ChatGPT and academia on accounting assessments

Indra Abeyesekera*
Charles Darwin University, Darwin, NT 0800, Australia

1. Introduction

With nights and days passing serenely, a big surprise as morning dawns on November 30, 2022, has made a lasting change in how we have taken what we know, to know, and beyond for granted. ChatGPT, a Generative Pre-trained Transformer, is a large language model developed by the Open Artificial Intelligence (OpenAI) company, the same organisation partnered with Microsoft and is part of the larger family of GPT (Yu, 2023). It has chatbot software with an application and web interface and uses large datasets, computing power, and algorithms to present it as an intelligent model. ChatGPT-3.5 version (referred to in this paper as ChatGPT), free for the public to use, has received over 175 billion training parameters (Homolak, 2023; Naveen, 2023; Yu, 2023). These parameters are components of a larger language model that describes the skills required to solve a particular problem, such as generating text (Lutheievich and Schmeizer, 2023).

The interface can simulate human conversations by accepting text and voice commands to operate without human assistance. ChatGPT processes information using rules and instructions (algorithms) chosen by the machine based on the learning it has carried out (Homolak, 2023; Yu, 2023). ChatGPT extends chatbot capabilities by combining them with deep learning and language models in its architecture (Radford et al., 2018).

1.1. Motivation

Artificial intelligence associated with ChatGPT has introduced opportunities for improving pedagogy. For example, it can provide personalised learning by creating educational content that is accessible to meet students’ interests, skills and learning goals (Javaid et al., 2023). Risks to academia arise from the increased propensity of students to compromise academic integrity by replacing student-required effort and cognitive resources with ChatGPT. Scepticism towards inaccurate information that students can rely on without questioning and increasing overreliance on it as a helpmate ChatGPT for learning has become an academic concern (Sok and Heng, 2023). Another perceived concern is the lack of algorithmic transparency of ChatGPT, which leads to concerns about auditable credibility (Diwedi et al., 2023).

However, ChatGPT has brought enormous benefits to stakeholders in mining unstructured data embedded in narratives (texts), visuals...
(photos and videos), and audio. The accuracy of the results increases with the volume of data processed and the computational resources used to run the algorithms (Dwivedi et al., 2023). The discriminated classifications undertaken to fine-tune unstructured data into meaningful information and present them as solutions to questions raised without human intervention at enormous speed are extremely helpful in advancing knowledge by making tacit explicit to meet each enquiry (Radford et al., 2018).

1.2. Aim

Against this backdrop, this study investigates the risks and opportunities in student academic accounting assessments. Accounting plays a crucial role in reducing uncertainty in control and making decisions in a given context and is a major field of study in academic institutions (Mellervik et al., 1987). The reason for choosing an accounting course is that the accounting syllabus comprises both narrative-based and mathematical-based learning. However, there is more skewness towards number-orientated learning and hence assessments.

The accounting syllabus contains various aspects of accounting knowledge, and financial accounting is fundamental for all students. Financial accounting information helps people evaluate the past, which is otherwise overshadowed by daily activities, supported by a fiscal-given context and is a major field of study in academic institutions crucial role in reducing uncertainty in control and making decisions in a

accessing tools such as ChatGPT in the context of assessments. The following section discusses the literature on artificial intelligence ecosystem to uphold high moral ground.

I. Abeysekera

2. Literature

2.1. What is it?

ChatGPT serves as a lighthouse with the metaphor of a sea route, where the light it surveys is perceived as a route for academic construction or destruction. ChatGPT provides the shining light, and the pathways are chosen by academia, defined by the five Ws: what it is being used, who uses it, where it has been used, when it is used, and why it is being used (Robertson, 1946). It is being used for almost any enquiry prompted by text and human voice through the chatbot; what it is used is for now, shaping the future of knowledge, as knowing of and knowing how (Abeysekera, 2021).

2.2. Why is it?

There is a good reason why ChatGPT used the chatbot approach. Humans cherish communication and interaction, and chatbots facilitate it. Chatbots are communication gateways, often with the support of an internet connection. Chatbot functionality embedded in ChatGPT can be embedded into various software applications used on technological devices such as laptops and communication devices such as mobile phones. ChatGPT provides access to its Knowledge Portal for Human Investigation to facilitate online and face-to-face learning. The natural language processing technology used in the chatbot can make it user-friendly to accept and create human-like conversational language processing, mimicking real-life instructor scenarios (Lin et al., 2022). It is simple to use; the basic version of ChatGPT version 3.5 is offered free of charge and can accept voice and text command instructions through the chatbot. Deep learning has been used to understand human subtleties and search the vast data available on the Internet to produce the most accurate solutions (Haleem et al., 2022).

