Routledge
Taylor & Francis Group

Innovations in Education and Teaching International

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/riie20

Relational assessment in a low-trust world

Philip Ramsey & Stefano Cataloni

To cite this article: Philip Ramsey & Stefano Cataloni (31 May 2024): Relational assessment in a low-trust world, Innovations in Education and Teaching International, DOI: 10.1080/14703297.2024.2360571

To link to this article: https://doi.org/10.1080/14703297.2024.2360571

Published online: 31 May 2024.

INNOVATIONS IN EDUCATION AND TEACHING INTERNATIONAL
https://doi.org/10.1080/14703297.2024.2360571

Routledge
Taylor & Francis Group

Relational assessment in a low-trust world

Philip Ramsey and Stefano Cataloni

School of Management, Massey University, Palmerston North, New Zealand

ABSTRACT

Prompted by concerns over student use of ChatGPT, staff teaching leadership and teamwork on a large university course experimented with an alternative way of assessing students' learning. Past assessment practices emphasised individual reflection and quality assurance. Aiming for a more relational approach that is aligned with the course's content, an oral review was employed. Staff conducted a review of learning with teams rather than individuals. Based on an in-depth staff review of the experience, this article explores the approach. The relational nature of the review was a dramatic departure from students' previous experience of assessment, which some students found disconcerting. Staff identified key lessons that can be applied to future oral review assessments. Lessons learned involved how to balance the twin goals of quality assurance and personalised learning. Staff recognise the need to explain the approach to assessment, starting early in the course.

KEYWORDS

Assessment; oral review; ChatGPT; relational; quality assurance; dilemma

Introduction

Prompted by concern over ChatGPT, staff at a New Zealand Business School reviewed the way that assessment was carried out in a leadership and teamwork course. While hard to detect plagiarism was an initial concern, we were interested in a continued effort to ensure the assessment reflects the nature of the subjects under consideration.

As professional educators we were aware that students not only learn from *what* they are taught. They are strongly influenced by social learning, including the behaviour modelled by staff. In a context where students may use AI to shortcut learning, staff may respond to the lack of trust in the university with approaches to assessment that are harsh and authoritarian. When teaching a course on collaboration, the danger exists that the behaviour and attitudes modelled during assessment are discordant with those staff espouse in class.

Educators Crawford et al. (2023) encourage higher education staff to respond to concerns over AI with character-based leadership through authentic assessment. This article explores how assessment might be made to harmonise with values associated with effective collaboration.

CONTACT Philip Ramsey P.L.Ramsey@massey.ac.nz School of Management, Massey Business School, Private Bag 11222, Palmerston North 4442, New Zealand

This article considers a key dilemma associated with assessment before discussing the role leadership plays in reconciling the dilemma. We then describe the nature of the course, and how the approach to assessment used in the past can be undermined by ChatGPT, which may also contribute to staff's increasing mistrust of students. We then discuss the experimental approach taken to assessment in the latest iteration of the course, the results achieved, and the lessons learned for the future.

An assessment dilemma

Assessment plays a complex role in higher education. In part it acts as a mechanism for quality control, enabling institutions to attest to the capability of graduates. Rigorous assessment thus enables a university to build a reputation within society, particularly with employers. Assessment is also an integral part of the learning process. Through assessment, staff can compare student work to desired learning outcomes, which forms a basis for feedback, reflection, and learning.

In regard to assessment, educators might pursue two quite different goals, which we frame as 'quality assurance' or 'personalised learning'. Quality assurance involves strengthening institutional reputation and places the emphasis on achievement of standards. It is 'objective' assessment, both in the sense that learning objectives are used as the basis for evaluating achievement, and that the assessment is not personal. When the outcomes achieved on a course are expressed as learning objectives, stakeholders, such as employers and funders, expect graduates to have attained a standard that is trustworthy. They trust that the assessment is consistent, carried out by emotionally neutral educators. In other words, the relationship that a particular student had with staff is not the basis for the grade given. The motivation for the quality assurance goal is teacher-centred and summative assessment aimed at certification of students is prioritised.

Using assessment and feedback as an integrated part of personalised learning, however, is a different goal involving different values. Learning may be strongly influenced by an emotionally engaged relationship between staff and student, with staff endeavouring to create an appreciative environment that generates psychological safety, optimism, and hope. As Harvey (1977) points out, while it sounds appealing to be assessed 'objectively', people want to be treated personally, not as objects. Hence, Zepke (2011) advises that academics incorporate activities that include both personal and relational elements which are relevant to students' lives and O'Donovan et al. (2021) suggest that students are more likely to read and find feedback more effective when it personalised.

