THE IMPLICATIONS OF AI FOR VOCATIONAL EDUCATION – IT'S COMPLICATED!

Bruno Balducci

OTAGO POLYTECHNIC AUCKLAND INTERNATIONAL CAMPUS

INTRODUCTION

It is barely two years since ChatGPT was launched, followed by a whole gamut of new chatbots (Rudolph et al., 2023), and already a great deal has changed, not only in the education sector but throughout industry and the world at large. How artificial intelligence (AI) will impact vocational education and training (VET), as well as job opportunities for graduates and the nature of work in the future can only be guessed at. But while specific consequences of generative and other forms of AI cannot be predicted with any degree of certainty, it is still possible for us to anticipate the conceptual shift that may come about through its rapid development, and to prepare ourselves intellectually for tackling the new challenges facing us in the coming decades.

To begin with, it is important to distinguish between the implications of the AI revolution from its actual effects on our working lives. Although we can expect things to be different in the way that education will function, it is the significance of such changes that are of interest in this article. When we talk about how teachers and students might be affected by AI today, we usually refer to "technology that (i) leverages deep learning models to (ii) generate human-like content (e.g., images, words) in response to (iii) complex and varied prompts" (Lim et al., 2023, p. 2). In education, this technology has started to be used as a learning tool, but it has also undermined the validity and reliability of assessments, albeit to a limited extent. While these issues are being dealt with, on the surface at least, through academic integrity policies and procedures, the real danger remains that the exponential growth of AI will pose a major threat to our assessment practices. The implication of this growth may very well be that assessment as we know it will become impossible. Or to put it differently, its meaning will have undergone a serious and far-reaching transformation.

The aim of this article is to explore the likely consequences of future AI on how assessment is to be defined and viewed if the current goals, outcomes and structures of vocational education remain what they are today. The key questions to be considered are:

- Who, or what, will be assessed?
- By whom?
- For what reason(s) will this assessment take place?
- Above all, what will be its perceived value in a world of work that is itself transformed by the new technology?
- In short, what would be the meaning of vocational assessment in such a world?

Giving answers to these questions may seem like science fiction. The social phenomenon of AI use, however, is now a reality. Claims are made about the detection of AI-generated material in student work, yet nobody can say how much of this material remains undetected (Weber-Wulff et al., 2023). Nor is anyone able to deny that "this is only the beginning of an era where generative models will play an instrumental role in content creation" (Martínez et al., 2024, p.2). There is a need to look at the problem of AI in educational assessment squarely and directly, without waiting for it to become overwhelming. The central problem that seems to be avoided in the literature and that should be addressed, therefore, is how the development of AI may impact the conceptual categories on which all our existing assessment practices are built.

WHAT IS ASSESSMENT?

Before turning to the implications of future AI for those who are assessed, those who assess them, and those who make use of recognised and quality-assured qualifications, what is currently meant by assessment needs to be established. The New Zealand Qualifications Authority (NZQA) focuses on "assessment for learning", that is to say "the skills and knowledge acquired by students and ākonga" in order to "provide them with accurate, fair, and valid judgments that can be used to inform their further learning" (NZQA, 2022, p.5). Meanwhile, the fuller definition of assessment given by NZQA rests on the following summary of its characteristics by the Center for Curriculum Redesign:

- A coordinated system

- Collections of demonstrated evidence

- Evidence collections supporting inferences and claims

- Claims about how well students are learning

- Claims about learners' developed and progressing competencies

- Useful and actionable feedback

- Feedback that guides individual and collective learning

- Feedback that supports integrated educational improvement

(Bialik et al., 2016, p.4)

The advantage of having such a broad and detailed definition is that it reveals how much is at stake once the origin of material submitted for assessment is in doubt. All the evidence, inferences and claims of students' learning, as well as all the feedback aimed at educational improvement, in a word an entire system is impacted. Since, according to NZQA, the judgements that are made through assessment may be useful in relation to further learning, the value of these judgements is necessarily compromised if they do not accurately reflect the skills and knowledge of the learners themselves.

