



Against anonymity: relational marking and awarding gaps

Cathy Elliott

University College London, UK

Abstract

Anonymous marking is often proposed as a means of tackling awarding gaps, on the basis that if markers do not know whose work they are marking, they will not be able to display bias. Nevertheless, despite the widespread use of anonymous marking in the UK higher education sector, awarding gaps and bias persist. In this article, I draw on my experience of using negotiated marking in assessment in an undergraduate class to argue that anonymity is not only an inadequate means of eradicating awarding gaps but that it is actively counterproductive to the types of strategies that might actually work. I argue that the sort of assessment literacy that promotes both high-quality student work and fair marking, is better promoted through relational marking. I define this as an approach to negotiated marking that requires and foregrounds deep, trusting relationships between students and their teachers and encourages continuous joint reflection. Anonymous marking is, I argue, not only ineffective on its own terms but also inimical to our ability to engage in the work of reducing awarding gaps. This is for two reasons. First, relational marking produces the circumstances that have been shown to reduce awarding gaps more generally, whereas anonymous marking works against them. Second, relational marking allows students to exercise power, including the ability to challenge the structures that produce awarding gaps. Whilst anonymity may (imperfectly) mask bias, only deep relationships can really allow institutions and tutors to work in partnership with students to reckon with bias and to change.

Keywords: ungrading; relational marking; anonymous marking; awarding gaps; assessment literacy; feedback.

Introduction

Negotiated marking is a practice that comes under the broad umbrella of 'ungrading'. It requires students to take on shared responsibility for arriving at a decision on their mark, under institutional circumstances in which marking must occur in order for students to progress and/or graduate. The question of awarding gaps frequently comes up early in discussions about negotiated marking; it is often raised by students. Colleagues, including my own external examiner, express similar concerns.

The reasons for these concerns are twofold but related. First is the worry that students who are disadvantaged in various ways may not be good at advocating for themselves. This is sometimes understood to be related to the erosion of confidence through experiences of discrimination. Second, negotiated marking is correctly seen as making impossible the one policy that has been widely put in place to combat awarding gaps: anonymous marking (Pitt and Winstone, 2018; Borkin, 2020; Smith and Whitworth, 2024).

There is no consensus on how awarding gaps might be tackled. However, as I will discuss, there is quantitative evidence suggesting that anonymous marking does not seem to make much difference. The aim of this article is not to provide a further such quantitative study, but rather a theoretical argument, bolstered by qualitative data from my own practice, about the pedagogies that underpin anonymous marking and why they might, paradoxically, actually make awarding gaps more likely. As an alternative approach, I propose *relational marking*, which is an approach to negotiated marking that requires deep, trusting relationships between students and their teachers and encourages continuous joint reflection. I suggest that this actively engages with the power structures that create awarding gaps and enables staff and students to work collaboratively towards recognising and challenging them.

I begin by outlining how awarding gaps affect students in the UK and discussing some of the possible causes. I then turn to the literature on anonymous marking to assess its usefulness and unpick some of the pedagogical assumptions underlying this now widespread practice. I next contrast this approach with relational marking. I start out by explaining more about my class and how I implemented a relational approach, showing how it allowed us to challenge questionable pedagogical practices associated with

anonymous marking. Finally, I offer evidence that relational marking might help us challenge awarding gaps more effectively than anonymous marking but would require us to dispense with the latter.

Awarding gaps

Awarding gaps in the UK are defined as the systematic difference between the number of 'good degrees' awarded to students of different demographics. In the UK, the awarding gap between white students and students of colour for a 'good degree' (defined as a 2.i or above) was 8.8% in 2020/21, whilst between white and, specifically, black students, it was 18.4% (Universities UK, 2022). When considering only first class degrees, the gap between white students and students of colour was 9.5% (Universities UK, 2022). My own institution, UCL, has comparatively small awarding gaps. In the same year, the overall awarding gap for 'good degrees' between white students and students of colour was 0.4%. However, for first-class degrees, the awarding gap between white students and students of colour was 4.3%, rising to 15.3% between white and black students (UCL, 2023).

The causes of awarding gaps remain poorly understood. However, they persist even after controlling for factors such as entry qualifications, parental educational background, age and degree programme (Mountford-Zimdars et al., 2015; Office for Students, 2021; UCL, 2023). Therefore, there must be something we are doing (or not doing) that leads directly to differential results between groups of students who were similar on entry.