Where learning objectives emphasise student outcomes, assessment can contribute to a learning journey where students experience the process of learning as engaging and meaningful (Martinez-Arboleda, 2021). When pursuing this goal, educators aim to be learner-centred, aiming to encourage lifelong, sustainable development (Boud, 2000). Formative assessment is prioritised to enable those involved – both students and staff – identify what future steps might be taken (Black & Wiliam, 2006).

Achieving either one of these goals is a course management challenge. Cultural experts Trompenaars and Hampden Turner (2002) make a distinction between management and leadership based on dilemmas such as this one. While managers may worry about achieving a goal, leaders are concerned that pursuit of one goal will make it more difficult to attain another that is equally important.

Because they represent markedly different values, it may not seem possible to pursue some goals in conjunction with others. According to Trompenaars and Hampden Turner (2002), leadership involves reconciling differences so that seemingly conflicting goals can be vigorously pursued without weakening the other. Sadly, rather than being reconciled, many differences become 'polarized', with people believing there must be an 'either/or' choice, rather than accommodating the neglected option.

Polarized assessment in a low-trust environment

While we hope that both goals can be attained, varied forces act to make reconciliation difficult. In particular, summative assessment, certification, and objective learning are overemphasised (Boud, 2000; Nieminen, 2020). Appraisal systems in academia have made staff feel pressured to conduct and publish research which requires time that might otherwise be spent on teaching, leading to a disconnect between what staff espouse and what they practice (Gray et al., 2017).

From a students' perspective, assessment may contribute to feeling overwhelmed and anxious (Crawford et al., 2023) at a time where they feel pressured to succeed, while in competition with classmates who they might suspect to be cheating, when the risk of being caught appears to be minimal (Perkins, 2023). A further complication occurs when changes to assessment practices are met by resistance from students, even when students are dissatisfied with the status quo (Åkerlind & Trevitt, 1999). Students may experience a dilemma when the desire to perform on assessment activities appears to conflict with their desire for learning (Ramsey et al., 2002).

Consequently, rather than adopting assessment practices that contribute to trusting relationships with students, staff often respond to concerns about assessment by implementing measures designed to identify and punish academic integrity transgressions (Moorhouse et al., 2023).

These conditions contribute to what has been described as an assessment 'arms race' (Eaton, 2022; Roe & Perkins, 2022), where students and academics are pitted against one another. On one hand, some students endeavour to use technology to shortcut assessments, while staff (Roe et al., 2023) increasingly rely on technology to detect new forms of plagiarism.

A vicious cycle can operate, where increasing pressure on students leads to more complex forms of academic dishonesty (Moorhouse et al., 2023). In turn, staff adopt more onerous forms of assessment, while students experience greater pressure. This cycle undermines both goals discussed earlier and increases the level of mistrust between students and staff as they both attempt to overcome the challenges and pressures that each party is creating. It also makes it more difficult for institutions to attest to whether students have achieved learning outcomes. Further, it weakens the student – staff relationships that may enable personalised learning, which are essential for a high-trust learning environment.

The 'arms race' term suggests that those involved in this vicious cycle may be experiencing escalating commitment, where adopting a course of action leads to subsequent decisions that ensures the decision maker continues in the same course (Bazerman, 2014). When decision makers (staff) notice they are part of a vicious cycle

involving escalating commitment, Bazerman advises them to pause, think rationally about the situation, and make sure that further action makes sense. This was the situation that staff at Massey University in New Zealand found themselves in when contemplating assessment changes for their undergraduate Leadership and Teamwork course.

The leadership and teamwork course

Fundamentals of Leadership and Teamwork is a required course for undergraduate students at Massey University's College of Business. The College recognises that graduates need to be capable of collaborating with others, within leadership roles and as part of a team.

The undergraduate course includes a large Distance cohort, and students range from recent school leavers to experienced business practitioners. For the course to be relevant to all students, building collaborative capability is treated as a journey. Informed by academic literature, students are encouraged to identify which next steps are relevant to them based on the collaborative challenges they face, and to experiment with and reflect on practices that have personal and practical value. This approach is in line with concepts of leadership that emphasise that leadership development is an on-going process as leaders are challenged by novel situations and dilemmas that require new practices (Koestenbaum, 2003; Trompenaars & Hampden Turner, 2002).

