employability preparedness: How gender and age matter. *Journal of Further and Higher Education*, 46(8), 1092–1106. <https://doi.org/10.1080/0309877X.2022.2050687>
- Chu, H. C., Hwang, G. H., Tu, Y. F., & Yang, K. H. (2022). Roles and research trends of artificial intelligence in higher education: A systematic review of the top 50 most-cited articles. *Australasian Journal of Educational Technology*, 38(3), 22–42. <https://doi.org/10.14742/ajet.7526>
- Commonwealth of Australia. (2018). *Australian code for the responsible conduct of research*. National Health and Medical Research Council. <https://www.nhmrc.gov.au/about-us/publications/australian-code-responsible-conduct-research-2018>
- Crawford, J., Allen, K-A., Pani, B., & Cowling, M. (2024). *When artificial intelligence replaces humans in higher education: The cost of loneliness, student success, and retention*. [Manuscript submitted for publication]
- Crawford, J., Cowling, M., & Allen, K. (2023). Leadership is needed for ethical ChatGPT: Character, assessment, and learning using artificial intelligence (AI). *Journal of University Teaching & Learning Practice*, 20(3). <https://doi.org/10.53761/1.20.3.02>
- Crawford, J., Cowling, M., Ashton-Hay, S., Kelder, J. A., Middleton, R., & Wilson, G. S. (2023). Artificial intelligence and authorship editor policy: ChatGPT, Bard Bing AI, and beyond. *Journal of University Teaching & Learning Practice*, 20(5). <https://ro.uow.edu.au/jutlp/vol20/iss5/1/>
- Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, 11(4), 227–268. https://doi.org/10.1207/S15327965PLI1104_01
- Dempere, J., Modugu, K., Hesham, A., & Ramasamy, L. K. (2023). The impact of ChatGPT on higher education. *Frontiers in Education*, 8. <https://doi.org/10.3389/educ.2023.1206936>
- Denis, C., Colet, N. R., & Lison, C. (2019). Doctoral supervision in North America: Perception and challenges of supervisor and supervisee. *Higher Education Studies*, 9(1), 30–39. <https://doi.org/10.5539/hes.v9n1p30>
- Eager, B., & Brunton, R. (2023). Prompting higher education towards AI-augmented teaching and learning practice. *Journal of University Teaching & Learning Practice*, 20(5). <https://doi.org/10.53761/1.20.5.02>
- Evans, T., Macauley, P., Pearson, M., & Tregenza, K. (2003). A decadic review of PhDs in Australia. In *Educational research, risks and dilemmas—Proceedings of the NZARE/AARE Joint Conference* (pp. 1–16). Australian Association for Research in Education. <https://www.aare.edu.au/data/publications/2003/eva03090.pdf>
- Geary, E., Allen, K. A., Gamble, N., & Pahlevansharif, S. (2023). Online learning during the COVID-19 pandemic: Does social connectedness and learning community predict self-determined needs and course satisfaction? *Journal of University Teaching & Learning Practice*, 20(1). <https://ro.uow.edu.au/jutlp/vol20/iss1/13>
- Hughes, R. T., Zhu, L., & Bednarz, T. (2021). Generative adversarial networks-enabled human-artificial intelligence collaborative applications for creative and design industries: A systematic review of current approaches and trends. *Frontiers in Artificial Intelligence*, 4, Article 604234. <https://doi.org/10.3389/frai.2021.604234>
- Igumbor, J. O., Bosire, E. N., Karimi, F., Katahoire, A., Allison, J., Muula, A. S., Peixoto, A., Otwombe, K., Gitau, E., Bondjers, G., Fonn, S., & Ajuwon, A. (2022). Effective supervision of doctoral students in public and population health in Africa: CARTA supervisors' experiences, challenges and perceived opportunities. *Global Public Health*, 17(4), 496–511. <https://doi.org/10.1080/17441692.2020.1864752>
- Janssen, V. (2012). Indirect tracking of drop bears using GNSS technology. *Australian Geographer*, 43(4), 445–452. <https://doi.org/10.1080/00049182.2012.731307>
- Jeyaraj, J. J., Wald, N., & Harland, T. (2021). Higher education teachers' experiences of becoming research active: Striving for university status in the Global South. *Asia Pacific Education Review*, 22, 417–425. <https://doi.org/10.1007/s12564-021-09688-8>