2.3. Who is it?

ChatGPT is a future-oriented support tool for knowledge production and reproduction (Haleem et al., 2022). Its wide application capabilities span industry sectors, covering a wide range of actual and potential users and facilitating increased productivity through machine help (Lin, 2023). The user-friendly chatbot that can facilitate learning has become a beneficial technological tool to academia but is foreseen by academics with apprehension as a risk to academic integrity, raising concerns about students’ increased propensity to compromise academic integrity and engage in plagiarism. Academic institutions constantly evolve their policies and procedures to maintain academic integrity in this dynamic ecosystem (Cotton et al., 2023).

2.4. Where is it?

The Silicon Valley-based OpenAI, ChatGPT, has garnered 1 million users in less than a week, exceeding the benchmarks achieved by Instagram in 10 weeks and Facebook in 43 weeks. It reached 100 million active users in two months, making it the fastest-growing consumer application (Dempere et al., 2023). The administrative areas of academic institutions have used ChatGPT to streamline their workflows and increase student retention (Lin, 2023). Any geographic location with internet access can access the free version of ChatGPT 3.5 at no cost, magnifying the potential to provide quality education endorsed by the United Nations Sustainable Development Goal (UN SDG) number 4 (United Nations, 2023).

2.5. When is it?

ChatGPT is used in academia for teaching, learning, and assessment. However, the pitfalls are that academia can delegate responsibility to the ChatGPT, which becomes the content owner of the work. It is a moral duty in academia to cite the source of information and concepts obtained from others who have contributed to understanding the materials or making the arguments presented. The rule followed is to cite the original source as much as possible when in doubt. These include quoting, paraphrasing (restating or rephrasing the original idea), summarising, presenting facts, statistics, dates, and information, and indebtedness (Boston University, 2022). Although the best practice is to cite, certain instances may not require citing, such as common knowledge, generally accepted or observed facts, or writing about yourself and your own experiences (Boston University, 2022). Because ChatGPT is an artificial intelligence, it has no legal personality to own intangible assets such as copyright.
under European and US law (European Commission, 2023).

2.6. The Research Gap

Increased student engagement, collaboration, and accessibility to ChatGPT has raised concerns for academia regarding whether students can compromise their academic integrity and plagiarism without citing the sources from which information was obtained for their assessable task (Cotton et al., 2023). The common knowledge domain is dynamic and constantly changes. For example, computers and calculators prevented students from memorising much of the calculations, which has now become common knowledge and is an accessible help tool in taking assessments (Kim and Wong, 2023). The knowledge generated by open artificial intelligence tools such as ChatGPT can bring out such changes to what is common knowledge, then as not required to be cited, but for now, it is a source that must be cited on moral grounds (Sison et al., 2023).

The question is whether ChatGPT is an effective human agent for student assessments. A study was conducted across several countries and several course units in the accounting syllabus and reported that ChatGPT performed less well, with a 47.5% score than the average student performance of 76.5% when assessments did not earn credits. When the assessments earned partial credit, ChatGPT scored 56.5%, but students scored less with a 15.8% score. It is not known how the ChatGPT score changed from no credit to partial credit. It is possible that they used a different set of assessment questions in the two scenarios, thus making the results not comparable (Wood et al., 2023). Furthermore, their study covered course units in accounting syllabuses in several countries, with each course unit contributing to variability because learning outcomes and their nature and assessment structures are different. In addition, contextual geographic variabilities can also influence individual assessment scores in course units. Apart from the noted study, there is little research that has tested ChatGPT to see whether it can provide accurate solutions to assessment questions. This study aims to refine the previously broad-brush findings by focussing on two selected financial accounting course units and to what extent ChatGPT can become a human agent.

3. Theoretical framework

Figure 1 shows the agency framework used in the study. ChatGPT and academia are two parties involved in the interplay in co-creating academic production with student assessments. There are two distinct stakeholders in academia: academics and students. The relationship between academia and ChatGPT constitutes social agency, on the presumption that ChatGPT and academia are social, and between them, they engage in human-to-human-like interactions (Atkinson et al., 2018). Academia takes a deontological position to knowledge production because right or wrong based on the greatest number of people who have access to knowledge (Kay, 2018). Academia takes a deontological position to knowledge promoting the highest good as developing reasoning in people as a duty – rights and wrongs are determined by utilitarianism (Musselbrook, 2013).