A key guiding principle on the course is that collaboration is relational (Rost, 1991), in that teammates, leaders and followers influence one another, shaping one another's behaviour as they work towards shared purposes. Therefore, to be effective at collaboration requires deepening one's understanding of relationships and involves building credibility with others (Kouzes & Posner, 1993). This relational approach requires self-awareness (Drucker, 1999), authenticity (Koestenbaum, 2002), character (Crossan et al., 2013), and capability at treating others with dignity (Hicks & Waddock, 2015).

These broad goals are supported by specific practices that students can use in collaborative work, such as reflection (Kiersch & Gullekson, 2021), questioning techniques (Adams, 2022), balancing task and social interactions (Levi, 2014), and receiving feedback (Stone & Heen, 2014). The intention of staff is to encourage students to (1) identify specific challenges, relevant to their collaborative efforts, (2) use leadership and teamwork concepts to make sense of these challenges, (3) identify practices that could make a positive difference for the future, and (4) experiment with these practices in team-based activities.

Staff are aware that the desired behaviour described above is different to what students may have practiced in other courses. Students' previous experience with teamwork at university will often have been in conjunction with assessments on courses that do not focus on teamwork practices. Many students report that, based on experiences at school and university, they dislike group-based assessment, and expect to have negative experiences with teammates. Capable students expect to shoulder extra work to compensate for the limitations of their teammates.

In our experience, many students follow an approach that gets the assessment completed while avoiding conflict with classmates. They will often divide the assessment tasks up so that individual tasks are completed before finally compiling their finished piece of work. The unintended result of this behaviour is that students avoid learning skills that

help them navigate collaborative challenges. On some courses grading practices reinforce this behaviour by evaluating individual contributions rather than the team's output.

If the validity of assessment processes is to be judged on the basis of consequences (Messick, 1989), these assessment-related behaviours are poorly aligned with the purpose of the Leadership and Teamwork course. Individualised approaches would be appropriate if the course aimed to teach a defined amount of knowledge and skill, expressed as learning objectives. Because the course aims to encourage students to improve their collaborative skills to meet emerging relationship challenges, a different approach to assessment is needed. Further, collaboration is undermined when leaders put others in a 'double bind' by espousing one way of acting while indirectly encouraging its opposite (Kets de Vries, 2011). It is inappropriate, therefore, for staff to explicitly advocate for relationships of trust while designing processes that assume students are untrustworthy.

The leadership and teamwork course aims to disrupt the individualised approach to collaboration in several ways. Teams are formed early in the course and the initial team-based activities are not directly connected to assessment. Students are assigned to teams randomly, so they have to deliberately and consciously build relationships. The difference between the individualised behaviour and the desired team behaviour is discussed explicitly, and students are repeatedly encouraged to experiment with collaborative practices. Finally, assessment is continually reviewed to determine whether it supports the behaviour we intend.

Revising assessment

Assessment for the course combines individual work with a team-based assignment. The first assignment is an individual written reflection on a student leadership experience, identifying ways they might develop future leadership capability. While working on the first assignment, students are already in teams, discussing leadership concepts together. For the second assignment, students work in teams to complete a written project, for which they receive a shared grade. During grading, staff consider both the content of the project and indications that the team has produced an integrated piece of work.

Prior to the recent revision, the final assignment was an individual written reflection on students' contribution to teamwork, basing it on the experience with the second assignment. The final assessment allowed students to 'recover' from problems that might have been experienced with the team-based work. For example, a student may have found themselves in a dysfunctional team that received a poor grade for their work together. Problems with the team, however, had a silver lining, in that it gave students more in the way of rich experiences that could be the basis for reflection and learning.

This approach allowed staff, to some degree, to reconcile the need for individual and team-based assessment. Reconciling the need to be both objective and relational with assessment became increasingly difficult. As discussed earlier, staff want to see students making progress on their own developmental journey. For many students, this is the first time they have encountered assignments based on personal reflection. Ideally, reflection involves using relevant concepts from academic literature to shed light on personal experiences. Some students discuss their experiences with little reference to literature. Others treat reflections like the essays they encountered on other courses, discussing concepts with little application to personal

experience. The idea of learning as a personal journey is also new and can be difficult for students. Rather than viewing challenges and mistakes as opportunities for learning, many have a 'fixed' mindset (Dweck, 2006) that makes them reluctant to acknowledge failure and are inclined to describe their personal experiences in heroic terms.