In a special edition with a focus on freedom to learn, it might be tempting to put forward an obvious, if glib, response to this: if marks are harmful, meaningless and arbitrary anyway (Blum, 2017), then an obsession with awarding gaps misses the point. Rather than trying to equalise the competition, this argument might run, we need to critique the whole corrupt enterprise that ranks and pits students against each other. This is too easy, however. After all, it would be troubling if a 'freedom to learn' approach could not address the structural effects of discrimination. I, therefore, want to suggest that we take awarding gaps seriously as a sign that there is something wrong, without fetishising the 'good degree' – and what this measure says about the judgements passed on students who do not get one – as an

end in of itself. It is therefore worth considering what we know about where these awarding gaps come from.

The Higher Education Funding Council for England (HEFCE) commissioned an extensive review of published and grey literature, alongside stakeholder interviews and fieldwork, to produce a significant study of differential outcomes in 2015 (Mountford-Zimdars et al., 2015). The study suggested that there is not a single cause of awarding gaps, but rather they are the effect of a range of different issues, many of which we might want to address for their own sake. They usefully classify these into four overarching areas. First, they address the curriculum (including teaching, learning and assessment practices) recommending a curriculum that is engaging to all social groups and assessment practices that are clearly explained, addressing the hidden curriculum, defined as the practices that students may not know about but which support them to succeed. As an example, a study at Leicester, Birmingham City University and the University of Wolverhampton recently developed a pedagogical approach that addresses the hidden curriculum directly with considerable success on a range of awarding gaps including race, disability and first-in-family students (Campbell and Duke, 2023). Second, relationships among staff and students are key, with belonging emerging as a crucial factor. Third is the importance of students' social, cultural and economic capital. Whilst tutors have limited ability to influence students' material resources, we can support students to develop the confidence and knowledge to use university systems of support, peer networks and the collaborative approaches to learning that characterise the successful attitudes of more privileged students. Finally, psycho-social processes are highlighted, with a particular emphasis on experiences of racism and micro-aggressions both from staff and other students. This includes the ongoing prevalence of deficit models that locate the reasons for awarding gaps within the students themselves, rather than wider social, economic and cultural processes.

I will return to the question of assessment pedagogies. However, first I will address a suspicion that is widespread, particularly among students, that marker bias is also a causal factor (Borkin, 2020). In particular, one driver of the move to anonymous marking in the UK was the National Union of Students' (NUS) campaign called 'Mark My Words, Not My Name' (NUS, 2008).

Anonymous marking: assessing the 'final product'

The HEFCE report discussed above states unambiguously that '[anonymous] assessment has not been found to reduce [awarding] gaps' (Mountford-Zimdars et al., 2015, p.82). This finding is borne out in other studies. Malouff and Thorsteinsson's (2016) meta-analysis of experimental studies of 'bias' – including some studies with experimental conditions going well beyond simply mentioning a student's race or gender, such as markers being told that a particular student was less intelligent than their peer group – found a very small but statistically significant effect of bias on marking. However, this already small difference was greatly attenuated if the marker was experienced or if a rubric was used. This study gives us compelling reasons not to insult our students, employ inexperienced markers without proper support, or fail to use a rubric, but does not provide overwhelming evidence for anonymous marking in the absence of these conditions.

Meanwhile, other studies do not find evidence that anonymous marking affects awarding gaps. Hinton and Higson (2017) is a particularly useful study, covering thousands of observations (n = 30,607) over twelve years, and finding no impact. Smaller studies by Cousin and Cuerton (2012) and Pitt and Winstone (2018) likewise find that it made no or negligible difference. Student campaigning on this issue might suggest that anonymous marking gives the impression of fairness, thereby at least promoting student trust in assessment. However, the same study by Pitt and Winstone (2018) also provides some (admittedly small-scale) evidence that, in practice, when students are asked about their perceptions, they do not believe that anonymously marked assessments are more fair.

There is one widely cited study, however, now famous beyond the small group of academics who take a particular interest in assessment: the case of screened auditions for symphony orchestras (Goldin and Rouse, 2000). The authors found, based on audition records from eight symphony orchestras from the 1950s to 1995, that the introduction of screened auditions has played a significant role in improving gender representation in orchestras. This finding is mentioned in the HEFCE report on anonymous marking (albeit via an interview with a stakeholder) (Mountford-Zimdars et al, 2015). There is also a widely repeated claim (see, notably, a 2014 TED Talk by Yassmin Abdel-Magied, which has been viewed over 2.8 million times) that hiring committees continue to discriminate against

female candidates, even in screened auditions, if the sound of high heels or a cough reveal their gender. This latter claim does not feature in the original article and I can find no empirical evidence for it whatsoever, but that does not stop it circulating. The striking nature of this story offers a 'folk' rationale for anonymous judgement.