ChatGPT provides access to structured knowledge to anyone and uses a utilitarian position because it determines right or wrong based on the greatest number of people who have access to knowledge (Kay, 2018). Academia takes a deontological position to knowledge reconfiguring the created moral agency. The morality of what is right and wrong is also combined with whether the greatest number of people have access to knowledge. It is in this where the tension is, where academia thinks that the academic duty is to assess knowledge- and skill-based competencies for certification.

3.1. Social agency

Social agency can influence the broader world through intentional actions (Billett, 2008). Social agency is a structure created by academia and ChatGPT; an agency is governed by rules, norms, power, and the knowledge required to take effective action. Social agency is constantly.

---

![Fig. 1. Agency Framework on Academia and ChatGPT Interaction.](image-url)
Becoming a human agent means that humans intentionally influence their functioning and life circumstances. In relation to academia, it is about how academia can intentionally direct scholarship.

There are four properties of human agency: First, intentionality. Academic institutions’ intentions influence and are influenced by social and moral agents in the ecosystem, such as ChatGPT, knowledge creation and assessments. Second, the forethought of visualising academia in the future through probabilistic scenarios helps academia to favourably respond to disruptions in future ecosystems. Third, self-regulating academic behaviour at single institutions or at more aggregated levels, such as within a country or a block of countries, vests them with greater power to act with resilience. Fourth, academia, through constant self-reflection, can determine and adjust toward sound pursuit and thought, for instance, to co-create knowledge with ChatGPT (Bandura, 2006).

3.4. Research questions

Based on these agency perspectives, academia can decide whether to allow ChatGPT to become a human agent for student assessments. It depends on the capabilities of ChatGPT to what extent it is effective in helping students. The help ChatGPT can offer can differ depending on the complexity of the task. For example, Introductory Financial Accounting is less complex than Advanced Financial Accounting. It also depends on the level of capabilities that ChatGPT has as a language processor versus a numerical processor. Against this backdrop, this study proposes the following four research questions. These are about how academia can intentionally direct scholarship.

1. RQ1. To what extent can ChatGPT provide accurate answers to accounting number-based assessment questions in the Introduction to Financial Accounting course unit?
2. RQ2. To what extent can ChatGPT provide accurate answers to the accounting narrative-based assessment questions in the Introduction to Financial Accounting course unit?
3. RQ3. To what extent can ChatGPT provide accurate answers to accounting number-based assessment questions in the Advanced (Corporate) Accounting course unit?
4. RQ4. To what extent can ChatGPT provide accurate answers to accounting narrative-based assessment questions in the Advanced (Corporate) Accounting course unit?

4. Methodology

Figure 2 shows the methodological approach. The two units tested are Introductory Financial Accounting and Advanced Financial Accounting. Each course unit comprises 10 questions: five narrative and five numerical questions. A question is entered separately into ChatGPT and ChatGPT4. These are explained in detail next.

4.1. Sampling frame

Since rules in social agency are a co-creation between academic institutions and ChatGPT, students’ assessments are determined by the rules set out by the academic institutions for students, where ChatGPT offers free knowledge. These rules can be largely classified into narrative-based and numerical-based assessments (Stone, 2021).

The study selected assessments that involve numbers because students can have ‘number anxiety’ and develop emotional reactions to arithmetic and mathematics. Number anxiety is separate from general anxiety and is not related to general intelligence. Those with lower and higher grades in mathematics tend to have higher number anxiety, whereas those in the middle range do not (Dowker et al., 2016; Dreger and Aiken, 1957). Therefore, such students will likely seek ChatGPT as a human agent in their assessments. Although students seek such help, it is not known whether ChatGPT is an effective human agent that presents correct answers.

Financial accounting numbers follow special rules with double-entry principles. The same number is recorded twice in two places as two whole numbers. In another situation, one whole and the other constituting two or more numbers to make it the whole. The third situation is where both numbers constitute two or more numbers to make the whole. Such number crafting has made accounting a skilled craft (Sangster, 2016). Skill mastery in double-entry bookkeeping is crucial (Sangster, 2022). A higher proportion of students, even those without number anxiety, can seek help to solve the assessment questions through ChatGPT.