While staff want to give students the benefit of the doubt where possible, they want to uphold collaborative values such as credibility and self-awareness. When student reflections suggest that team experiences were trouble-free, staff naturally suspect that this is not a true reflection of actual team dynamics. It may be due to a lack of self-awareness on the part of the student, fear that they will be penalised for describing what really happened, or perhaps that they were never fully engaged in the team processes. High performing students may also tend to underplay problems caused by classmates who make minimal contributions to projects. Staff have observed that international students seem particularly reluctant to alert assessors to any kind of dysfunction in their teams.

The emergence of ChatGPT prompted staff to revise the way assessment was conducted. While it may be more difficult for AI to produce work that sounds authentically personal and reflective, staff already made allowance for the many students who struggle to find the right tone. Given the relatively large size of the course, much of the grading was done by inexperienced PhD students, who are less able to detect work produced through AI.

Given the growing mistrust between staff and students, largely generated by the proliferation of AI use, a default response could have been to replace the written assessment with an examination. While this option is relatively straight forward, it would have resulted in a misalignment between the type of assessment and the relational nature of the course and its learning objectives. Moreover, an examination may have placed further strain on the student-teacher relationship and eroded trust even further.

As for the balance between objective and relational assessment, in this context some students will take advantage of the concern staff have for relationships and the willingness to give students the benefit of the doubt unless there is clear evidence of a breach of academic integrity. The challenge facing staff, however, is how to strengthen the objectivity of assessment without weakening the relational values of the course. Staff did not want to undermine assessment and learning for those who were committed to a learning journey on the assumption that some of their classmates were not trustworthy.

In response, staff revised the final assignment, the teamwork reflection. Rather than a written assignment, staff conducted an 'oral review' of teamwork with each team. Teams met with staff for 12 to 15 minutes and answered questions related to their team processes and individual contributions. Teams either met with staff in person or on Zoom. Students had the review's purpose and process explained to them and had the opportunity to prepare both individually and as a team.

Research on oral assessments has generally been limited to what has taken place in second language acquisition courses. Joughin (1998) posits that oral assessments are more authentic than most other assessments and are generally better at preventing plagiarism. They are also effective at assessing student understanding of key concepts while facilitating critical thinking (Gent et al., 1999). In line with the course's learning objectives, oral assessments are more appropriate for assessing emotive and personal issues, along with reflective practice (Huxham et al., 2012).

Literature on group oral assessments is limited, other than highlighting concerns about potential problems. For example, a talkative candidate may feel restrained by a group of reticent members, while a quiet candidate may be even quieter in the presence of more sociable members (Van Moere, 2006). Relatedly, Ockey's (2009) on oral group assessments found that more assertive candidates were awarded higher grades than less assertive candidates.

Design and implementation of the oral review

The desire for the oral review assessment to be both relational and objective, while avoiding potential problems identified in literature, shaped the design of the process and how it was implemented. This included the design of the marking rubric and staff preparation.

The marking rubric focused on three broad areas. Firstly, *the extent to which the student gave evidence of self-awareness and practice orientation*. We were interested in whether students could give clear, insightful descriptions of what they had learned and the changes this would mean for future teamwork practice. Further, could they show that they had documented their reflections and learning. Next, *the ability of the student to link their learning to teamwork theory*. Could they give a clear explanation of how their insights were related to scholarly literature? Finally, *the extent to which student contributed to the assessment discussion*. Was it evident that students were engaged in the conversation, such as taking the initiative to contribute? Did they demonstrate supportive behaviour, such as listening when teammates were talking?

Staff met before the oral assessment took place, to agree on how the oral reviews should run and how grading should be processed and assigned. In line with Åkerlind and Trevitt's (1999) advice to proactively manage potential student resistance to an unorthodox assessment, staff were encouraged to assume that each student started with a C-grade, and then note when individuals said or did something to warrant a higher or lower grade. Since this was a new and experimental assessment for everyone involved, taking this approach provided staff with a benchmark to refer to and a clearer understanding of what was expected to achieve a passing grade. With multiple staff in different locations and across different course offerings, this benchmark may have contributed to higher levels of consistency and confidence in the assigned grades.