The wide influence of this study makes it worth considering in greater detail, not least for the entry point it offers into thinking about some of the assumptions around anonymous marking. There are two questions to be addressed here. The first is whether the study bears scrutiny. The second is what subsequent debates about screened auditions within the community of orchestral musicians might help us understand anonymous judgements more broadly.

First, then, the statistician, Andrew Gelman, has looked in detail at the original study, noting that it has been overinterpreted and, according to Goldin and Rouse's *own* discussion, the results are 'not very impressive at all' (Gelman, 2019, n.p.). He also notes, following a discussion of statistical significance and the way that standard errors would be dealt with now in comparison to 2000, that 'we can't really trust these numbers' (Gelman, 2019, n.p.). Therefore, the Goldin and Rouse article is only one study and a contested one at that. I can find no other studies of screened auditions. Given what else we know about anonymous judgement, I, therefore, suggest that we should be very cautious about how far we continue to allow this article to influence policy.

Second, there have been intriguing debates among musicians about the usefulness of screened auditions. Maia Jasper White (2020), for example, argues that screened auditions place undue importance on the sound made by a musician (on one occasion) over other relevant considerations. In an argument that will resonate with advocates of ungrading, she notes that judgements about 'who sounds best' are deeply subjective. In a field where everyone auditioning is already supremely well qualified, there is a lot of luck involved. Therefore, she argues, we should also consider what else candidates offer. Alongside an audition, perhaps the recruitment process should include the evaluation of other artistic skills and attributes like generosity, character, charisma, humour, interpersonal skills, ideas and so on. Most of these things simply cannot be assessed anonymously. However, she suggests that racial and gender diversity would follow

naturally from this reconceptualisation of what orchestras mean by 'the best' candidate, as the range of different skills needed would create opportunities for different sorts of people.

I draw attention to this chiefly because it is still common to assume that anonymous marking probably is not doing any harm. Even Gelman (2019) suggests at the end of his polite demolition of the Goldin and Rouse (2000) article that there is nothing wrong with screened auditions as such. Meanwhile, Winstone and Boud (2022) do lay out some significant disadvantages to anonymous marking because of the way it depersonalises feedback, but instead of recommending we do away with the practice, they propose that we 'disentangle' marking from feedback and that only the mark should be anonymous.

However, White's argument seems to bear extrapolation to the ways in which we likewise only pay attention to the 'finished product'. The pedagogies associated with anonymous marking, after all, assume that the finished product is the only thing that matters insofar as it is the only thing the marker may know about when making summative judgements. It is also assumed the marker is the only person qualified to judge that product, thus ceding all power to them. This casts the student as the passive recipient of both marks and feedback. Anonymous marking also both assumes and perpetuates a culture of mistrust and suspicion between students and teachers, undermining the possibility of strategic alliances. Pedagogies of anonymous marking assume that bias is intractable and can only be mitigated, not challenged, by either students, or tutors' own self-reflection, or both together in dialogue. Finally, these pedagogies assume it is possible to arrive at a 'correct' mark outside of context or relationship and that 'fairness' is the primary value that transcends everything else. In the next section, I will seek to challenge these assumptions and argue against anonymous marking in favour of a more relational approach, which – I suggest – is more likely to help us succeed in closing awarding gaps than what seems to be, at the very least, the highly questionable model of anonymous assessment.

Relational marking: an alternative approach

Progressive educators have been thinking about how to go beyond a finished product in assessment for a long time. For example, Hughes' (2014) approach to 'ipsative assessment' foregrounds development and improvement. Relatedly, Cope et al (2021)

propose 'ergative assessment'. They argue that shifting our attention from the 'product' to the skills, effort and dispositions used in producing academic work seems particularly pressing in the era of large language models, where the generation of the 'finished product' can increasingly be done by the bots (see also Smith and Whitworth, 2024).

Both ipsative and ergative assessment have much in common with ungrading approaches, including contract grading (Cordell, 2019), which rewards the amount of effort and work put in, and portfolio-based approaches that rely on the ongoing development of, often multimodal, products across the course of a module with assessment based on peer, self and instructor evaluation (Blum, 2020a). There is by now a significant body of literature showing how these approaches can be used in a variety of disciplines, including STEM subjects, where awarding gaps seem to be particularly problematic (Royal Society, 2021). For these authors, negotiating a mark on the basis of effort and process is seen as a good alternative to marking the final product (for example, Blum, 2020b; Stommel, 2023; and see Jarvis, 2020 on ungrading in STEM subjects).

Building on these approaches, I propose relational marking: an approach that directly involves students in self-assessment and the negotiation of their own mark in collaboration with their teacher. I will discuss my own practice and pedagogy in my class to explain how this approach might reduce awarding gaps.