This study sets its sampling limits within the units of the financial accounting course. The two units selected were the Introductory Financial Accounting course, which has a lower technical complexity, and the Advanced Financial Accounting course, which has a higher technical complexity.

4.2. Data

Data comprise a sample of multiple choice quizzes provided for students obtained from the textbook for Introductory Financial Accounting (Weygandt et al., 2013), and questions developed by the author for Advanced Financial Accounting influenced by prescribed and relevant textbooks (Abeysekera, 2007; Deegan, 2019). Although textbook release versions are a few years before the study, the validity of the questions is intact because the questions are based on testing the concepts and applications of accounting standards that have remained largely the same. The questions are pre-validated to meet course unit rigour standards through student testing. Completing the Introductory Financial Accounting course is a prerequisite to undertaking the Advanced Financial Accounting course.

Introductory Financial Accounting is a core course unit in the accounting syllabus where students learn the fundamentals of financial account processing and reporting on a single entity. Advanced Financial Accounting is the most advanced form of financial accounting that
emphasises consolidation (group) accounting or business combination.

The study tested ten multiple-choice questions, 5 of them number-based and the other five narrative-based. They led to testing the four research questions. Each multiple-choice question contained four answer choices in which one answer was correct. The study introduced each multiple-choice question with an introductory statement, ‘Please give the correct answer’, into the ChatGPT (version 3.5) chatbot. The ChatGPT solution was checked to determine whether it produced an accurate answer.

The topics in each course unit were selected randomly, and multiple-choice questions were selected randomly for each topic. The six topics in Introductory Financial Accounting were accounting equations, adjusting entries to accounts, inventory, accounts receivables, and liabilities. The topics included in Advanced Financial Accounting were submitted tax, business combination, intragroup inventory transactions, and intragroup noncurrent asset transactions. Where possible, narrative and numerical data were obtained to match the topics.

This study obtained questions for Introductory Financial Accounting from the end-of-chapter questions provided in the textbook. Advanced Financial Accounting questions were obtained from questions attempted by students who were tested across two Australian universities. They represent real assessment questions. In addition to such objective validations, student feedback across several years that supports subjective validations suggests that they are rigorous and too difficult to obtain correct answers. The study established content and face validity by authoritative sources of textbooks as referenced assessment materials. It established discriminant validity by the choice of topics that are independent of each other. The questions have high reliability as there is only one correct answer. The study strengthened reliability further by the fact that the questions originated or were adopted from prescribed textbooks that several Australian universities use for these course units.

5. Results

5.1. Introductory financial accounting course unit

Table 1 summarises the results of ChatGPT performing as a human agent for a student in the multiple choice assessment questions. In relation to RQ1, can ChatGPT provide accurate answers to the accounting number-based assessment questions in the Introduction to Financial Accounting course unit? The findings showed a score of 4 out of 5, leaving it in the 80th percentile. In relation to RQ2, can ChatGPT become a human agent for students to provide accurate answers to accounting narrative-based assessment questions in the Introduction to Financial Accounting course unit? The findings showed a score of 4 out of 5, leaving it in the 80th percentile. The results show that there is no discriminatory difference in ChatGPT between narrative-based and numerical-based in providing solutions to financial accounting questions. ChatGPT obtaining wrong answers to the questions is analysed next using the ChatGPT output produced for those two questions.

ChatGPT got the Question 2 answer wrong. The question was about where the supplies expenses have been recorded in full as an expense, but at the year-end, unconsumed expenses become an asset to be carried forward to the following year. However, ChatGPT assumed that the payment for supply expenses was initially capitalised as an asset. The accounting principle used by ChatGPT was incorrect because expenses are assumed to be consumed in the reporting period, and when they are paid for or purchased, they are immediately shown as expenses.

ChatGPT got the answer to Question 6 wrong. The question required identifying an event that would not be recorded in the accounting records. ChatGPT identified all events that must be recorded. However, in the answer choices, one choice answer was that staff is terminated, but the cost was not given in the answer to include in the accounting records. ChatGPT assumed that the cost was known and noted it as an event that must be included in the accounting records.

When providing answers to the Introduction Financial Accounting Questions, the questions were within the ChatGPT Artificial Intelligence boundaries. However, it got two of them wrong because it misunderstood the fundamentals behind the questions and made incorrect underlying assumptions.