With most teams, a pair of staff members conducted the interview. One would take the lead asking questions while the other took notes, assessing students against the rubric. Questions were initially directed to the team as a whole, asking about their team process. They included questions like, 'What were some of the key challenges you faced and how did you adapt your process to meet the challenge?' and 'Tell us about any specific concepts from the course you decided to employ in your teamwork'.

The review also involved asking individuals what they had learned. Questions included, 'What concepts did you personally find valuable?', 'What will you use in the future?', and 'Was there anyone in the team who contributed something you really appreciated?'

At the end of the oral review staff would thank the team, mention any specific strengths that were evident, and encouraged them to keep practicing skills they had

mentioned. When the team had left the staff member taking notes would summarise judgements made about grades. Staff would compare impressions, finalise a grade and make notes for personalised written feedback.

Results

The oral review process ran as designed. There were significant savings of staff time and costs to the university because of the difference in time taken for the review as opposed to marking individual written reflections. The class mean mark for the assessment was very similar to that in the written assessment used previously. Requests for extensions are common for written assignments, so it was notable that all students attended their designated times for oral review. This suggests that the process also involved a reduced assignment load for students at a point in the semester when all courses are demanding assessed work.

Staff involved in the oral review met online afterwards to evaluate the effectiveness of the process and to consider whether it was an appropriate fit with the aims and content of the course and identify where improvements could be made. Because this process reviewed personal practice, it was not necessary to apply for ethics approval. The authors report that there are no competing interests to declare.

Eight of the nine staff involved in the assessment took part in the critique. Five were full time experienced academics, two were experienced tutors, and one was a PhD student with minimal tutor experience. Those present commented on their experiences using the PMI technique (De Bono, 2007), where decisions are viewed in terms of plus, minus, and interesting aspects, to explore a topic fully, without being trapped into a fixed position. They recorded important conclusions on a Miro Board.

There was general agreement that the process was attractive in terms of staff workload. Not only was it efficient in terms of the time taken per student assessment, the process was energised and engaging for both students and staff. For staff, this was a refreshing contrast to grading large numbers of written assignments. Working in pairs enabled 'real-time' moderation of grades, which also reduced staff workload.

Staff viewed the assessment process as valid; even though judgements regarding grades were made quickly, they were based on observable evidence of student competence. Mismatches between what students said about what they had learned and how they behaved towards others in the team were obvious, which otherwise may not have been as noticeable in a written reflection. Staff could see the dynamics of teams and differences between teams firsthand, rather than relying on written descriptions. Comments included:

> *The oral reflection seemed to consolidate what they learned and how they could develop within a team setting in a more authentic way than what a written reflection could ever do.*
> *No need for AI checking here.*
> *[Students were] anxious if they hadn't internalized concepts. It was harder to fake it than with a written assessment.*
> *I liked the real sense of how some teams had built relationships and learned to navigate difficulties etc.*

The process was endorsed as generally effective, particularly in terms of the mix of general questions about team process and personal questions. All teams had been given opportunity to prepare together for the review, and it was evident that some had taken advantage of preparation time and others had not. Well prepared teams often began the review in a 'scripted' way that was somewhat artificial but allowed them to demonstrate their ability to support one another while working together. Staff appreciated that effort and, at the same time, adjusted their questions to elicit more spontaneous answers from students. Staff also adjusted questions to draw out students who had difficulty expressing themselves, or whose answers were similar to those given by teammates.

The combination of prepared and spontaneous answers provided valuable insights into student learning. Comments included:

> Scriptedness at the start showed that they had prepared.
> Sometimes the group doesn't feel 'real', but once asked a real question then the masks came off. Sometimes there was a robotic answer, then the real person was more evident.
> One useful question was, 'What surprised you?' This made teams laugh, made them relaxed and thoughtful.
> I liked the way that there were spontaneous outpourings of appreciation for each other at the end of the review!

A 'minus' issue that several staff pointed out was to do with timing and group sizes. Larger groups – with five or six members – did not have as much speaking time per person. Staff at times felt rushed while trying to ensure each member could contribute to the conversation. Having teams talk about the general lessons learned first and following with the personal lessons meant that time pressure was felt around the discussion of individual contributions.

Staff decided that future iterations of the oral review could be extended by around five minutes, to ensure that reticent members have adequate time and opportunities to contribute to the discussion and encourage greater participation. Further, once the team leaves the session, setting an additional five-minute window for staff to discuss and agree on individual grades would reduce the rush to process the grade.