The class is a level 6 Politics module taken by a maximum of 40 second- and final-year students each year. It is optional and advertised as covering gender, sexuality and race. Perhaps for this reason, there tend to be many more women than men in the class and a large number of LGBTQ+ students opt to take the module. As expected for a university in London with a large international student population, the class is very ethnically diverse, and usually around half the class are students of colour.

The assessment for the class is a multimedia portfolio that students must work on each week. It is heavily scaffolded, with a range of suggestions and questions that students may respond to, but they are invited to 'be creative, daring and original'. My suggested assessment criteria are provided from the start, although these are negotiable and students are invited to shape how work is judged and evaluated. Students receive ongoing feedback on their portfolios from me and each other. The feedback is always framed in

reference to the stated assessment criteria, but students also have the latitude to question, contest or redefine them. As such, most of the assessment is formative and continuous, with students developing their critical understanding of the assessment criteria throughout the term.

This approach accords substantially with Winstone and Boud's (2022) recommendations on disentangling assessment and feedback. However, where I depart from them and adhere more closely to ideas about ungrading is in my approach to the final mark. Throughout the term – starting from Week 1 when we read Susan Blum's (2017) *Inside Higher Ed* article, which explains the many well-established problems with grading – the whole class has in-depth conversations about the role that marks have played in our lives. We discuss the way that marks have distracted us from learning, taking risks and being creative, and the atmosphere of mistrust and competition they create. This requires some emotional vulnerability from both me and the students, and it is not uncommon for a few tears to be shed in the classroom or during office hours as it dawns on students how much marks have dominated their lives, poisoned their relationships, and ruined their enjoyment of learning.

Discussions about feedback and quality of work are very much informed by these critical conversations, so when students ask me what mark I would give a piece of work or whether it is 'good', I consistently reply with a question such as, 'Why does that matter to you?' or 'What do you think?' and then redirect the discussion towards the assessment criteria. At the end of term, following lots of qualitative, dialogical feedback on their work, with opportunities to improve it, and discussions about mutual trust and responsibility, students are asked to fill in a self-assessment questionnaire. Having reflected carefully on their effort and explained how well their work meets the assessment criteria, they give themselves a numerical mark. I reserve the right to change that mark if necessary, but it almost never is. They do not, by any means, all give themselves a first-class mark and their own qualitative self-assessments are generally clear-eyed and thoughtful, particularly since I tweaked the questionnaire to encourage more reflection on room for improvement. I ask the students for permission to quote from their self-assessment and portfolio, and after two years this is now an incredible trove of data that enables me to reflect on their experiences of the module.

It is obvious that such a process could never be carried out anonymously for practical reasons, but I also want to highlight the ways in which the pedagogy that underlies it is philosophically in complete opposition to the approach that informs anonymous marking. In my class, the teacher and the students are on the same side and support one another to have in-depth conversations about what 'good work' looks like. Students are active partners with responsibility for their own learning and for evaluating how well they are doing. Students are asked to reflect on how hard they have worked and how far they have come, so that process, work and development matter as much as the 'final product', although the final portfolio is often a piece of work they are much prouder of than their essays for other classes. Relationship is valued above other considerations. We acknowledge, and have in-depth conversations about, the fact that bias and discrimination may occur, but that we can reflect on them, challenge them and strive to do better. The trusting relationships make it possible to do that reflective work together. Students are encouraged to bring their own knowledge and expertise to the table and they know I expect to learn from them, which I invariably do. My pitch to them about why they might listen to me is not that I hold the power to evaluate, reward or penalise them, but instead that I am on their side, and I have been doing this sort of work for a long time. I, therefore, suppose that I might be able in some ways to do it well. Finally, there is an acknowledgement that there is no such thing as an 'objective' or 'correct' mark that could be arrived at outside of context and relationship, since academic work is always the outcome of both.

Thinking about the putative causes of awarding gaps laid out by HEFCE, this approach might be expected to mitigate them. First, students have a great deal of control over what they are learning, so they are welcome to concentrate on the subject matter that is of most interest to them, so long as their work meets mutually agreed assessment criteria. They are also able to shape the assessment in ways that make the most sense to them and the assessment criteria are not just explained to students, but actively discussed with them in the hope they can define and shape them for themselves. This gives them an active role in making any hidden aspects of the curriculum visible and explicit. Second, collaborative relationships are promoted, with mistrust and competition explicitly challenged. Third, students are heavily encouraged to use the support systems available to them because feedback, peer support, and independent, scaffolded work are made central.