5.2. Advanced financial accounting course

In relation to RQ3, can ChatGPT become a human agent for students to provide accurate answers to accounting number-based assessment questions in the Corporate (Advanced) Accounting course unit? ChatGPT got two of five questions correct, leaving it in the 40th percentile. In relation to RQ4, can ChatGPT become a human agent for students to provide accurate answers to accounting narrative-based assessment questions in the Advanced Accounting course unit? The results show that there is no discriminatory difference in ChatGPT between narrative-based and numerical-based in providing solutions to financial accounting questions. ChatGPT got 3 out of 5 questions correct, leaving it in the 60th percentile.

ChatGPT got the Question 13 answer wrong on the intragroup inventory transaction. It showed the correct calculation of intragroup unrealised profits to be eliminated as a consolidation adjustment. However, it eliminated the unrealised profits in full ($10,000), rather than the remaining, which has in fact been calculated ($5,000). A possible algorithmic selection could have occurred because $5,000 was not a choice answer. However, the correct answer is related to eliminating sales of $50,000 because the initial trading transaction must also
be eliminated. However, ChatGPT had no idea that sales-purchase intragroup transactions must be eliminated, as this was not discussed in the ChatGPT solution. The ChatGPT learning competency was inadequate to deal with it.

ChatGPT got the answer to Question 14 wrong. ChatGPT demonstrated a straightforward understanding of business combinations but fell short of the deep level of understanding required to solve such a problem. For instance, ChatGPT knew that a debit transaction must be eliminated by a credit entry in consolidation accounting for a consolidation adjustment. Hence, it debited the Accumulated Depreciation of $500,000 because it is a credit amount. However, in situations where a net credit balance is required for accumulated depreciation, that amount is shown. However, ChatGPT was less competent in solving the problem. Further, ChatGPT was less than competent to realise that assets sold at a gain within the group give rise to deferred tax assets rather than a deferred tax liability.

ChatGPT got the answer to Question 15 wrong. ChatGPT understood how to mathematically calculate the minority interest in consolidation. However, it considered only the fair value of equity in the parent entity, which is based on the parent entity concept of consolidation. Business combinations follow the concept of the group entity and not that of the parent entity, meaning that the equity of the group entity is considered when calculating the share of minority interest.

ChatGPT got the answer to Question 17 wrong. ChatGPT showed a linear understanding of the concept of control as actual control but was incapable of knowing that control capacity is the criterion used to decide whether that entity or person has a controlling interest in another entity.

ChatGPT got the answer to Question 20 wrong. ChatGPT showed a less than capable understanding of the effect of revaluation on business combinations. The chosen answer is valid for ordinarily occurring revaluation transactions in single entities. However, revaluations are considered special event transactions for consolidation. Higher asset revaluations in consolidation accounting delay income tax expense as a delayed tax liability. ChatGPT could not realise this.

5.3. Additional analysis

The subscribed version is multimodal, where it processes images, has more processing power, and has improved accuracy. It is estimated to have received training on more than a trillion parameters (Naveen, 2023). This study reviewed the questions using the ChatGPT4 (ChatGPTPlus) version accessible by subscription, with a focus on the questions that produced incorrect answers.

5.3.1. Introductory financial accounting course unit

Table 3 summarises the results of the Introductory Financial Accounting course assessment answers. ChatGPT4 produced a different but incorrect answer to Question 2. That said, the software followed logical but incorrect assumptions about how much supplies were unused and should be capitalised as an asset at the end of the year. ChatGPT4 produced all correct answers. In reviewing Question 6, ChatGPT4 provided a reasonable and correct answer. ChatGPT4 passed the test to be in the 90th percentile.

5.3.2. Advanced financial accounting course unit

Table 4 summarises the results of the answers to the Advanced Financial Accounting course assessment using ChatGPT4 with a focus on the questions in which ChatGPT gave incorrect answers.

ChatGPT4 produced an incorrect but different answer to ChatGPT for Question 13. ChatGPT4 provided more explanation. Calculates unrealised profits in intragroup inventory that must be eliminated. However, it did not contain the knowledge that all intragroup transactions, such as sales and cost of sales, must also be eliminated. ChatGPT4 produced the correct answer to Question 14 in which ChatGPT produced a wrong answer and provided a detailed reasoning for the calculation. ChatGPT4 selected the same answer choice as ChatGPT, which was a wrong answer, showing that it does not deeply understand this minority interest calculation in consolidation accounting. In Question 17, ChatGPT4 produced the same incorrect answer as ChatGPT. Both versions did not understand that the capacity to control an entity rather than actually controlling it is the criterion to consider whether another entity is controlled. ChatGPT4 produced the correct answer and provided a reasoned explanation for its choice answer selection, whereas ChatGPT got it wrong. ChatGPT4 produced a score of in the 70th percentile.