Other suggestions for improvement related to the mix of questions asked and ways to manage the balance of scripted and spontaneous answers. Staff agreed that having an extensive range of questions should address this issue and including more complex questions. This strategy extends high achieving students while providing a stronger basis for grade differentiation.

The idea to change the assessment emerged once the semester had begun and was 'fleshed out' further as the semester progressed. The students who did best in the oral review were those who had internalised leadership and teamwork concepts by incorporating them into course activities throughout the semester. Therefore, an additional key area for improvement was to ensure that in future the assessment is explained early in the semester with repeated reminders to encourage students to engage in a real learning process.

Increasing assessment clarity through explicit explanations early in the semester has a dual purpose: to increase student preparedness and reduce possible resistance to pedagogical innovation. These explanations involve explaining the purpose and value (Cheng et al., 2013) of the assessment, while also increasing student awareness by way of

training (Åkerlind & Trevitt, 1999) in workshops. Spending more time focusing on the assessment's structure and requirements should also make students feel more comfortable with the assessment. In turn, these measures may enhance student-teacher trust and reduce adversarial relationships. Finally, student perception of objectivity can be improved by strengthening and explaining the criteria used by staff during the oral review.

Discussion

As an initial effort to reconcile the interdependent goals of being both relational and objective, the oral review delivered promising results. This was a new experience for both staff and students. The staff involved were satisfied with the objective rigour of the assessment, especially given the threat to written assessments posed by AI. By engaging in the process, nine staff members were left with enhanced skills, a positive attitude towards the process, and ideas on how they can further improve similar reviews.

Staff went into the review assuming that it would be difficult for students to achieve a failing grade for the assessment. Given the assessment's experimental approach, staff were ready to err on the side of being overly relational. During the reviews, however, some students gave clear evidence of having made little or no effort to contribute to the course or to their team. In these cases, it was clear that these students did not see themselves in a relationship of any significance; a failing grade was objectively appropriate and did not harm the 'low trust' relationship that the student had chosen to be in. Staff were more confident in making this determination than they might have been in questioning the authenticity of a written assignment where cheating was suspected.

While the process was novel to students, they will presumably encounter more assessments like this in the future, involving alternatives to written work. Integrating the assessment into the course, including explanations of why we take the approach, will likely result in students having a deeper understanding of dilemmas related to learning and to teamwork.

With more exposure to oral assessments, students get to experience some of the more advanced teamwork concepts. For example, Kahane (2017) explains that teams are living systems, just as individuals within the team are living systems. Teams and individuals – and larger systems like organisations and communities – all must balance the need for collective action and personal contribution. An individual within the team must meet their own needs while simultaneously attending to the needs of the team. The oral review calls on students to give evidence of their own capability while allowing their team as a whole to perform, thus reflecting this 'real-world' tension.

A small number of students contacted the course coordinator after the oral review, questioning assigned grades. Students questioning an assessment grade is, of course, not unusual. The common theme in this instance appeared to be that students had compared the grades they had received with teammates and were concerned about equity, particularly given the limited time allowed for the review. They felt, based on grades given in other courses, that they would have received a higher grade if the assessment were in written format.

This concern can be understood as a predictable cultural dynamic that occurs when, historically, one way of acting in a dilemma has been valued more highly than its opposite

(Hampden Turner & Trompenaars, 2000). The favoured goal is taken for granted as the culturally appropriate approach. Individualistic and objective assessment may be readily accepted by students simply because it has been established as culturally appropriate behaviour (Bower, 1966) and students who have learned to succeed in this culture are comfortable with the approach. Our experience was that students valued the opportunity to discuss the reasons for the unfamiliar approach to assessment, explained in relation to the course content.

Conclusion

Assessments of can create dilemmas for staff. The rapid proliferation of AI has created a potential 'arms race' as academics continue to develop strategies to respond to the challenges these software create. Staff at Massey University's School of Management recognised that an adversarial relationship was looming, as they became concerned that some students may rely entirely on ChatGPT to complete their assessments. Not wanting to create a low-trust environment, staff replaced a written assessment with an oral review to assess both objective and relational features that reflect the course's learning objectives.