Most importantly, though, relational marking challenges precisely the psychosocial processes, including racism, that produce awarding gaps. This is because power relations are explicitly discussed, including being able to talk honestly about awarding gaps. The practice of critiquing the logic of marking gives students the opportunity to engage with and challenge the systems that produce awarding gaps in the first place. However, they can only do so with scaffolding: students are supported to understand debates about what 'good' work looks like, so they can subsequently explain how their work meets those criteria or why the ways in which good work is currently conceptualised are more likely to produce inequalities. To give a brief example, I have sometimes seen markers criticise students for writing in the first person (as I am doing now). When my students ask about this, we start with a conversation about the relevant assessment criterion: 'ability to write for an audience (please specify)'. This allows us to consider why certain readers may prefer the disembodied third person. A student thus equipped with assessment literacy understands that that marker is not making an objective judgement on their work. Rather, they are taking a particular – and contestable – view about the importance of acknowledging positionality and reflexivity in academic work (see, for example, Schwartz-Shea and Yanow, 2013). That student is then able to decide to what extent they agree with the feedback. Thus, the reflection provoked by relational marking actively allows students to analyse the specific criteria against which their work is judged. They can therefore make arguments about whether those criteria are useful or perhaps even whether they contribute to certain sorts of marginalisation related to awarding gaps. In the example about using the first person, a student may conclude that this marker subscribes to views about objectivity that deny that the personal is political, a position they – along with many feminist thinkers – may well disagree with. This leaves them better equipped to challenge such judgements analytically, even if they cannot change the intellectual edifice of power in academia overnight.

What happened?

I did not collect demographic data from the students, but as I did not mark anonymously, I am fairly confident that I know whether each of them identifies as a man, woman or non-binary and whether they are white or students of colour. Whilst, admittedly, there may have been mistaken identities here or there, on aggregate I am confident there were only

minuscule differences in the average marks between these groups of students, and that they ran in favour of women, non-binary students and students of colour. This is clearly small-scale and anecdotal evidence, but nevertheless might do something to quell the concerns of my external examiner who worried about the potential for awarding gaps in my approach.

More qualitatively, on the basis of what they wrote in their self-assessments and the quality of their work, I can report that some of my hunches about the potential of this approach to addressing the causes of awarding gaps did materialise. Students did take more risks in their work for this class, including trying out arguments they were not sure about, engaging in making different forms of media, including video and audio, and putting forward creative writing or their own artworks for assessment. This meant that their work was often better and less boring, which made marking a pleasure. Many wrote this was not something they had the opportunity to do in their other classes and that the format emboldened them to try things out. They also explored what mattered to them, with many students writing about race, colonialism, sexuality, social class and other areas where they felt they had lived expertise, which they could then understand even better by engaging with relevant academic literature and ideas. They also engaged with ongoing feedback from me and each other. They reported feeling happier and benefiting from a sense of community, with almost all students reporting that they had made a friend in the class. Most importantly, they took responsibility for their own learning and many reported feeling that they had learned more than they had in other classes.

The approach also enabled students to think critically about power relations. For example, one particularly serious and hardworking student who had been very focused on getting high marks proposed a new assessment criterion: 'have fun'. This challenge to the culture of stress and anxiety that was causing her real unhappiness was a moment of intervention in power relations, and a reminder that she, like all people, loves learning and that it should be fun, despite overbearing contrary pressures. Her work was incidentally also excellent when judged against the other assessment criteria I had suggested.

Two other students wrote about how the class enabled them to take back some power and responsibility for their own learning. One wrote:

I remember we were constantly asking Cathy, 'Is this OK?' and time after time she'd reply, 'What do you think?' Suddenly we had to think about our own learning – it was like coming out of autopilot and being in control of a plane!

I cannot think of a metaphor that more effectively communicates the sense of control that students gained from the approach. The other student wrote: 'I asked myself for a long time, "What? Does? She? Want?" until I finally thought about what I wanted out of this module and my degree'.

This displacement from taking her cue entirely from a teacher to assuming responsibility for herself is again a striking example of how relational marking empowers students. The first of these two students also noted that she was now much less likely to take marks in her other classes personally. She had decided to focus more on the feedback to see to what extent she agreed with the mark she had been awarded: a great illustration of taking responsibility and developing her own interdependent academic judgement.

Conclusion

In this article, I have argued that despite concerns that negotiated marking may exacerbate awarding gaps, in fact, *if* we take a relational approach, it should logically challenge them. I have proposed the most commonly adopted approach to eradicating awarding gaps – anonymous marking – has been shown to be rather ineffective on its own terms and also requires a pedagogy which is insufficient and counterproductive in challenging the power relations that produce awarding gaps in the first place. In contrast, I have proposed an alternative relational pedagogy of assessment and spelled out theoretically how it might challenge awarding gaps. With a particular focus on my own practice, I have then shown how this can work at module level and the subtle but impactful ways in which it allows students to understand and challenge power relations. This approach promotes assessment literacy and students' own independent, critical understanding of what makes 'good work' good. I hope this example provides a call to action and a model for adoption, further reflection and more research.