WHO IS ASSESSED?

This leads to the first key concept to be analysed - that of the assessee. Once it is officially acknowledged that the work on which the judgements of assessment are based may not be entirely generated by the learner being assessed, but may also contain the work of AI, the implication is that our understanding of what constitutes an assessee has become drastically different. In accordance with the aim in this study of conceptualising future AI in relation to the vocational education structures of today, it can be observed that assessed learners in such a world would be taken as something more than separate individuals. As a system, assessment is no longer able to assess the skills and knowledge of people as independent entities. While it is obvious that students remain learners, they are changed nevertheless for the purposes of assessment into a sort of hybrid, part human and part AI. The equivalent transformation at a physical level would be to have radically-altered humans, or cyborg-athletes competing in the Olympic Games (Lopez Frias, 2016). This is not to say that such a thing can never happen; but if it does, it will be more than individual athletes that are changed as a consequence.

The strongest objection to this view must be that even the most sophisticated AI will amount to nothing more than a tool, like language and mathematics, and thus would not be assimilated to human intelligence (HI). A tool that helps us to think or create or solve problems does not, by definition, automatically become a thinker or creator or problem-solver. However, AI is different from a tool; it is more comparable to a machine that can operate itself, as well as perform tasks for an individual user. Unless the assessor is equipped with criteria that are sufficiently precise and effective, and able to monitor the conditions in which an assessment product is generated, it will become impossible to differentiate between the work of AI and HI (at least without time-consuming verification procedures). Unless we see radical changes in many different disciplines (including vocational ones), assessment processes will no longer be able to identify the evidence required for inferences, claims and feedback on an individual's own progress. It will either have to be accepted that the AI/HI distinction is no longer important in the accreditation of individual learning, or new modes of assessment will need to be devised.

A further problem arising from the advent of AI will be that, as with cyborg-athletes, the transformational nature of the new technology could have an adverse effect on the fairness of the system as a whole. While academic reflection on AI's impact in tertiary education stresses the importance of inclusiveness (Lodge et al., 2023), the acceptance of AI-enhanced performance might still create an affordability gap in relation to assessment. To put it simply, there is no guarantee that some learners will not be disadvantaged. As with all large-scale technology transformation, there is the risk of existing inequalities being reinforced and social disparity increasing through education as a result.

WHO ASSESSES?

If the concept of an assessee is no longer stable in the face of future AI, then neither is that of the assessor, whose role it is to make inferences and claims about what has been learned on the basis of the evidence available. When challenged, assessors should be able to explain how their judgements are supported by the relevant assessment rubrics. Some of this work is undoubtedly linguistic, requiring that the wording of criteria is interpreted in order to be applied, but for the most part learner performance is assessed in relation to epistemological constructs and standard practices that constitute the reality of the discipline(s) involved. The implication of AI for this process may be twofold:

- In a world where AI is set to become as ubiquitous as the internet, the assessor's initial task must be to differentiate between appropriate and illegitimate AI use (after which, all appropriate uses of the technology may be evaluated).
- The introduction of AI in assessment roles for which it can be trained undermines the function of the assessor as a reliable judge of non-symbolic, non-linguistic knowledge and skills.

To focus on the first of these, it might help to imagine there is an assessment product ready to be assessed at this moment. Whether it be written or visual, digital or physical need not be a concern in this context. It is now the assessor's own expertise in assessing whether (and how) AI has been used by the assessee that should be taken into account. To make matters worse, there are numerous factors involved in assessment design (timeframe, format, learner independence, etc) which could make such decisions extremely difficult to make. The issue may be raised as to whether in future an assessor in this situation will be capable of deciding not only whether but also how AI might have been used, a question that transcends the current debate on detection and integrity. With the likely development of AI for specialised purposes in most (if not all) employment sectors (Dwivedi et al., 2021), it can be assumed that assessors in vocational education will pass judgement on the use of AI by individual students. But will they be in a position to determine the legitimacy of this use? Will they be qualified to evaluate its effectiveness? There can be little doubt that the professional assessment of AI use must be carried out by experienced practitioners. As anticipated by Lodge et al. (2023, p.7) "There will be a widespread need to build academic and professional staff capabilities for working with AI."

