African Journal of Inter/Multidisciplinary Studies 2024 **|** *6(S3): 1-11* **|** *DOI: https://doi.org/10.51415/ajims.v6i1.1528*

RESEARCH ARTICLE:

Interrogating Assessment in the Age of Generative AI

Margaret Blackie[1], Sioux McKenna[2], Neil Kramm[3] and Nicola Pallitt[4]

Received: 21 June 2024 **| Revised:** 30 October 2024 **| Published:** 05 December 2024

Reviewing Editor: Dr. Charlie Reis, Xi'an Jiaotong-Liverpool University

Abstract

Generative AI (GenAI) has foregrounded important educational issues. In this conceptual paper we build on Blackie's work on knowledge-building and assessment practices to argue that responding to GenAI in educationally sound and sustainable ways requires engagement with the purpose of a higher education. A university education should transform students' understanding of the world, themselves, and their relationship to the world in discipline-specific ways. We need to understand what it means to be gain expertise in a field before we can consider what it means to assess competence in that field. In this article, we draw on the discipline of Chemistry to reflect on the nature of the target knowledge and knowing and how this then aligns to the approach to assessment. It is only then that we can consider how GenAI might positively or negatively affect the development and demonstration of competence. In addition, because education is inherently relational, we also have to endeavour to nurture and assess such competence on a consciously created foundation of trust.

Keywords: *generative AI; assessment; trust; knowledge-building; critical AI literacy*

Introduction

Generative Artificial Intelligence (GenAI) has foregrounded important issues associated with assessment. In this paper, we argue that the initial knee-jerk responses in late 2022 and early 2023, indicated by such titles as 'Death of the college essay' and 'Will colleges survive the age of AI?' point to problems that predate the release of ChatGPT. Such news stories suggest that higher education learning can be reduced to the ability to create credible sounding texts and when software programs can produce these in ways that might fool those responsible for measuring student performance, there is no longer a point to higher education. Unfortunately, these concerns thrive where universities are increasingly positioned as training centres responsible for credentialling of individuals for industry (Shore and Wright, 2015). Unless we understand what a higher education is for, we are unable to respond to changes in society in meaningful ways. We will always be afloat and at the mercy of agendas external to the academic project (Kramm and McKenna, 2023).

In this paper, we suggest that understanding the nature and value of knowledge is central to the academic project, and it is central to teaching and assessing for epistemic access and with epistemic justice. If we are not clear as to the nature of what we are teaching and the kinds of knowing we expect of our students, we are unlikely to teach in ways that make such knowledge and knowing accessible to all. Furthermore, we argue that it is such understandings that should guide our responses to change, including the rise of GenAI. We offer the case of Chemistry to illustrate how our assessment decision-making needs to take the nature of the target knowledge into account.

[1]Rhodes University, mags.blackie@ru.ac.za | https://orcid.org/0000-0002-2465-4987
[2]Rhodes University, s.mckenna@ru.ac.za | https://orcid.org/0000-0002-1202-5999
[3]Rhodes University, n.kramm@ru.ac.za | https://orcid.org/0000-0002-4173-5586
[4]Rhodes University, n.pallitt@ru.ac.za | https://orcid.org/0000-0003-1362-670X

DURBAN UNIVERSITY OF TECHNOLOGY
INYUVESI YASETHEKWINI YEZOBUCHWEPHESHE
AFRICAN JOURNAL OF INTER/MULTIDISCIPLINARY STUDIES
OPEN ACCESS

Conceptions of Assessment

In a system in which higher education is often positioned as simply about credentialling, assessment is seen as the hurdle over which to jump to attain the certificate required by the job market (see Figure 1). Swiecki *et al*. (2022) refer to this as the standard assessment paradigm where a 'predefined set of items (e.g., problems or questions) is used to infer claims about students' proficiency in one or more traits. The data used for these inferences are typically sparse, and student learning may not be the focus of the assessment' (p.1). These kinds of assessment practices lend themselves to 'outsourcing'. 'Outsourcing' can take the form of cutting and pasting chunks of text from elsewhere and then using paraphrasing sites to avoid the surveillance of text-matching software through to purchasing assessments from paper mills. Such phenomena have been known for decades (Lancaster and Clarke, 2008). Most recently, outsourcing has become accessible to anyone with an internet connection and the willingness to sign up for the experiment that is exemplified by ChatGPT. If assessment is seen as an obstacle to attaining the necessary credential (Figure1), the student may perceive the outsourcing of the task as an occasional necessity when they encounter difficulties (Ahsan *et al*., 2022). The student may acknowledge the breach in academic integrity, but this is easily justified in terms of the benefits of attaining the pass (McKenna, 2022). The value proposition of higher education in this scenario is the certificate. To say that this is a very hollow construction of higher education is an understatement.