Some colleagues may argue that my approach to relational marking is very specific to the social science discipline in which I work, or that it would not be appropriate for large

classes or inexperienced tutors, or even (as an audience member at a recent conference suggested) that there is something particular about my own character that enables it in my class. I disagree. First, I have taken my cue from inspiring work on contract grading in STEM subjects, including Ryan Cordell's (2019) work on teaching data analysis and Clarissa Sorensen-Unruh's Chemistry classroom (Jarvis, 2020). Approaches to relational marking in circumstances where there really is an 'objective right answer' to assessment questions will differ from my more qualitative approach to assessment criteria, but the onus could still be placed on the student to figure out – with scaffolding and in a supportive relationship – how good they are at arriving at that right answer. However, colleagues in STEM subjects may also wish to think widely about whether there is more to their disciplines than 'right answers on the test'. Second, Courtney Sobers' (Rutgers School of Arts and Sciences, 2021) work with teaching assistants in large modules demonstrates that negotiated marking is possible in these circumstances. Sobers is also supporting the next generation of tutors to engage in ungrading in her role as a senior colleague and mentor: this is important if we want to embed a relational approach. As for character, I subscribe to the Foucauldian view (Dreyfus and Rabinow, 1982) that the person we are is the result of the practices and discourses we engage in. Do we want to be a person who coerces students with the threat of a bad mark or the promise of a good one? Do we want a relationship of mistrust that relies on anonymity to protect staff and students? Or could we take the risk of *cultivating* a professional character full of care, trust and joy, which encourages students to move out of a 'consumer' identity towards membership in a learning community (Symonds, 2020)?

What sort of person do you want to be? Could relational marking help you get there?

Another argument I sometimes encounter is that this approach takes up time tutors may not have. It is true that relational marking takes time. Personally, I took out a summative midterm assessment in order to spend more time giving formative feedback. I also spend relatively little time giving summative feedback, because students have had their feedback already, given when it could make the most difference. This is less stressful and more interesting for me and the students. I also want to argue that tutors in general would benefit from spending time doing work that is fun and rewarding. It is just more pleasurable to give formative feedback to attentive, motivated students who are excited to learn than it is to mark lots of identical, risk-free, defensive essays written to please an imagined

version of the teacher. Sometimes, in our efforts to save a few minutes, we make our own lives more miserable than they need to be. To put it another way, when one of my colleagues looked at me joyfully doing my marking while he suffered through his, he decided to adopt my approach, saying: 'I'll have some of what she's having'.

I want to conclude with an example of how relational marking can challenge the very distribution of self-confidence that so worries sceptics. One of my students wrote:

I am a brown, disabled woman and I have always been taught that I don't matter much. But I have awarded myself a first class mark for this module because I have worked hard, thought about the assessment criteria, and I think I deserve it.

Having read her excellent portfolio, and witnessed the way she engaged with feedback to improve it, I am firmly in agreement with her. This moving testimony suggests the skill and confidence to make, and argue for, judgements about what constitutes good work, and to advocate for oneself and one's work, is a skill that can be taught like any other. However, this work cannot be done under the cloak of anonymity: our job is to foster and develop the trusting, mutual relationships within which the crucial task of challenging awarding gaps can finally take place.

A common misunderstanding of my teaching practice is that I allow students to 'mark their own work'. If this were true, it might be justified to worry that less advantaged students might, on average, award themselves lower marks because of lower overall self-confidence or prior experience of getting lower marks. However, I hope this article and, particularly, the above example make it clear that the stakes, and potential rewards, are much higher than that. I have advocated here for a thoroughgoing relational pedagogy that treats students as members of the academic community, who can make judgements *in collaboration* with their teachers and each other. As such, they have a right to be part of the conversation about academic values and to contribute to ongoing discussions – which will always be contested and fraught with potential conflict – about what constitutes worthwhile work. They are not born knowing how to do this, but they can be taught how to exercise that right and, in doing so, perhaps make their own contribution to closing awarding gaps.