It is evident that ChatGPT4 has undergone deeper training and has corrected anomalies by considering a wider breadth and depth of data and information. However, there is more training to undergo, and it is toward perfection.

5.3.3. Question-type performance

Table 5 summarises the ChatGPT and ChatGPT4 performance in producing correct answers to question types such as numerical and narrative questions where there were five numerical and five narrative questions in the course unit assessment. There were no notable differences in performance between question types, although ChatGPT4
performed better than ChatGPT.

6. Discussion

6.1. Pedagogical instruction

The solutions provided by ChatGPT showed that it is a solution provider rather than a teacher or instructor. Those who are competent in the topic area can understand these solutions. It does not scaffold (a metaphor used for erecting a building with support) knowledge to progressively teach novice learners to develop unassisted competencies (van de Pol et al., 2015). Furthermore, the findings showed that ChatGPT is not a foolproof solution provider, especially when questions have discipline-specific underlying assumptions and increased technical and task complexity.

ChatGPT can be constructive to a competent learner who has reached the competency level to further develop critical understanding. As research has indicated, high achievers can have a fear of numbers, and they can benefit by using ChatGPT solutions as validation checks for their learning. It can be destructive to an incompetent learner and can serve as a platform to simply find the solution either as a deliberate attempt or as a channel to ease their fear of numbers (Dowker et al., 2016; Dreger and Aiken, 1957).

ChatGPT does not provide scaffolding for novice learners to take over their learning and develop individual competencies to be less or not reliant on it (Van de Pol et al., 2011). ChatGPT becoming a scaffolding-learning platform can be counterproductive to a profit-making business model such as ChatGPT, as it requires more customers to use its services while maintaining a continuing dependency.

6.2. Deception versus detection

The competition for academia is two-fold. They arise not from the fact that artificial intelligence such as ChatGPT can be trusted, but rather from academic institutions that can trust academia to uphold the moral agency of dutyfulness. First, regarding harnessing the full potential of ChatGPT to promote academic learning, the problem for academia is the seemingly perceived deception of academic assessments by students. Detection software has not caught up fast enough to do so. Academic institutions rely on the social agency of software providers to take care of compliance tasks. Instead, academic institutions must develop the moral agency of inculcating academia to become dutiful in scholarship.

Second, to promote and ensure that ChatGPT becomes a constructive tool for academic learning, addressing it requires academia to inculcate values of moral agency that students willingly embrace to become lifelong learners. The willing moral acceptance by academia must then be supported by a social agency in which norms about knowledge are co-created in the academic ecosystem, where ChatGPT exists and is a friendly social agent. This embrace expands the human agency of academia to intentionally and proactively influence the functioning of scholarship in academic institutions.

Until these issues are resolved, temporary measures are to shut down the Internet when undertaking student assessments, consider such assessment scores as hurdles to completing the course unit, and invite late students to ensure that they produce solutions to assessment questions by treating ChatGPT as a foe social agent.

Such measures are temporary because the fundamental issue to resolve is the ontology of knowledge and the epistemology of knowledge—such as nature, justifications, and limitations of knowledge and how it should be assessed to certify academic learning (Hathcoat et al., 2018). These can lead to continuously evolving debates because knowledge epistemologies are perspective-based, objective, subjective, and constructive. Such resolutions can lead to eliminating the ‘cheat factor’ from assessments, which is considered a major risk to academia due to ChatGPT’s presence in the learning ecosystem and its acceptance as an unreserved social agent.

7. Conclusions

This original study investigated whether ChatGPT can become a human agent for a student in two financial accounting course unit assessments comprising multiple choice questions: Introductory Financial Accounting and Advanced Financial Accounting. The assessment questions comprised numerical and narrative questions. The findings showed that ChatGPT could provide correct answers at the 80th percentile and ChatGPT4 at the 90th percentile in the Introductory Financial Accounting course unit. The ChatGPT produced correct answers at the 50th percentile and ChatGPT4 at the 70th percentile. There was no notable difference between numerical and narrative questions performance by ChatGPT and ChatGPT4 in producing correct answers. The implications are discussed next.