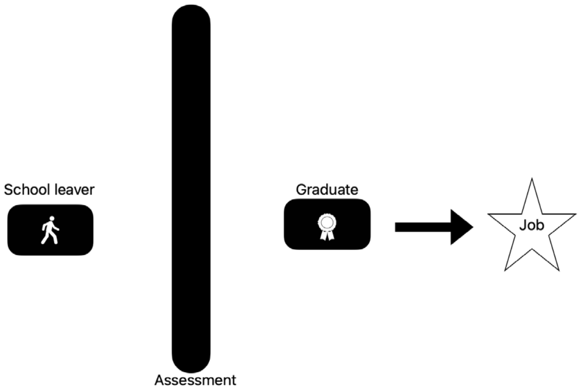

Figure 1: Assessment as an obstacle which needs to be navigated

Ashwin (2020) argues for the transformational value of higher education which comes from engagement with disciplinary knowledge. And indeed, in a seven-year longitudinal study following chemistry and chemical engineering students through their degrees and on into their lives after graduation, Ashwin and his team have shown that 'different bodies of knowledge look different and make the world look different' (Case and Ashwin, 2024). The purpose of higher education in this understanding is about coming to know a body of knowledge from the inside (Case and Ashwin, 2024). Assessments here are about giving the student an opportunity to explore their understandings and to receive external validation of their own internal sense of becoming. In this value proposition, outsourcing the assessment task makes no sense because it undermines the opportunities that assessment provides for engagement with disciplinary knowledge.

Yucel and Blackie (2024) have developed a model for understanding the relationships between disciplinary knowledge, the nature of the knowledge, assessment and evaluative judgement required to afford the kind of transformative higher education envisioned by Ashwin (2020) (see Figure 2).

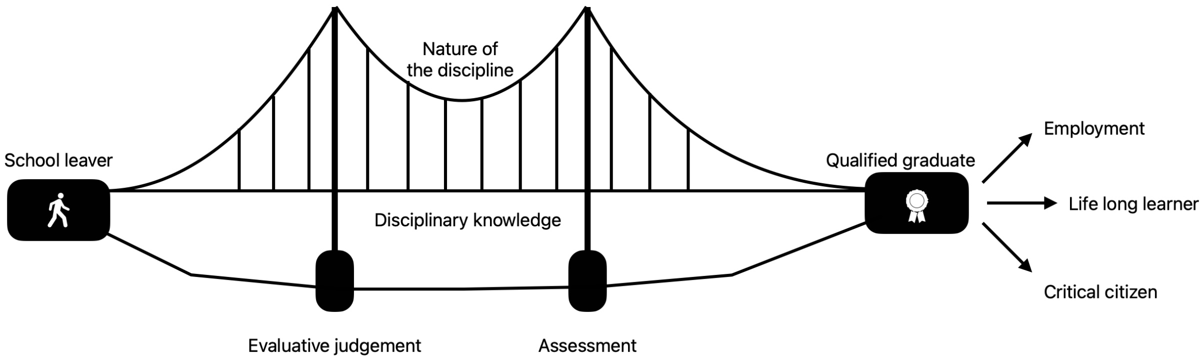

Figure 2: Model of the relational nature of assessment practices

In this model, disciplinary knowledge is the road deck of the suspension bridge. The cables are the nature of the discipline, which is necessary to develop the 'insider view' where one can judge a knowledge claim appropriately in this discipline (Maton, 2014). The two support towers are the twin pillars of evaluative judgement and assessment. To become qualified in this discipline, the student must develop an internal sense of what they know and what it means *to know* in this discipline – this is evaluative judgement (Tai *et al*., 2018). However the development of reliable evaluative judgement is hard to achieve without feedback from those who are experts in this discipline. This external feedback is achieved through appropriate assessment practices.

Significantly, this model shows that assessment practices must be aligned to the knowledge structure of the discipline. Blackie (2022) has shown that a clear understanding of the knowledge structure of a discipline (in that case, Chemistry) is necessary to ensure that students can use performance on an assessment as a reliable indicator of their grasp of the discipline. Blackie also showed that, in the case of organic chemistry, the ability to deal with procedures was problematically taken as a proxy for understanding the underlying principle. For example, the ability to balance a chemical equation was taken as a proxy for understanding the principle of the conservation of mass. To make the knowledge structure more visible Blackie developed the 'epistemic assessment framework'. This aided the development of more appropriate assessment tasks.

Epistemic Assessment Framework

The epistemic assessment framework was developed for use in chemistry (Blackie, 2022) but it has the potential for adaptation for other disciplines (see Table 1). We have here augmented the original chemistry epistemic assessment framework with suggested 'restrictions' for associated assessment tasks The 'taxonomy of restrictions' suggests that there are three kinds of resources which can be independently restricted – people, tools and information (Dawson *et al*., 2023). There is a spectrum of restrictions which ranges from completely closed (unavailable) to completely open (available).

Table 1: Epistemic assessment framework with the suggested restrictions for assessment tasks

Category	**Kind of knowledge**	**Example of a question in chemistry**	**Examples of restrictions which apply in chemistry**
Vocabulary	Knowing the fact	Information which must be learned	Information – no information sources allowed Tools – no tools allowed People – no collaboration
Simple procedure	Knowing how	Give the product of a reaction	Information – Periodic table Tools – calculator People – no collaboration
Complex procedure	Knowing how	Multistep synthetic procedure	Information – Periodic table Tools – calculator People – no collaboration
Principle	Knowing why	Knowledge applied to explain a new scenario with known results	Information – Periodic table Tools – calculator People – no collaboration
New problem	Powerful knowledge	Use of knowledge to solve a problem not previously encountered	Information – Periodic table, textbook, other texts and multimedia Tools – calculator People – classmates

The first point evident from the epistemic assessment framework is that different categories of knowledge require different kinds of assessment. For example, the ability to define specialist vocabulary may best be tested in a closed book, invigilated setting. However, testing the capacity to solve a new problem is far better suited to an open environment where engagement with different resources is possible, as would be the case in the workplace or other social setting. Given that the graduate will almost exclusively use knowledge in a completely open environment, exposing them to a variety of restrictions, including situations with none, seems vital. Indeed, students who are not offered opportunities to engage in open environments as part of their assessment process might suffer in their development of the necessary evaluative judgement in this regard.