Whereas the first implication is about our ability to adapt to AI, the second is about AI's ability to adapt to our practices. Training AI to assess human performance is not the same as enabling it to make informed and principled judgements in the way that assessors can be trained to do. Unless AI is to evolve into a real intelligence of its own, the best it can offer must remain a simulacrum of HI (no matter how advanced or even superior it may be) and there is a world of difference in that regard between human-like and human. The behaviour that AI can imitate is essentially of a formal nature, more or less equivalent to following protocols, solving maths problems or writing to impress. It is true, of course, there is a place in vocational education for learning routines and procedures. There are many other educational outcomes, however, especially at tertiary level, where the aim is also to assess other kinds of knowledge and skills. The behaviours through which this kind of learning is applied require new objectives in new contexts in order to be observable. Human assessors with experience of achieving similar objectives have the ability to gain an understanding of what an assessee has done, or at least attempted. AI, on the other hand, has no such understanding. All it can "observe" is the form, the appearance of things, without any way of judging the content. But the appeal of technology-assisted practices being what it is, there is now the temptation of adapting assessment to AI's capabilities. Instead of AI adapting to our ways, it could end up encouraging us to adapt our practices to what it can do, thereby devaluing the role of assessors and their work.

WHY ASSESS?

According to Meyer (2009, p.7), the purpose of assessment in tertiary education is three-fold (my italics):

"*a. feedback on learning (students and teachers)*

b. selection and progression decisions (students, teachers and institutions)

c. quality assurance and accountability (institution-wide responsibilities)."

The function of feedback, as already indicated, is to guide learning and support educational improvement, both of which are impaired if evidence of the students' knowledge and skills on which assessment is based becomes unreliable. This in turn will have a knock-on effect on progression and quality.

But as well as problems of reliability, AI raises questions about the validity of assessments, i.e. whether they are "actually measuring what they are intended to measure" (Meyer (2009, p.7). In vocational programmes, this involves the measurement of "functioning knowledge", where students are required "to exercise active control over problems and decisions in the appropriate content domains" (Biggs and Tang, 2011, pp.120-121). Now whether, first of all, a student can be said to have achieved learning outcomes by means of their own actions and intentions when they have used AI becomes a moot point. The learner's dependence on its capabilities may limit their own so profoundly as to make it debatable that they are able to have "active control". The only way for assessment to remain valid in such conditions would be to accept the human-AI

synergy of assessees. The purpose of assessment would likewise have a dual aspect: human-focused assessment would be about student learning, whereas AI-focused evaluation would be about technological efficiency.

Secondly, "appropriate content domains" in the world of future AI will, by necessity, have to include the appropriate use of AI technology in accordance with current professional practice in the relevant fields. VET would cease to be vocational if this were not the case. This could entail i) not assessing tasks that can be performed as well (or even better) by AI as they are by humans, and ii) assessing tasks that cannot be performed as well by AI (preferably with the widest possible margin). Educators already need to ensure that they do not teach their students how to perform human tasks that either have been or are in the process of being replaced by AI activity. As employers adapt their operations to take advantage of the opportunities offered by AI, or perhaps to resist the intrusion of AI into their industry, vocational programmes may have to be aligned to both approaches in order that their graduates are given the best possible chances of employment. This would have to be underpinned by a robust understanding of how AI is affecting the labour market, whether as a substitute for, complement to, or creator of new jobs (Cazzaniga et al., 2024).

Recently proposed measurements of AI exposure (Felten et al., 2021) and AI complementarity (Pizzinelli et al., 2023) could help to foster such an understanding. The first provides an index of the overlap between AI's capabilities and the tasks ordinarily performed in a given type of job. The second is an extension of the first that includes "a broad set of factors beyond tasks, related to the social and physical context in which work is performed" (Pizzinelli et al., 2023, p.5). No doubt these definitions will be further elaborated or refined as AI continues to displace, enhance and develop professional practice.