Taking an approach to assessment that is both objective and relational can be viewed as a change to the organisational culture that is familiar to both staff and students. While some students might be ready to change, others may resist innovations that they see as putting grades at risk. However, change is likely to be easier if the innovation is explained and issues of alignment between course intentions and assessment are discussed explicitly with students.

To prevent an adversarial relationship and minimise mistrust, staff need more than a desire to align assessment practices with the aims of the course. Experience with this innovation highlighted the value of proactively addressing potential student concerns early and explicitly, providing an explanation of the intended purpose and value of the oral review, along with details regarding its process and criteria.

Presumably, students will also become more comfortable with the approach as an increasing number of courses take similar approaches and they encounter a wider range of assessment strategies during their studies. With the increasing accessibility and advancements of AI raising concern over the validity of purely objective approaches to assessment, we expect that more academics will withdraw from the 'technology arms race' by experimenting with alternative assessment strategies.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributors

Phil Ramsey is a Senior Lecturer at Massey University in New Zealand. He teaches Organisational Learning, Leadership and Teamwork. He also consults with Senior Leadership Teams in schools.

Stefano Cataloni is a Lecturer and doctoral researcher at Massey University in New Zealand. He teaches in the area of Leadership and Teamwork.

ORCID

Philip Ramsey http://orcid.org/0000-0003-3609-6254
Stefano Cataloni http://orcid.org/0009-0002-2879-1038

References

Adams, M. (2022). *Change your questions, change your life* (4th ed.). Berrett-Koehler Publishers.
Åkerlind, G. S., & Trevitt, A. C. (1999). Enhancing self-directed learning through educational technology: When students resist the change. *Innovations in Education and Teaching International*, *36*(2), 96–105. https://doi.org/10.1080/1355800990360202
Bazerman, M. (2014). *The power of noticing: What the best leaders see*. Simon & Schuster.
Black, P., & Wiliam, D. (2006). Developing a theory of formative assessment. In J. Gardner (Ed.), *Assessment and learning* (pp. 206–229). Sage.
Boud, D. (2000). Sustainable assessment: Rethinking assessment for the learning society. *Studies in Continuing Education*, *22*(2), 151–167. https://doi.org/10.1080/713695728
Bower, M. (1966). *The will to manage*. McGraw-Hill.
Cheng, F. P., Wang, K. H., & Lin, I. C. (2013, May 7–9). Measuring the adoption and resistance of e-learning by students. *The International Conference on E-Technologies and Business on the Web* (EBW2013) (pp. 247–249). The Society of Digital Information and Wireless Communication, Bangkok, Thailand.
Crawford, J., Cowling, M., & Allen, K. (2023). Leadership is needed for ethical ChatGPT: Character, assessment, and learning using artificial intelligence (AI). *Journal of University Teaching & Learning Practice*, *20*(1), 1–7. https://doi.org/10.53761/1.20.01.01
Crossan, M., Mazutis, D., Seijts, G., & Gandz, J. (2013). Developing leadership character in business programs. *Academy of Management Learning & Education*, *12*(2), 285–305. https://doi.org/10.5465/amle.2011.0024a
De Bono, E. (2007). *Tactics: The art and science of success*. Profile Books.
Drucker, P. (1999). Managing oneself. *Harvard Business Review*, *77*(2), 65–74.
Dweck, C. (2006). *Mindset*. Random House.
Eaton, S. (2022). The academic integrity technological arms race and its impact on learning, teaching and assessment. *Canadian Journal of Learning and Technology*, *48*(2), 1–9. https://doi.org/10.21432/cjlt28388
Gent, I., Johnston, B., & Prosser, P. (1999). Thinking on your feet in undergraduate computer science: A constructivist approach to developing and assessing critical thinking. *Teaching in Higher Education*, *4*(4), 511–522. https://doi.org/10.1080/1356251990040407
Gray, D., Smart, K., & Bennett, M. (2017). Examining espoused and enacted values in the AACSB assurance of learning. *Journal of Education for Business*, *92*(5), 255–261. https://doi.org/10.1080/08832323.2017.1335278
Hampden Turner, C., & Trompenaars, F. (2000). *Building cross cultural competence: How to create wealth from conflicting values*. Yale University Press.
Harvey, J. (1977). Organizations as Phrog Farms. *Organizatioinal Dynamics*, *5*(4), 15–23. https://doi.org/10.1016/0090-2616(77)90037-7
Hicks, D., & Waddock, S. (2015). Dignity, wisdom, and tomorrow's ethical business leaders. *Business and Society Review*, *121*(3), 447–462. https://doi.org/10.1111/basr.12094
Huxham, M., Campbell, F., & Westwood, J. (2012). Oral versus written assessments: A test of student performance and attitudes. *Assessment & Evaluation in Higher Education*, *37*(1), 125–136. https://doi.org/10.1080/02602938.2010.515012
Joughin, G. (1998). Dimensions of oral assessment. *Assessment & Evaluation in Higher Education*, *23*(4), 367–378. https://doi.org/10.1080/0260293980230404
Kahane, A. (2017). *Collaborating with the enemy*. Berrett-Koehler Publishers.
Kets de Vries, M. (2011). *The hedgehog effect: Executive coaching and the secrets of building high performance teams*. Jossey-Bass.