Powerful knowledge is knowledge that can be abstracted from its initial context to be put work to make sense of the unthinkable or not-yet-thought (Wheelahan, 2007). Importantly, we need to distinguish between 'knowledge of the powerful' which acquires its power through status and colonial legacies, and 'powerful knowledge' (Young, 2009b). Drawing on the work of Muller (2024: 7), we argue that the term powerful knowledge 'is a standard bearer not only for the 'knowledge turn' but more specifically for the knowledge-centric position at the heart of it'.

Illustrating the use of the epistemic assessment framework with chemistry, Blackie (2022) argues that one can assist students in engaging meaningfully with disciplinary knowledge through a range of assessment tasks. This includes introducing a 'vocabulary test' with a minimum pass mark of 80%. Blackie (2022) argues that this ensures that students have the necessary foundation to make sense of content in lectures and that this assists in lowering the cognitive burden in high-stakes assessments such as closed-book exams. Chemistry has a strongly hierarchical knowledge structure in Bernsteinian terms (Bernstein, 1996; Blackie, 2022). This means that there are several fundamental underlying principles. It is inconceivable that one could claim 'mastery' of chemistry without a grasp of these underlying principles. Blackie (2022) illustrates, through a retrospective analysis of organic chemistry exams, that an entire paper can problematically be made up of simple and complex procedures. Thus, one can infer that generations of students were graduating with degrees in chemistry without any guarantee that those students had grasped the 'insider view' of chemistry (Case and Ashwin, 2024). This is not to say that no students attained this insider view; it is simply that there was no way of knowing whether any students attained it. The use of the epistemic assessment framework made the problem with established assessment practices visible.

The epistemic assessment framework cannot be transferred from one discipline to another without adaptation. To make the adaptations, it is necessary to have a clear understanding of the architecture of the knowledge. There are many ways to conceive this; here, we draw on Bernstein's (1996) notions of hierarchical and horizontal knowledge structures. Hierarchical knowledge structures are identified by their use of foundational principles. For example, it is inconceivable that one could gain a degree in chemistry without a strong grasp of the Periodic Table. Different branches of chemistry continue to draw on this foundational knowledge even though each branch may also have distinct second-tier principles which are necessary for that specialism. Horizontal knowledge structures operate slightly differently. The object of investigation is likely to be agreed upon (although the boundaries may vary). Sociology has society as its object of study, whereas psychology focuses on the individual. Social psychology could share concepts with both psychology and sociology. In these disciplines, there is usually a larger suite of principles that can be drawn on. So, not all sociology degrees will draw on the same principles. Still, when one looks at sociology taught in several different departments, there is likely to be an overlapping common core (see Figure 3). This means that even horizontal knowledge structures have core concepts that should be identifiable within the teaching of a specific course. In addition, how one demonstrates grasping the key concepts may vary, so how one labels the categories of knowledge may require adaptation for intuitive use.

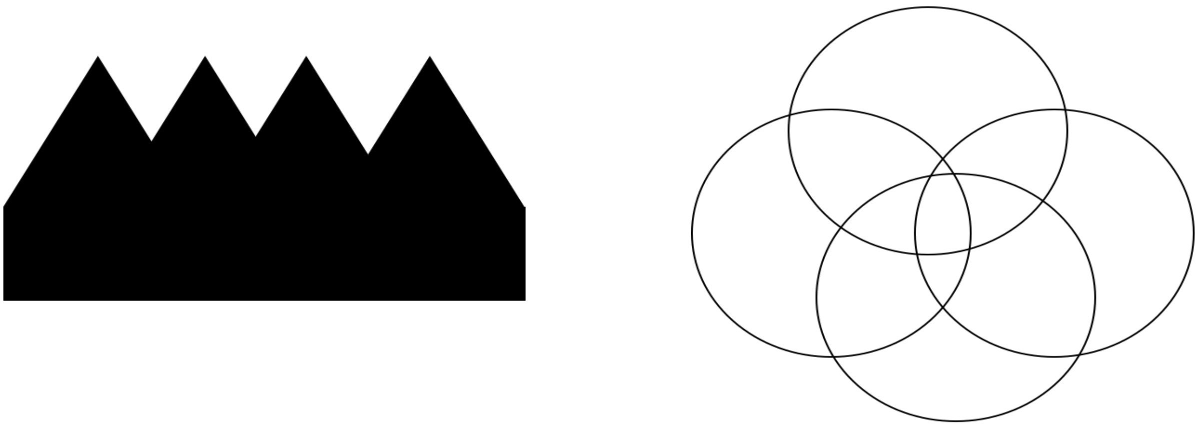

Figure 3: Representations of Bernstein's concept of hierarchical and horizontal knowledge structures

On the left of Figure 3 is a typical hierarchical knowledge structure, such as chemistry, where different specialisations build on common foundational concepts. On the right is a representation of a horizontal knowledge structure, such as sociology, where there is an overlapping central core of fundamental concepts, but different branches of sociology may draw on other concepts that are not directly linked (Schwemmer and Wieczorek, 2019). In the case of sociology, the student needs to take on the core of fundamental concepts but then engage with multiple, sometimes oppositional theories and approaches that vary across different schools of thought (Schwemmer and Wieczorek, 2019). Importantly, regardless of the school, the sociology student needs to cultivate a very particular critical gaze if they are to succeed (Van Krieken *et al*., 2020). Assessment needs to enable opportunities to develop evaluative judgement in this regard.