THE FUTURE ROLE OF VOCATIONAL ASSESSMENT

Having looked at the potential impact of AI on our understanding of what constitutes an assessee, an assessor and the purposes of assessment, the next point to be considered is how tertiary qualifications might be viewed in future. To recap, these are the fundamental changes to assessment that have been envisaged so far in this article:

- assessees being able to make use of AI's capabilities
- assessors having the capacity to judge how AI has been used
- assessment that is (partly) focused on AI complementarity skills

What would be the value in the workplace of education and training based on assessment practices of this kind? Employers, it must be remembered, are not uncritical of the vocational qualifications that job applicants include in their CVs. Nor are students going to be satisfied with vocational courses that do not adequately prepare them for the real world of work. For the qualifications of tomorrow to be fit for purpose, they will need to accurately reflect the ways in which human intelligence and skills may be affected by AI.

The initial requirement for any vocational qualification to gain acceptance in the labour market must be that it gives a true indication, not of AI's capabilities, but of an applicant's knowledge and skills. Hence, if the evidence, inferences and claims of assessment about someone's ability to make use of AI are not reliable, the qualification to which they relate will be deemed, at best, irrelevant. This places the onus on assessment writers to ensure that their designs are effective instruments not only for allowing the use of AI, but also for identifying and judging the application of complementarity skills.

In addition, the learning outcomes of VET programmes that such assessments would be aiming to assess must enable students to develop their technical, creative and critical skills through their own interaction with AI technology. As well as social and physical factors in occupations, the concept of AI exposure can be extended to vocational qualifications and learner competencies (Mandala Partners and Future Skills Organisation, 2023). The degree to which qualifications are exposed to AI will play a significant part in determining their relative value in the eyes of employers. It seems highly likely that, in the future, VET will be required to demonstrate appropriate AI exposure in student learning, together with valid assessment processes that are able to discriminate between different kinds of AI use by assessees.

Finally, the knowledge and skills that vocational graduates may bring to the workplace will undoubtedly be judged in the wider context of their adaptability to the "fourth industrial revolution." This socio-economic transformation that began at the turn of the century is "not only about smart and connected machines and systems. (...) Occurring simultaneously are waves of further breakthroughs in areas ranging from gene sequencing to nanotechnology, from renewables to quantum computing. It is the fusion of these technologies and their interaction across the physical, digital and biological domains that make the fourth industrial revolution fundamentally different from previous revolutions." (Schwab, 2016, p.12). The extent to which vocational education can adequately prepare students for these new work environments will be crucial to its future role. The implication for the way students will be assessed, and for learning and teaching in general, is that the use of AI in VET programmes will be combined with other major technological innovations appropriate to the different vocational fields.

CONCLUSION

This article has sought to articulate the basic concepts of assessment in education and training in light of the disruptions likely to be caused by future AI. From this perspective, it is the whole system of assessment that would appear in danger of being subverted. The autonomy of learners, competence of assessors and validity of assessments are all called into question. But the assumption throughout has been that vocational education itself would not undergo any radical alterations to mitigate and even overcome these problems. In fact, deep structural changes over time, perhaps in direct response to the issues raised by early incarnations of AI, must also be considered and anticipated for a better picture of the shape of things to come.

Consequently, to say that we need more research into the possible impact of future AI on the VET sector would be something of an understatement. From the perspective of this study, however, the following areas of enquiry stand out as important in the effort to mitigate and counteract the negative implications of the technological revolution that is already underway:

- how to ensure reliability and validity in vocational assessments
- how to develop staff capabilities in the appropriate use of AI in assessment tasks
- how to design assessments focused on AI complementarity skills

The best starting point for this research work is most probably to conduct a rigorous study of AI use in all the economic sectors that are relevant to our programmes. Reliable data is needed from which information on the use of AI by professionals can be extracted, and the specific types of knowledge and skills which they apply can be identified. It would be on the basis of such findings that secure vocational assessments could be modelled and designed, as well as professional development for academic staff and policies for the integration of AI complementarity into student programmes.