Acknowledgements

I am grateful to participants at the KCL/UCL Freedom To Learn conference on 5 April 2024, the RefleX seminar at Université Libre de Bruxelles on 11 June 2024, the Political Studies Association Teaching and Learning Network conference in Durham on 5 and 6 September 2024 and the University of Leeds Assessment Matters seminar on 11 November 2024, for all their helpful questions, discussion, and support for earlier drafts of this article. Thank you also to two (ironically!) anonymous referees and the editors of this special edition for very useful feedback, which has made the article better. I am also grateful to numerous colleagues at UCL including everyone in the Freedom to Learn network, the Queer Pedagogy gang, my friends in the Arena Centre and the Faculty of Social and Historical Sciences, supportive second markers and external examiners, and especially Fergus Green, for all their support for, and useful constructive comments on, my unusual teaching and assessment practice. My friend, colleague and co-conspirator, Martin Compton, has supported me from the start to make 'ungrading' a reality in my teaching – it would not have happened without him. Cheers to Mike for proofreading, editing and reliably asking the right questions; you sharpened my thinking and improved my prose. And, above all, thank you with all my heart to three cohorts of truly fantastic students who remembered that they came to university to learn and repaid my trust beyond what I could ever have hoped. You all deserve to think of yourselves as so, so much more than just a grade.

I did not use generative AI technologies in the creation of this manuscript.

References

Blum, S.D. (2017) 'Ungrading,' *Inside Higher Ed blog*, 13 November. Available at: <https://www.insidehighered.com/advice/2017/11/14/significant-learning-benefits-getting-rid-grades-essay> (Accessed: 11 July 2024).

Blum, S.D. (2020a) 'Just one change (just kidding): ungrading and its necessary accompaniments', in S.D. Blum (ed.) *Ungrading: Why Rating Students Undermines*

Learning (and What to Do Instead). Morgantown, WV: West Virginia University Press, pp.53-73.

Blum, S.D. (2020b) (ed.) *Ungrading: Why Rating Students Undermines Learning (and What to Do Instead)*. Morgantown, WV: West Virginia University Press.

Borkin, H. (2020) *Unconscious Bias Literature Review: Bias in Assessment*. Available at: <https://www.advance-he.ac.uk/knowledge-hub/unconscious-bias-literature-review-bias-assessment> (Accessed: 11 July 2024).

Campbell, P. and Duke, B. (2023) *An Evaluation of Racially Inclusive Practice in Assessment Guidance Intervention on Students' and Staffs' Experiences of Assessment in HE: A Multi-University Case Study*. Available at: <https://doi.org/10.25392/leicester.data.23579565.v1>

Cope, B., Kalantzis, M. and Searsmith, D. (2021) 'Artificial intelligence for education: knowledge and its assessment in AI-enabled learning ecologies', *Educational Philosophy and Theory*, 53(12), pp.1229-1245. Available at: <https://doi.org/10.1080/00131857.2020.1728732>

Cordell, R. (2019) 'How I contract grade,' *Ryan C. Cordell blog*, 7 December. Available at: <https://ryancordell.org/teaching/contract-grading/> (Accessed: 11 July 2024).

Cousin, G. and Cuerton, D. (2012) *Disparities in Student Attainment (DiSA)*. Available at: <https://wlv.openrepository.com/server/api/core/bitstreams/f69a8ec1-3163-4463-9033-55283f7f025b/content> (Accessed: 21 March 2025).

Dreyfus, H.L. and Rabinow, P. (1982) 'Afterword: The Subject and Power' in *Michel Foucault: Beyond Structuralism and Hermeneutics*. London: Routledge. Available at: <https://doi.org/10.4324/9781315835259>

Gelman, A. (2019) 'Did blind orchestra auditions really benefit women?' *Statistical Modeling, Causal Inference, and Social Science blog*, May 11. Available at:

<https://statmodeling.stat.columbia.edu/2019/05/11/did-blind-orchestra-auditions-really-benefit-women/> (Accessed: 11 July 2024).

Goldin, C. and Rouse, C. (2000) 'Orchestrating impartiality: the impact of "blind" auditions on female musicians', *American Economic Review*, 90(4), pp.715-741. Available at: www.doi.org/10.1257/aer.90.4.715

Hinton, D.P. and Higson, H. (2017) 'A large-scale examination of the effectiveness of anonymous marking in reducing group performance differences in higher education assessment', *PLoS ONE*, 12(8), e0182711. Available at: <https://doi.org/10.1371/journal.pone.0182711>

Hughes, G. (2014) *Ipsative Assessment: Motivation through Marking Progress*. London: Palgrave Macmillan.

Jarvis, C.L. (2020) 'Chemistry educators try "ungrading" techniques to help students learn', *Chemical and Engineering News*, 26 April. Available at: <https://cen.acs.org/education/undergraduate-education/Chemistry-educators-try-ungradingtechniques-help/98/i16> (Accessed: 15 October 2024).