Thus far, the discussion on assessment practices has been restricted to the need for assessment to be aligned to the nature of the target disciplinary knowledge. However, this is insufficient to ensure that a student gains the ‘insider’ view of the discipline. To this end, they must also understand how knowledge claims may be legitimately made within this discipline. For example, how one draws from sources in chemistry and sociology are different. In chemistry, the emphasis is on mapping out the prior research on the issue on which the contribution will build. In sociology, the emphasis is on positioning the study within a school of thought or theory and justifying the methods in alignment to this.

In these post-truth times of ‘fake news’ having an understanding of the foundation of a knowledge claim i.e. the nature of the discipline, is increasingly important. Understanding what sources of information can be regarded as reliable and understanding how data can be interpreted within a discipline is a necessary part of what it is to be a critical citizen. Matthews (2007), a strong proponent of teaching the nature of science, refers to this as ‘critical scientific literacy’. As GenAI becomes pervasive, this capacity to judge the quality of work is becoming more and more important (Bearman *et al*., 2024) and links to the need for evaluative judgement and critical AI literacies.

Development of Evaluative Judgement

The development of evaluative judgement has been largely overlooked when considering teaching, learning and assessment practices in higher education. Bearman *et al*. (2024) argue that the development of evaluative judgement is essential ‘if humans are not to give up their position as arbiters of quality’. Tai *et al*. (2018) state that the heart of evaluative judgement is ‘coming to know what “good” looks like’ and thus constitutes both an internal understanding of what should count as quality and an understanding of one’s contribution. We concur, and add, that this is discipline-specific. We also add that this must include the nature of a knowledge claim within a field. Whilst we agree with Sadler (1989) that expert evaluative judgement is tacit, holistic and cannot be reduced to its parts, this does not mean that explicitly recognising some of the components will not aid a person in developing evaluative judgement. Without making this visible, we are unlikely to achieve, the much-desired, epistemic access (Morrow, 2009). Furthermore, unless we make this visible, it remains beyond critique. Making the discipline-specific terms of evaluative judgement explicit is thus also crucial if we are to achieve epistemic justice (Fricker, 2013). Hence it is valuable to the student to make visible the kinds of knowledge which work together to develop the disciplinary knowledge as described above.

There are several ways in which evaluative judgement can be developed to ensure that students engage at the kinds of levels expected by that field. Tai *et al*. (2018) suggest five pertaining to assessment – self-assessment, peer assessment, feedback, rubrics, and exemplars. Bearman *et al*. (2024) build on this work, showing how these different kinds of activities can be aided by GenAI. They illustrate how GenAI can be used in three ways for each of the five avenues (see Figure 4 below). Firstly, assessing GenAI ouputs, by evaluating their usefulness, completeness, and inclusion of hallucinations and so on. Secondly, assessing GenAI processes, including a basic understanding how GenAI works, and thirdly, GenAI assessment of student evaluative judgement. Bearman *et al*. (2024) argue that encouraging students to use GenAI in different ways helps them develop their evaluative judgement whilst harnessing the power of GenAI in meaningful ways to build confidence and depth of understanding rather than bypassing engagement with knowledge. Training students how to engage with AI-generated material in such a meaningful way will assist in developing critical AI literacies which will be increasingly necessary in the decades to come.

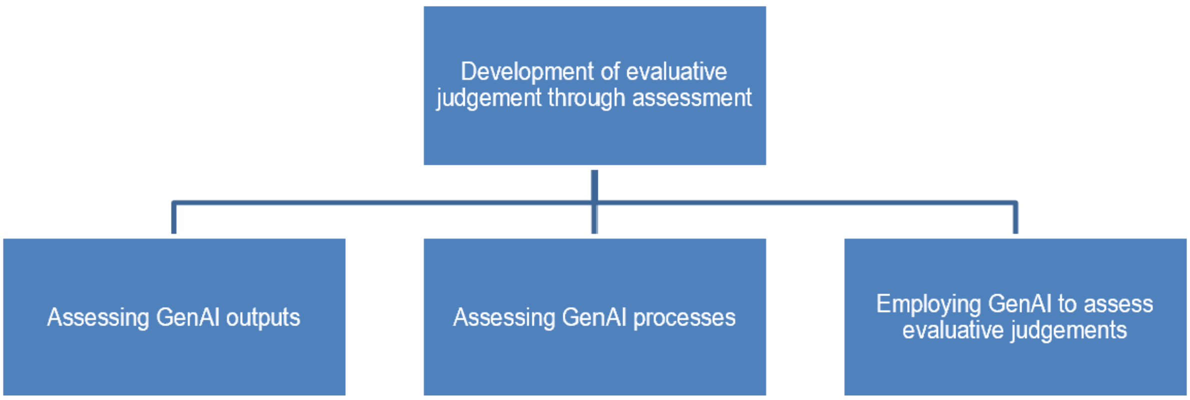

Figure 4: Three ways in which to engage students with GenAI (adapated from Bearman *et al*., 2024)

Who is Assessing?