- Jiang, S., & Yan, X. (2020). Research on the effect of supervisor feedback for undergraduate thesis writing. *English Language Teaching*, 13(1), 43–50. <https://doi.org/10.5539/elt.v13n1p43>
- Kasneji, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., ... Kasneji, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, Article 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Kaur, A., Kumar, V., & Noman, M. (2022). Partnering with doctoral students in research supervision: Opportunities and challenges. *Higher Education Research & Development*, 41(3), 789–803. <https://doi.org/10.1080/07294360.2020.1871326>
- Kellerman, B. (2022, December 28). Leader of the year – 2022. *BK Barbara Kellerman*. <https://barbarakellerman.com/leader-of-the-year-2022/>
- Kelly, A., Sullivan, M., & Strampel, K. (2023). Generative artificial intelligence: University student awareness, experience, and confidence in use across disciplines. *Journal of University Teaching & Learning Practice*, 20(6). <https://doi.org/10.53761/1.20.6.12>
- Kenny, J., & Fluck, A. E. (2022). Emerging principles for the allocation of academic work in universities. *Higher Education*, 83(6), 1371–1388. <https://doi.org/10.1007/s10734-021-00747-y>
- Khosa, A., Wilkin, C., & Burch, S. (2023). PhD students' relatedness, motivation, and well-being with multiple supervisors. *Accounting Education*. <https://doi.org/10.1080/09639284.2023.2179889>
- Kift, S., Nelson, K., & Clarke, J. (2010). Transition pedagogy: A third generation approach to FYE – A case study of policy and practice for the higher education sector. *Student Success*, 1(1), 1–20. <https://doi.org/10.5204/intifyhe.v1i1.13>
- Kumar, N. N., Summerell, E., Spehar, B., & Cranney, J. (2021). Experiences of honours research students and supervisors during the COVID-19 pandemic: A pilot study framed by self-determination theory. *Frontiers in Education*, 6, Article 758960. <https://doi.org/10.3389/feduc.2021.758960>
- Kung, T.H., Cheatham, M., Medenilla, A., Sillos, C., De Leon, L., Elepaño, C., Madriaga, M., Aggabao, R., Diaz-Candido, G., Maningo, J., & Tseng, V. (2023). Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. *PLOS Digital Health*, 2(2), e0000198. <https://doi.org/10.1371/journal.pdig.0000198>
- Lee, Y., Wehmeyer, M., Palmer, S., Williams-Diehm, K., Davies, D., & Stock, S. (2011). The effect of student-directed transition planning using a computer-based reading support program on the self-determination of students with disabilities. *Journal of Special Education*, 45(2), 104–117. <https://doi.org/10.1177/0022466909358916>
- Levesque-Bristol, C. (2023). Fostering students' learning experiences in higher education – Reflections from student-centered pedagogy and course transformation. In L. Waller & S. Waller (Eds.), *Higher education – Reflections from the field*. IntechOpen. <https://doi.org/10.5772/intechopen.110327>
- Lodge, J. M., Thompson, K., & Corrin, L. (2023). Mapping out a research agenda for generative artificial intelligence in tertiary education. *Australasian Journal of Educational Technology*, 39(1), 1–8. <https://doi.org/10.14742/ajet.8695>
- Longden, B., & Yorke, M. (2009). Institutional rankings, marketing, and the needs of intending students. In J. Knight (Ed.), *Global perspectives on higher education* (pp. 47–64). Brill Publishing. https://doi.org/10.1163/9789087908164_005
- McCallin, A., & Nayar, S. (2012). Postgraduate research supervision: A critical review of current practice. *Teaching in Higher Education*, 17(1), 63–74. <https://doi.org/10.1080/13562517.2011.590979>
- Meer, N. M., & Chapman, A. (2014). Assessment for confidence: Exploring the impact that low-stakes assessment design has on student retention. *The International Journal of Management Education*, 12(2), 186–192. <https://doi.org/10.1016/j.ijme.2014.01.003>
- Mishler, E. G. (1991). *Research interviewing: Context and narrative*. Harvard University Press.
- Narayan, A. K., Northcott, D., & Parker, L. D. (2017) Managing the accountability-autonomy tensions in university research commercialisation. *Financial Accountability and Management*, 33(4), 335–355. <https://doi.org/10.1111/faam.12127>
- Nasiri, F., & Mafakheri, F. (2015). Postgraduate research supervision at a distance: A review of challenges and strategies. *Studies in Higher Education*, 40(10), 1962–1969. <https://doi.org/10.1080/03075079.2014.914906>