Malouff, J.M. and Thorsteinsson, E.B. (2016) 'Bias in grading: a meta-analysis of experimental research findings', *Australian Journal of Education*, 60(3), pp.245-256. Available at: <https://doi.org/10.1177/0004944116664618>

Mountford-Zimdars, A., Sabri, D., Moore, J., Sanders, J., Jones, S., and Higham, L. (2015) *Causes of differences in student outcomes: Report to HEFCE by King's College London, ARC Network and The University of Manchester*. Available at: https://dera.ioe.ac.uk/id/eprint/23653/1/HEFCE2015_diffout.pdf (Accessed: 11 July 2024).

National Union of Students (2008) *Mark My Words, Not My Name: The Campaign for Anonymous Marking*. Available at: <http://samairaanjum.weebly.com/uploads/1/0/5/2/10526755/markmywordsbrief1-1.pdf> (Accessed: 11 July 2024).

Office for Students (2021) *Differences in Student Outcomes*. Available at:

<https://www.officeforstudents.org.uk/data-and-analysis/differences-in-student-outcomes/ethnicity/> (Accessed: 11 July 2024).

Pitt, E. and Winstone, N. (2018) 'The impact of anonymous marking on students' perceptions of fairness, feedback and relationships with lecturers', *Assessment and Evaluation in Higher Education*, 43(7), pp.1183-1193. Available at:

<https://doi.org/10.1080/02602938.2018.1437594>

Royal Society (2021) 'STEM sector must step up and end unacceptable disparities in Black staff and students' academic progression and success', *News from the Royal Society*, 25 March. Available at: <https://royalsociety.org/news/2021/03/stem-ethnicity-report/> (Accessed: 18 October 2024).

Rutgers School of Arts and Sciences-Newark (2021) *Chart a Course: Ungrading with Courtney Sobers*. 30 June. Available at:

<https://www.youtube.com/watch?v=YadoqOs6UpA> (Accessed: 16 October 2024).

Schwartz-Shea, P. and Yanow, D. (2013) *Interpretive Research Design: Concepts and Processes*. London: Routledge. Available at:

<https://doi.org/10.4324/9780203854907>

Smith, P. V. and Whitworth, D. (2024) 'Anonymous assessment: is it still worth it?' *Teaching in Higher Education*, pp.1-9. Available at:

<https://doi.org/10.1080/13562517.2024.2363891>

Stommel, J. (2023) *Undoing the Grade: Why We Grade and How to Stop*. Denver, Colorado: Hybrid Pedagogy, Inc.

Symonds, E. (2020) 'The depersonalised consumer subjectivity and its effect on fostering meaningful relationships between undergraduates and academics in higher education', *Critical Studies in Education*, 62(5), pp.624-640. Available at:

<https://doi.org/10.1080/17508487.2020.1755330>

TED (2014) *Yassmin Abdel-Magied: What does my headscarf mean to you?* December.

Available at:

https://www.ted.com/talks/yassmin_abdel_magied_what_does_my_headscarf_mean_to_you (Accessed: 11 July 2024).

UCL (2023) *BAME Awarding Gap Project: Evaluation Report*. Available at:

https://www.ucl.ac.uk/teaching-learning/sites/teaching_learning/files/bame_awarding_gap_project_evaluation_report_-_may_202374.pdf (Accessed: 11 July 2024).

Universities UK (2022) *Closing The Gap: Three Years On*. Available at:

<https://www.universitiesuk.ac.uk/what-we-do/policy-and-research/publications/features/closing-gap-three-years/introduction> (Accessed: 11 July 2024).

White, M.J. (2020) 'Eyes wide shut – the case against blind auditions', *NewMusicBox*, September 10. Available at: <https://newmusicusa.org/nmbx/eyes-wide-shut-the-case-against-blind-auditions/> (Accessed: 11 July 2024).

Winstone, N. and Boud, D. (2022) 'The need to disentangle feedback and assessment in higher education', *Studies in Higher Education*, 47(3), pp.656-667. Available at: <https://doi.org/10.1080/03075079.2020.1779687>

Author details

Cathy Elliott teaches in the Political Science department at UCL and has more than 15 years' experience as an educator and educational leader. Her research interests are in the pedagogy of politics, particularly the role of joy, freedom, the imagination and creativity in education. She is currently Professor and Vice Dean Education in the Faculty of Social and Historical Sciences at UCL and is an active educator and educational researcher. She is the co-founder and co-Director of the UCL Centre for the Pedagogy of Politics and was

co-Convenor of the PSA Teaching and Learning Network from 2022 to 2025. She is also a co-editor and regular writer for the APSA/PSA award-winning ALPS blog.

Licence

©2025 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International Licence (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0>. Journal of Learning Development in Higher Education (JLDHE) is a peer-reviewed open-access journal published by the Association for Learning Development in Higher Education (ALDinHE).