7.1. Implications for academia

7.1.1. Implications for knowledge

By making knowledge more accessible, ChatGPT has lifted the benchmark of assumed knowledge to all stakeholders, including the scholarly community comprising academics and students. Automated help provided by ChatGPT has invalidated the Bloom taxonomy, which is centred on the student remembering, understanding, applying, analysing, and evaluating knowledge.

In Bloom’s taxonomy, the apex produces new and original knowledge. However, ChatGPT has fallen short of creating original knowledge beyond what is known (Adams, 2015). Because there is no commonly accepted single definition of originality, it is necessary to agree on it in given contexts. The context could be whether it is a research- or coursework-based course, and within them, the course units introduce finer contexts into the course. Originality is discipline-specific; in accounting courses, it is different from that in medical courses. Originality is also topic-specific. In accounting courses, management accounting, financial accounting, and accounting theory evoke different originality. Commonly, creative skills and lateral thinking contribute to originality (The University of Melbourne, 2012). These are skill sets that academia must foster to develop future students.

7.1.2. Learning objectives and assessments

The learning objectives are assessed in the course units. A refocus in academia on learning objectives should shift to the originality of knowledge. ChatGPT can process unstructured data, convert it into information, and share it as knowledge. It can meet some criteria that meet originality, such as synthesising information differently and providing a new interpretation using known information. However, there are originality tasks that ChatGPT is incapable of meeting, at least for now. These include testing someone else’s idea in a different context or developing a research tool. Hence, academic focus on learning and assessment should shift toward original knowledge.

The structuring assessment must focus on facilitating and upholding academic integrity, rather than providing assessments that increase
compromising academic integrity, and it is the responsibility and can become a performance indicator of academicians rather than students. Assessment structures that facilitate the protection of academic integrity lead to the efficiency and effectiveness of the resource use of academic institutions. Enforcing compliance for compromising academic integrity leads to academic institutions’ resource consumption with little value creation towards pedagogical and research indicators.

7.1.6. Benefitting from open innovation

Although not purposefully built for academia, by opening up its systems and processes to accommodate ChatGPT, academia stands to appropriate value to benefit from an open innovation to which it has not directly contributed or partnered for its development (Almirall and Casadesus-Masanell, 2010). The human-like output generated by artificial intelligence in ChatGPT as a pre-trained transformer, with prompts given to it through its Chatbot, has raised questions about the created authorship. In previous chapters on open innovation in organisations, the clearer demarcations of resource contributions by insiders and outsiders enabled a clearer dissection of value creation resulting from the combined efforts. ChatGPT has blurred lines of value creation as it uses preexisting content to create value, value appropriation by ChatGPT users, and value ownership as to who owns the rights to the ChatGPT output; answers to these ethical and legal challenges are likely to unravel over time (Lee, 2023; Zirpoli, 2023). Satisfactory resolution of ethical dilemmas by academia can lead to the extent of associating and embracing ChatGPT. But for now, ChatGPT has highlighted the benefits of open innovation for academia to facilitate student learning, where academia stands to benefit from decreased administrative, teaching, learning, and research costs by gaining access to vast knowledge (Ray, 2022).

7.2. Implications for ChatGPT

What would the large language model look like if ChatGPT were a model developed by academia? Although it is a philosophical question, it raises a question of reflection on its business model and moral consciousness. Although the business model has no unified definition, one definition refers to it as representing a firm’s core logic and strategic choices for creating and capturing value within a network (Shafer et al., 2005, p. 202). Strategic decision making is a key aspect of the business model. Such reflections are not intended to decrease the amount of money generated but rather to increase it. However, such reflections can provide scenario-based solutions as to how to produce the set amount of money by differently contributing to social agency, moral agency, and human agency. ChatGPT can champion social agency, increasing the opportunities for academia to optimise resources, thereby reducing the resources spent on acquiring and disseminating knowledge, leaving academia with more funds to redirect to developing students’ skills. ChatGPT can champion moral agency by sharing information and developing co-patented software with academia to enhance student learning in the context of academic institutions.