In the knee-jerk reaction to the emergence of GenAI, many have called for the return to the invigilated closed-book examination (Kumar *et al*., 2024). However a multimodal assessment strategy is a far better response. As we have shown above, the knowledge structure itself suggests that more than one kind of assessment with a variety of levels of restriction on the use of people, tools, and information is an essential part of the development of reliable evaluative judgement and an understanding of both the core concepts in the discipline and how valid knowledge claims can be made in the discipline. There is an additional aspect to this – what is the mode of presentation and who is assessing? We need to explore more ways of reliably assessing and giving meaningful feedback to students. We have given pride of place to academic writing, but do we consider the importance of academic speech? Surely the capacity to read and write in a discipline must be augmented by the capacity to 'speak' the discipline? One of the challenges in assessing speaking is that it takes substantially longer and must be done in 'real-time'. However, strategies such as peer assessment and multimodality may be leveraged to ensure that this aspect of learning is also assessed.

Peer assessment requires clear guidelines and/or rubrics to ensure that classmates know what they are looking for. It is also important to ensure that the feedback given is appropriate; therefore, moderation of peer assessment is beneficial. This can include 'marking the marking' – giving bonus marks for peer assessors who reliably give meaningful feedback (Weaver and Esposto, 2012). Creating a system where peer assessment is one part of the assessment strategy can aid in developing students' evaluative judgement – knowing what quality looks like assists students in recognising whether they are capable of producing quality. Even students who demonstrate a strong grasp of the content may not have sufficiently developed evaluative judgement to know this. Peer assessment can then be used to widen the modalities of assessment and, at the same time, allow the active development of evaluative judgement. In many fields, it may be appropriate to consider peer-led learning as an assessment approach (Chin, 2016). The link between peer assessment and evaluative judgement has received attention in recent years. Ibarra-Sáiz *et al*. (2020) look at the links between peer assessment and evaluative judgement, feedback, self-regulation and the quality of the assessment. They show the iterative links between the quality of peer assessment, the development of evaluative judgement, and the quality of feedback given. Sridharan *et al*. (2019) note that students are more able to give accurate judgements on the quality of peers' work in formative assessments where the marks given do not count towards the final grade.

Sridharan *et al*. (2019) also report on developing a self-assessment strategy. The context of this is to get some differentiation of marks in a group assignment setting. Here, they recommend introducing two weighted factors. The 'relative performance factor asks the individuals to rate their evaluation of the relative contributions of all members of the group relative to the final group score – this is reported as a fraction, and the individual's mark is then weighted relative to the group according to this factor. (The presumption here is that all members of the group will have their marks raised by working together, so the group score is taken to be 1 and all members of the group contribute some fraction of 1). The second factor is the 'self and peer assessment factor' – this factor takes into account the difference (if any) between the self-assessment and the assessment by peers. This does not necessarily alter the mark but gives the student a good sense of their accuracy in self-assessment (Author 1 was part of a teaching team which very successfully used both of these methods for several years).

Lifelong Learning

It is important to recognise that the advent of GenAI will substantially shape the world of work. Beyond this, GenAI is now embedded in some social media platforms and web browsers. Therefore, a higher education focused only on job readiness will likely become an increasingly problematic value proposition. Today's jobs are not likely to be the jobs of tomorrow, so developing the disposition needed for lifelong learning is increasingly important. Blackie *et al*. (2023) have shown that some of the students who used the epistemic assessment framework reported that it helped them not only understand better what was required of them in the context of the organic chemistry course but that it also helped them think differently about their other subjects. Using this model as a foundation for teaching, learning and assessment practices in an undergraduate degree may well assist students in coming to understand what it is to be educated. That is to gain an 'insider view' of what it is to know in a particular discipline. It is this disposition that will best set the student up for lifelong learning. As Blackie *et al*. (2023) have shown, developing an understanding of knowledge building in one discipline can be transferable to other disciplines. With the emergence of GenAI, learning how to master a new discipline seems to be an essential component of any educational endeavour.

Harnessing Gen-AI to Assist with Developing the Insider View

As Bearman *et al*. (2024) have shown, generative AI can be used in many ways to augment learning. In the South African context Rhodes University (2024) has developed guidelines for using GenAI. Table 3 is extracted from the document and suggests ways that GenAI can be used by both students and lecturers to augment teaching and learning.