- Neumann, M., Rauschenberger, M., & Schön, E. M. (2023). We need to talk about ChatGPT: The future of AI and higher education. *SerWisS*. <https://doi.org/10.25968/opus-2467>
- Newstead, T., Eager, B. & Wilson, S. (2023). How AI can perpetuate – or help mitigate – gender bias in leadership. *Organizational Dynamics*. Advance online publication. <https://doi.org/10.1016/j.orgdyn.2023.100998>
- O'Dea, X., & O'Dea, M. (2023). Is artificial intelligence really the next big thing in learning and teaching in higher education? A conceptual paper. *Journal of University Teaching & Learning Practice*, 20(5). <https://doi.org/10.53761/1.20.5.05>
- Open AI. (2022, November 30). *Introducing ChatGPT*. <https://openai.com/blog/chatgpt>
- Perkins, M. (2023). Academic Integrity considerations of AI large language models in the post-pandemic era: ChatGPT and beyond. *Journal of University Teaching & Learning Practice*, 20(2). <https://doi.org/10.53761/1.20.02.07>
- Popenici, S. (2023). The critique of AI as a foundation for judicious use in higher education. *Journal of Applied Learning and Teaching*, 6(2). <https://doi.org/10.37074/jalt.2023.6.2.4>
- Pucciarelli, F., & Kaplan, A. (2016). Competition and strategy in higher education: Managing complexity and uncertainty. *Business Horizons*, 59(3), 311–320. <https://doi.org/10.1016/j.bushor.2016.01.003>
- Reddy, K. S., Xie, E., & Tang, Q. (2016). Higher education, high-impact research, and world university rankings: A case of India and comparison with China. *Pacific Science Review B: Humanities and Social Sciences*, 2(1), 1–21. <https://doi.org/10.1016/j.psr.2016.09.004>
- Retraction Watch. (2023, March 8). *Ob-gyn loses PhD after committee finds he made up research*. Retrieved March 15, 2023, from <https://retractionwatch.com/2023/03/08/ob-gyn-loses-phd-after-committee-finds-he-made-up-research/>
- Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education?. *Journal of Applied Learning and Teaching*, 6(1). <https://doi.org/10.37074/jalt.2023.6.1.9>
- Ryan, R. M. (2023). *The Oxford handbook of self-determination theory*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/97801976000047.001.0001>
- Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. The Guilford Press. <https://doi.org/10.1521/978.14625/28806>
- Sambell, K., & Hubbard, A. (2004). The role of formative 'low-stakes' assessment in supporting non-traditional students' retention and progression in higher education: Student perspectives. *Widening Participation and Lifelong Learning*, 6(2), 25–36. <https://www.ingentaconnect.com/content/openup/jwpl/2004/00000006/00000002/art00004>
- Sawatsky, A. P., O'Brien, B. C., & Hafferty, F. W. (2022). Autonomy and developing physicians: Reimagining supervision using self-determination theory. *Medical Education*, 56(1), 56–63. <https://onlinelibrary.wiley.com/doi/full/10.1111/medu.14580>
- Shaw, J. (2009). The diversity paradox: Does student diversity enhance or challenge excellence?. *Journal of Further and Higher Education*, 33(4), 321–331. <https://doi.org/10.1080/03098770903266018>
- Spinrad, M. L., Relles, S. R., & Watson, D. L. (2022). Not in the greater good: Academic capitalism and faculty labor in higher education. *Education Sciences*, 12(12), Article 912. <https://doi.org/10.3390/educsci12120912>
- Stokel-Walker, C., & Van Noorden, R. (2023). What ChatGPT and generative AI mean for science. *Nature*, 614(7947), 214–216. <https://doi.org/10.1038/d41586-023-00340-6>
- Thorp, H. (2023). ChatGPT is fun, but not an author. *Science*, 379(6630), 313–313. <https://doi.org/10.1126/science.adg7879>
- Tice, D., Baumeister, R., Crawford, J., Allen, K. A., & Percy, A. (2021). Student belongingness in higher education: Lessons for Professors from the COVID-19 pandemic. *Journal of University Teaching & Learning Practice*, 18(4), Article 2. <https://doi.org/10.53761/1.18.4.2>
- United Nations Educational, Scientific and Cultural Organization. (2020). *Towards universal access to higher education: International trends*. UNESCO IESALC. <https://unesdoc.unesco.org/ark:/48223/pf0000375686>

- Wisker, G., McGinn, M. K., Bengtsen, S. S., Lokhtina, I., He, F., Cornér, S., Leshem, S., Inouye, K., & Löfström, E. (2021). Remote doctoral supervision experiences: Challenges and affordances. *Innovations in Education and Teaching International*, 58(6), 612–623. <https://doi.org/10.1080/14703297.2021.1991427>
- Wisker, G., Robinson, G., & Shacham, M. (2007). Postgraduate research success: Communities of practice involving cohorts, guardian supervisors, and online communities. *Innovations in Education and Teaching International*, 44(3), 301–320. <https://doi.org/10.1080/14703290701486720>
- Yang, H., & Kyun, S. (2022). The current research trend of artificial intelligence in language learning: A systematic empirical literature review from an activity theory perspective. *Australasian Journal of Educational Technology*, 38(5), 180–210. <https://doi.org/10.14742/ajet.7492>
-

Corresponding author: Joseph Crawford, Joseph.crawford@utas.edu.au

Copyright: Articles published in the *Australasian Journal of Educational Technology* (AJET) are available under Creative Commons Attribution Non-Commercial No Derivatives Licence ([CC BY-NC-ND 4.0](https://creativecommons.org/licenses/by-nc-nd/4.0/)). Authors retain copyright in their work and grant AJET right of first publication under CC BY-NC-ND 4.0.

Please cite as: Cowling, M., Crawford, J., Allen, K.-A., & Wehmeyer, M. (2023). Using leadership to leverage ChatGPT and artificial intelligence for undergraduate and postgraduate research supervision. *Australasian Journal of Educational Technology*, 39(4), 89-103. <https://doi.org/10.14742/ajet.8598>