REFERENCES

Bialik, M., Martin, J., Mayo, M. & Trilling, B. (2016). Evolving Assessments for a 21st Century Education. Center for Curriculum Redesign.

Biggs, J. & Tang, C. (2011). Teaching for Quality Learning at University. Society for Research into Higher Education.

Cazzaniga, M., Jaumotte, F., Li, L., Melina, G., Panton, A.J., Pizzinelli, C., Rockall, E. & Tavares,M.M. (2024) Gen-AI: Artificial Intelligence and the Future of Work. IMF Staff Discussion Note.

Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., Duan, Y., Dwivedi, R., Edwards, J., Eirug, A., Galanos, V., Ilavarasan, P. V., Janssen, M., Jones, P., Kar, A. K., Kizgin, H., Kronemann, B., Lal, B., Lucini, B., & Medaglia, R. (2021). Artificial Intelligence (AI): Multidisciplinary Perspectives on Emerging challenges, opportunities, and Agenda for research, Practice and Policy. *International Journal of Information Management*. 57, 101994. https://doi.org/10.1016/j.ijinfomgt.2019.08.002

Felten, E., M. Raj, and R. Seamans (2021). Occupational, industry, and geographic exposure to artificial intelligence: A novel dataset and its potential uses. *Strategic Management Journal*. *42*(12), 2195–2217. https://doi.org/10.1002/smj.3286

Lim, W. M., Gunasekara, A., Pallant, J. L., Pallant, J. I. & Pechenkina, E. (2023). Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators. *International Journal of Management Education. 21*(2). https://doi.org/10.1016/j.ijme.2023.100790

Lodge, J. M., Howard, S., Bearman, M., Dawson, P, & Associates. (2023). Assessment reform for the age of Artificial Intelligence. Tertiary Education Quality and Standards Agency.

Lopez Frias, F. J. (2016). The defining components of the cyborg: cyborg-athletes, fictional or real? *Sport, Ethics and Philosophy. 10*(1), 97–111. https://doi.org/10.1080/17511321.2016.1171249

Mandala Partners & Future Skills Organisation. (2023) Impact of generative AI on skills in the workplace. Future Skills Organisation.

Martínez, G., Watson, L., Reviriego, P., Hernández, J.A., Juarez, M., Sarkar, R. (2024). Towards Understanding the Interplay of Generative Artificial Intelligence and the Internet. In: Cuzzolin, F., Sultana, M. (eds) Epistemic Uncertainty in Artificial Intelligence. Epi UAI 2023. *Lecture Notes in Computer Science*, vol 14523. Springer, Cham. https://doi.org/10.1007/978-3-031-57963-9_5

Meyer, L. (2009). Developing assessment policy: A guide for tertiary institutions. Victoria University of Wellington.

NZQA. (2022). Aromatawai and the Principles of Assessment.

Pizzinelli, C., Panton, A. J., Tavares, M. M. M., Cazzaniga, M., & Li, L. (2023). Labor market exposure to AI: Cross-country differences and distributional implications. International Monetary Fund.

Rudolph, J., Tan, S. & Tan, S. (2023). War of the chatbots: Bard, Bing Chat, ChatGPT, Ernie and beyond. The new AI gold rush and its impact on higher education. *Journal of Applied Learning and Teaching. 6*(1). https://doi.org/10.37074/jalt.2023.6.1.23

Schwab, K. (2016). The Fourth Industrial Revolution. World Economic Forum.

Weber-Wulff, D., Anohina-Naumeca, A., Bjelobaba, S., Foltýnek, T., Guerrero-Dib, J., Popoola, O., Šigut, P. & Waddington, L. (2023). Testing of detection tools for AI-generated text. *International Journal for Educational Integrity*. 19(26). https://doi.org/10.1007/s40979-023-00146-z