Table 3: Ways in which generative AI can be used to augment teaching and learning

Role	Description	Example of implementation
Possibility engine	AI generates alternative ways of expressing an idea	Students write queries in ChatGPT and use the regenerate response function to examine alternative responses.
Socratic opponent	AI acts as an opponent to develop an argument	Students enter prompts into ChatGPT following the structure of a conversation or debate. Teachers can ask students to use ChatGPT to prepare for discussions.
Collaboration coach	AI helps groups research and solve problems together	Working in groups, students use ChatGPT to find information to complete tasks and assignments.
Guide on the side	AI acts as a guide to navigate physical and conceptual spaces	Lecturers use ChatGPT to generate content for classes/courses (e.g., discussion questions) and advice on how to support students in learning specific concepts.
Personal tutor	AI acts as a tutor for students by providing immediate feedback on progress	ChatGPT provides personalized feedback to students based on information provided by students or lecturers (e.g., test scores).
Co-designer	AI assists throughout the design process	Lecturers ask ChatGPT for ideas about designing or updating a curriculum (e.g., rubrics for assessment) and/or to focus on specific goals (e.g., how to make the curriculum more accessible).
Exploratorium	AI provides tools to play with, explore and interpret data	Lecturers provide basic information to students who write different queries in ChatGPT to find out more. ChatGPT can be used to support language learning.
Study buddy	AI helps the student reflect on learning material	Students explain their current level of understanding to ChatGPT and ask for ways to help them study the material. ChatGPT could also help students prepare for other tasks (e.g., job interviews).
Motivator	AI offers games and challenges to extend learning	Lecturers or students ask ChatGPT for ideas about extending students' learning after providing a summary of the current level of knowledge (e.g., quizzes, exercises).
Dynamic assessor	AI provides educators with a profile of each student's current knowledge	Students interact with ChatGPT in a tutorial type dialogue and then ask ChatGPT to produce a summary of their current state of knowledge to share with their teacher/for assessment.

For students struggling to grasp the core concepts used in a discipline, GenAI can be tremendously helpful in giving simplified versions and offering suggestions for making sense of these ideas. This can create the scaffolding necessary to allow knowledge-building to occur. For students who have grasped the basics and are keen to push themselves, GenAI allows for the development of more sophisticated connections. A lecturer can therefore guide students to use GenAI in ways tailored to individual ability. This can support the development of competency in both disciplinary knowledge and the nature of the discipline. GenAI can enable the student to evaluate their own learning and develop more self-directed approaches to learning. As discipline-specific GenAI continues to be developed, there is significant potential for it to sharpen students' levels of critical appraisal. If students are encouraged to reflect on the ways in which they are using GenAI, this can also aid in the development of evaluative judgment. GenAI can also be used to give formative feedback. However, meaningful use of this mode does require reasonably skilful use and is likely to be beyond the capacity of students who are struggling academically. Thus, understanding the affordances, limitations and costs of these tools is essential to use GenAI meaningfully and ethically. For this, we need Critical AI literacies.

Critical AI Literacies

Muller (2024) argues that a knowledge-centric or powerful knowledge approach to education understands that education and knowledge are both essentially public goods, in that they are non-rivalrous and have a fundamentally shared and collaborative nature. As public goods, knowledge and education need to serve more than the individuals involved in the endeavour and seek to benefit society and the planet at large. Taking this understanding of knowledge and education means that critical AI literacy should include critical theory and critical pedagogy to develop an awareness of social justice issues and the dispositions needed to address these (Bali, 2024).

Bali (2024) identifies the elements of critical AI literacies as understanding how GenAI works, recognising inequalities and biases within GenAI, examining ethical issues, crafting effective prompts, and assessing appropriate uses of GenAI. Ng *et al*. (2021) and Bali (2024) suggest that existing terms, such as digital literacies, require expansion to be relevant to an evolving digital landscape and potential implications of AI. UNESCO defines digital literacy broadly as 'the confident and critical use of a full range of digital technologies for information, communication and basic problem-solving in all aspects of life' and that it 'is underpinned by basic skills in ICT: the use of computers to retrieve, assess, store, produce, present and exchange information, and to communicate and participate in collaborative networks via the Internet'. However, ethical considerations and an awareness of bias and inequalities are absent from this definition.

While such concepts are useful, their framings in higher education have been less so. Critical AI literacies should not be treated as generic 'skills' that are, at the same time, novel in the sense of being disconnected from other literacy practices and academic literacies (Boughey, 2022). The impact of posthumanist and new materialist thought has encouraged new ways of thinking about literacies underpinned by an ethical orientation at the heart of the materialist shift (Prinsloo and Krause, 2023). Critical AI literacies are complex assemblages that include discipline-based academic literacies. Returning to Bearman *et al*. (2024), what the human brings is the capacity to judge quality on the basis of disciplinary knowledge, the nature of the discipline, and the ethico-political implications of the use of that knowledge. We need to know about ChatGPT and other tools, understand their limitations and how the data outputs are generated. As part of using it effectively, we also need to know how to refine our prompts and that different kinds of prompts can enable different uses, from understanding concepts better and using AI tools as tutors to idea generation and writing improvement. Students who are not first-language English speakers need additional support – the language of outputs may read well and be enticing to use as is. Still, it may have incorrect information or 'botshit' (Hannigan, 2024). It is crucial for students to understand how GenAI works so that they are not swayed by the confidence of AI generated texts and graphics. There are also inequalities when it comes to using AI tools – access to paid vs unpaid versions, and abilities to design and refine prompts. Many universities have useful guidelines for students but ability to use these guidelines is impacted by these inequalities.