7.3. Implications for the Society

7.3.1. Inclusive and quality education (UN SDG 4)

ChatGPT has enormous potential to contribute to the Sustainable Development Goals of the United Nation. SDG 4 of the UN ensures quality, inclusive, and equitable education and promotes lifelong opportunities for all. The United Nations has indicated that by 2030, 300 million students will lack basic numeracy and literacy skills. One in six countries achieves the universal secondary education completion target, and 84 million young people and children will be out of school (United Nations, 2023). ChatGPT has enabled the capacity to bridge knowledge gaps found in different corners of the world, with sub-Saharan Africa experiencing the greatest disadvantage with additional associated issues compounded by less access to drinking water, electricity, computers, and the Internet. However, enabling education through ChatGPT requires people to have Internet access. In 2020, only 60% of the world’s population was online, and in 2022, it will be 65%. Infrastructure unavailability and unaffordability restrict ChatGPT’s ability to play a role in emancipating people through education (Ritchie et al., 2023; United Nations, 2023). Governments, international and bilateral donor agencies, and NGOs can provide greater access to new information and communication technologies by finding cost-effective internet applications in rural settings (James, 2005). Internet access reduces educational inequality (Korkmaz et al., 2022).
7.3.2. Establishing partnerships for sustainable development (UN SDG 17)

Earlier in this paper, it was pointed out that human agency is co-created through social and moral agency. Academic institutions are locally focused in specific towns, cities, and countries, as many still rely on brick and mortar as centres of educational excellence. However, academia has somewhat overlooked the marginalised populations identified in the UN SDGs to meet their educational needs because their business models cannot support them, requiring academia and ChatGPT to co-create a value-creation pathway for the marginalised population. Such partnerships are consistent with the UN SDG 17 Partnerships for the Goals; in this case, the partnership supports the UN SDG 4 Quality education. It requires the adoption of a unified strand of ethics that is dutiful to the greatest number of people.

Free software sites such as The R Project for Statistical Computing will likely release similar large language models such as ChatGPT. Search engines are also likely to develop such models and have third-party models as additional software for their search engines. These future endeavours will progressively define the dynamic concept of knowledge from which society can benefit.

7.4. Limitations and future research

7.4.1. Limitations

The findings must be interpreted within the context of the multiple-choice questions in the two financial accounting course units used as an illustration and testing. The accounting syllabus has nine technical and professional competency areas: 1) accounting systems and processes, 2) financial accounting and reporting, 3) audit and assurance, 4) business law, 5) economics, 6) finance and financial management, 7) management accounting, 8) quantitative methods, 9) taxation, 10) information and communications technology, and 11) business acumen (CPA Australia, 2023). This study examined financial accounting and reporting. The findings can differ in assessment testing in other competency areas. Second, the findings are valid at the testing date because large language models continuously undergo deep learning; consequently, findings can differ when tested on a future date. The findings apply to ChatGPT testing as large language models with different or exceeding capabilities can be developed. Third, the results are generalizable to course units with similar learning objectives and assessment contents.

7.4.2. Future research

Future research can test ChatGPT capabilities across discipline-specific course units and academic disciplines in which courses are offered that are not covered in this study. They relate to the other ten competency areas in the accounting syllabus. Second, a study can investigate ChatGPT capabilities for long computational questions with varying complexity to understand the algorithmic limits. Third, this study can investigate the academic response to large-language open artificial intelligence models through the moral development of academic human agency. Fourth, the academic response to the evolution of ChatGPT in the future, whether it is an individual or collective institutional response, can become an interesting investigation. Artificial intelligence software such as Grammarly is widely used by academic institutions, and individual academic institutions selectively use several other artificial intelligence software. Fifth, it would be interesting to determine whether academia actively collaborates with large language model providers or becomes a passive responder. SDG 17 of the United Nations encourages strengthening the means to implement and revise global partnerships for sustainable development. The partnership between academia and an artificial intelligence-driven large language model can provide inclusive and equitable quality education and lifelong learning opportunities for all (UN SDG 4), which can have positive implications for reducing inequality (UN SDG 10) and increasing gender equality (UN SDG 5).

Ethical statement/approval

Ethical review and approval did not apply to this study because there were no human or animal interactions.

Funding statement (Grant support)

This study has not received any funding.

Author contributions

The author confirms being the sole contributor to this work and has approved it for publication.

Declaration of Competing Interest

The author declares no conflict of interest.

Data Availability

The data used in this study are available online at: https://doi.org/10.6084/m9.figshare.24265567

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.joimtc.2024.100213.

References


Boston University, 2022. When to cite. School of Public Health. (https://www.bu.edu/sph/students/student-services/student-resources/academic-support/communication-resources/when-to-cite/#:~:text=You%20must%20cite%20the%20source%2C,