Kiersch, C., & Gullekson, N. (2021). Developing character-based leadership through guided self-reflection. *The International Journal of Management Education*, *19*(3), 100573. https://doi.org/10.1016/j.ijme.2021.100573

Koestenbaum, P. (2002). *Leadership: The inner side of greatness*. Jossey-Bass.

Koestenbaum, P. (2003). *The philosophic consultant*. Jossey-Bass/Pfeiffer.

Kouzes, J., & Posner, B. (1993). *Credibility*. Jossey-Bass.

Levi, D. (2014). *Group dynamics for TEAMS* (4th ed.). SAGE Publications.

Martinez-Arboleda, A. (2021). Ipsative assessment: Measuring personal improvement. InT. Beaven & F. Rosell-Aguilar (Eds.), *Innovative language pedagogy report* (pp. 77–82)). Research-publishing.net. https://doi.org/10.14705/rpnet.2021.50.1239

Messick, S. (1989). Validity. InR. L. Linn (Ed.), *Educational measurement* (3rd ed. pp. 13–104). Macmillan.

Moorhouse, B. L., Yeo, M. A., & Wan, Y. (2023). Generative AI tools and assessment: Guidelines of the world's top-ranking universities. *Computers and Education Open*, *5*, 100151. https://doi.org/10.1016/j.caeo.2023.100151

Nieminen, J. (2020). Disrupting the power relations of grading in higher education through summative self-assessment. *Teaching in Higher Education*, *27*(7), 892–907. https://doi.org/10.1080/13562517.2020.1753687

Ockey, G. J. (2009). The effects of group members' personalities on a test taker's L2 group oral discussion test scores. *Language Testing*, *26*(2), 161–186. https://doi.org/10.1177/0265532208101005

O'Donovan, B. M., den Outer, B., Price, M., & Lloyd, A. (2021). What makes good feedback good? *Studies in Higher Education*, *46*(2), 318–329. https://doi.org/10.1080/03075079.2019.1630812

Perkins, M. (2023). Academic integrity considerations of AI large language models in the post-pandemic era. *Journal of University Teaching & Learning Practice*, *20*(2), 1–24. https://doi.org/10.53761/1.20.02.07

Ramsey, P., Franklin, P., Ramsey, D., & Wells, R. (2002). Rethinking grades for sustainable learning. *Innovations in Education and Teaching International*, *39*(2), 117–123. https://doi.org/10.1080/14703290252934568

Roe, J., & Perkins, M. (2022). What are automated paraphrasing tools and how do we address them? *International Journal for Educational Integrity*, *18*(1), 1–10. https://doi.org/10.1007/s40979-022-00109-w

Roe, J., Renandya, W. A., & Jacobs, G. M. (2023). A review of AI-powered writing tools and their implications for academic integrity in the language classroom. *Journal of English and Applied Linguistics*, *2*(1), 3. https://doi.org/10.59588/2961-3094.1035

Rost, J. (1991). *Leadership for the 21st century*. Greenwood Publishing.

Stone, D., & Heen, S. (2014). *Thanks for the feedback*. Penguin Random House.

Trompenaars, F., & Hampden Turner, C. (2002). *21 leaders for the 21st century*. McGraw-Hill.

Van Moere, A. (2006). Validity evidence in a university group oral test. *Language Testing*, *23*(4), 411–440. https://doi.org/10.1191/0265532206lt336oa

Zepke, N. (2011). Inclusive teaching: Making space for the difference. InN. Zepke, D. Nugent, & L. Leach (Eds.), *Reflection to transformation* (pp. 99–116). Auckland, NZ: Dunmore.