We need to be able to evaluate these outputs and think critically about the appropriateness and the inherent biases that are part of these responses, as well as how to use AI ethically. We need to be conscious of the limitations in its use and the threats in reliance on it in ways that constrain the likelihood of enjoying a personally transformative relationship to knowledge that is at the heart of higher education. And for this, we need trust.

The Importance of Trust

The discussion thus far relies on shared understandings of the purposes of a higher education. Where higher education is understood only in economic terms, as a transaction whereby the commodity of a qualification is sold to students through their payment of fees and completion of tasks, then our responses to GenAI will be defensive and punitive. Our responses in such a case would be to manage the threat of use of GenAI through the police-catch-punish approach used in relation to plagiarism. If however, we agree with the argument built in this article thus far, that a higher education is about enabling students to enjoy a transformative relationship with powerful knowledge that changes their understanding of the world and their role within it (Ashwin, 2020), then we need to consider how we will illuminate the role of knowledge and the need for critical AI literacies. This article has offered a deliberation of these issues and argued that if we are to truly empower our students, we need to consider teaching in ways that align with the target knowledge. Deliberations as to what forms of assessment to use need to be informed by what kind of knowledge is being assessed and what we hope our students will be able to do with that knowledge. All of this requires not only a deep concern for higher education as a public good, but also a consideration of the role of trust in the educational relationship.

In Pedagogy of the Oppressed, Freire (1970) stresses that trust is fundamental to the educational endeavour. He focuses on the role played by dialogue in building trust. Mutual respect, care, and commitment are vital to such trust-building dialogue. Because education is about taking on knowledge in critical and challenging ways, it requires that dialogue underpins our exchanges. However, such dialogue can only be meaningful if we can trust each other's intentions. Transformative education requires a willingness to engage, to try out new ideas and practices, and to make mistakes. It thus requires vulnerability. Where people are afraid and defensive, they cannot allow themselves to be vulnerable. If students are in an environment where mistakes are derided or punished, where expectations are opaque, where stakes are always high, they cannot allow the vulnerability so central to a transformative education. We need to consciously build trust into our educational spaces so that students can participate in ways that allow a transformative education to take place. This requires that we develop assessments that are not all high stakes, that nurture different kinds of knowledge, and that we provide the kind of formative feedback that can nourish the development of evaluative judgement. It requires time and a focus on the relational in the educational endeavour.

Trust is essential to cooperation. It is important for students to find academics and the university trustworthy if they are to engage meaningfully and deeply with their studies. But if students believe that they are always being measured and found wanting; are not trusted by academics and are guilty until proven innocent; are seen to be inherently lazy or corrupt, they cannot possibly trust us. The resultant lack of trust makes transformative education all but impossible. Instead, we are left with a shallow understanding of the educational endeavour as entirely transactional. The academics provide the content and monitor the students to ensure that they have not cheated in demonstrating their mastery of the content. The university then provides the commodity of a qualification. As we have argued this is the antithesis of empowering students to build their evaluative judgement and enjoy a transformational relationship to knowledge.

Conclusion

We have shown the connections between assessment practices, trust, critical AI literacies and the transformative power of a quality education. The value proposition of a higher education cannot be a certificate. It must be transformational engagement with disciplinary knowledge. Without this possibility of gaining the 'insider view' of knowledge which facilitates the shaping of a student's way of being in the world, the investment in an undergraduate education is substantially eroded. We are at a point in history where we need to ask again what it means to be human. This paper has not answered that question, but we hope that we have demonstrated that one important aspect is the relationship with knowledge. Higher education institutions have a moral obligation to sustain the possibility of the development of this relationship for the next generation. This will not happen by accident. It requires careful consideration, a foundation of trust, a deep understanding of knowledge structures and an assessment strategy which can facilitate the development of this relationship.

References

Ahsan, K., Akbar, S. and Kam, B. 2022. Contract Cheating in Higher Education: A Systematic Literature Review and Future Research Agenda. *Assessment and Evaluation in Higher Education*, 47(4): 523-539.

Ashwin, P. 2020. *Transforming University Education: A Manifesto*. London: Bloomsbury Publishing.

Bali, M. 2024. Where are the Crescents in AI? Available: https://blogs.lse.ac.uk/highereducation/2024/02/26/where-are-the-crescents-in-ai/ (Accessed 22 June 2024).

Bearman, M., Tai, J., Dawson, P., Boud, D. and Ajjawi, R. 2024. Developing Evaluative Judgement for a Time of Generative Artificial Intelligence. *Assessment and Evaluation in Higher Education,* 49(6): 893-905.

Bernstein, B. 1996. *Pedagogy, Symbolic Control and Identity: Theory, Research and Critique*. London: Taylor and Francis.

Blackie, M. A. L. 2022. Knowledge Building in Chemistry Education. *Foundations of Chemistry*, 24(1): 97-111.

Blackie, M. A., Arnott, G. and Kaschula, C. H. 2023. Engaging Organic Chemistry Students in Knowledge Building. *Journal of Chemical Education*, 100(9): 3302-3308.

Boughey, C. 2022. Not there Yet: Knowledge Building In Educational Development Ten Years on. *Teaching in*

Higher Education, 27(8): 992-1004.

Case, J. and Ashwin, P. 2024. Graduates, Knowledge, and Society: Reflections on a Seven-Year Study of the Educational Outcomes of Science and Engineering. Available: https://www.researchcghe.org/events/cghe-webinar-series-webinar-4-graduates-knowledge-and-society-reflections-on-a-seven-year-study-of-the-educational-outcomes-of-science-and-engineering/ (Accessed 22 August 2024).

Chin, P. 2016. Peer Assessment. *New Directions in the Teaching of Physical Sciences*, 3: 13-18.

Freire, P. 1970. *Pedagogy of the Oppressed*. London: Routledge.

Fricker, M. 2013. Epistemic Justice as a Condition of Political Freedom? *Synthese*, 190(7): 1317-1332.

Ibarra-Sáiz, M. S., Rodríguez-Gómez, G. and Boud, D. 2020. Developing Student Competence through Peer Assessment: The Role of Feedback, Self-Regulation and Evaluative Judgement. *Higher Education*, 80(1): 137-156.

Kramm, N. and McKenna, S. 2023. AI Amplifies the Tough Question: What is Higher Education Really for? *Teaching in Higher Education*, 28(8): 2173-2178.

Kumar, R., Eaton, S. E., Mindzak, M. and Morrison, R. 2024. Academic Integrity and Artificial Intelligence: An Overview. In: Eaton, S. E. eds. *Second Handbook of Academic Integrity*. Switzerland: Springer Nature.

Lancaster, T. and Clarke, R. 2008. The Phenomena of Contract Cheating. In: Roberts, T. ed. *Student Plagiarism in an Online World: Problems and Solutions*. New York: IGI Global, 144-159.

Maton, K. 2014. *Knowledge and Knowers: Towards a Realist Sociology of Education*. London: Routledge.

Matthews, M. 2007. Science Education and the Scientific Revolution: A Way to Learn about Science. *Review of Science, Mathematics and ICT Education*, 1(1): 49-61.

McKenna, S. 2022. Plagiarism and the Commodification of Knowledge. *Higher Education*, 84(6): 1283-1298.

Morrow, W. E. 2009. Bounds of Democracy: Epistemological Access in Higher Education. Available: https://philpapers.org/rec/MORBOD (Accessed 27 June 2024).

Muller, J. 2024. The Palimpsests of Knowledge. Available: https://www.tandfonline.com/doi/epdf/10.1080/00220272.2024.2397981?needAccess=true (Accessed 12 August 2024).

Ng, D. T. K., Leung, J. K. L., Chu, S. K. W. and Qiao, M. S. 2021. Conceptualizing AI literacy: An Exploratory Review. *Computers and Education: Artificial Intelligence*, 2: 1-11.

Prinsloo, M. and Krause-Alzaidi, L. 2023. Culture and Literacy: Situated Social Literacies. Available: https://humanities.uct.ac.za/sites/default/files/content_migration/humanities_uct_ac_za/246/files/Culture%2520and%2520literacy%2520early%2520draft.pdf (Accessed 22 June 2024).

Rhodes University. 2024. CHERTL Leads in Supporting AI in Higher Education. https://www.ru.ac.za/education/latestnews/chertlleadsinsupportingaiinhighereducation.html (Accessed 25 August 2024).

Sadler, D. R. 1989. Formative Assessment and the Design of Instructional Systems. *Instructional Science*, 18(2): 119-144.

Schwemmer, C. and Wieczorek, O. 2020. The Methodological Divide of Sociology: Evidence from Two Decades of Journal Publications. *Sociology,* 54(1): 3-21.

Shore, C. and Wright, S. 2015. Governing by Numbers: Audit Culture, Rankings and the New World Order. *Social Anthropology,* 23(1): 22-28.

Sridharan, B., Tai, J. and Boud, D. 2019. Does the Use of Summative Peer Assessment in Collaborative Group Work Inhibit Good Judgement? *Higher Education*, 77(5): 853-870.

Swiecki, Z., Khosravi, H., Chen, G., Martinez-Maldonado, R., Lodge, J. M., Milligan, S., Selwyn, N. and Gašević, D. 2022. Assessment in the Age of Artificial Intelligence. *Computers and Education: Artificial Intelligence*, 3: 1-10.

Tai, J., Ajjawi, R., Boud, D., Dawson, P. and Panadero, E. 2018. Developing Evaluative Judgement: Enabling Students to Make Decisions about the Quality of Work. *Higher Education*, 76(3): 467-481.

Van Krieken, R., Habibis, D., Smith, P., Maton, K., Martin, G., Churchill, B., West, B. and Hanson, E. 2020. *Sociology: Themes and Perspectives*. 7th ed. Sydney: Pearson Publishers.

Weaver, D. and Esposto, A. 2012. Peer Assessment as a Method of Improving Student Engagement. *Assessment and Evaluation in Higher Education*, 37(7): 805-816.

Wheelahan, L. 2007. How Competency-Based Training Locks the Working Class Out of Powerful Knowledge. *British Journal of Sociology of Education,* 28(5): 637-651.

Young, M. 2009b. What are Schools for? In: Daniels, H., Lauder, H., Porter, J. and Hartshorn, S. eds. *Knowledge, Values and Education Policy: A Critical Perspective*. London: Routledge, 10-18.

Yucel, R. and Blackie M. A. L. 2024. Constructing Assessment Practices for Knowledge Building in Science, *Teaching in Higher Education,* 1-17. 10.1080/13562517.2024.2411678