Institution level awarding gap metrics for identifying educational inequity: useful tools or reductive distractions?

Katharine Elizabeth Hubbard[1,2]

Abstract

Equity is increasingly seen as a core value for higher education systems around the world. (In)equity is often measured through construction of achievement gaps, quantifying the relative outcomes of two populations of students. Institution-level gaps are embedded in the policy landscape of HE, becoming performance metrics in their own right. These gap metrics increasingly inform the actions of governments, regulators, institutions and educators. This theoretical article scrutinises the technical and conceptual construction of achievement gaps through using the dominant UK conception of the institution level degree classification 'awarding gap'. Drawing on Adam's Equity Theory of Motivation, Rawls's Distributive Justice and the Capability Approach as theoretical perspectives, I highlight multiple structural weaknesses in the conception of the awarding gap. I illustrate the implications of this metric by analysing simulated awarding gap data for a fictional institution, and through the perspectives of five idealised stakeholders. I identify multiple technical and theoretical limitations of the institution level awarding gap metric, including examples where the threshold-based nature of the awarding gap fails to capture statistical differences between groups, thereby undermining its utility in identifying inequity. I call on the sector to develop metrics that more accurately capture (in)equity of outcomes and align better with theoretical frameworks, thereby creating more powerful explanatory metrics that can inform meaningful action.

Keywords Equity · Achievement gap · Awarding gap · Higher education policy · Student outcome metrics

Introduction

UN Sustainable Development Goal 4 establishes a global ambition to *"Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all"* (United Nations, 2015*)*. However, higher education (HE) systems around the world perpetuate or even reinforce societal inequalities (Universities UK and National Union of Students, 2019; Lamb,

✉   Katharine Elizabeth Hubbard
     k.hubbard@hull.ac.uk

1   Teaching Excellence Academy, University of Hull, Hull, UK

2   School of Natural Sciences, University of Hull, Hull, UK

✌ Springer

2020; Bertolin & McCowan, 2022; Li & Jackson, 2023). Educators, institutions and governments are addressing educational inequalities, aspiring to ensure that students are not disadvantaged due to their social background, current circumstances or demographic characteristics (Universities UK and National Union of Students, 2019; Cabral-Gouveia et al., 2023; Cagliesi et al., 2023). However, achieving genuine educational equity is a major challenge requiring significant resourcing, pedagogical intervention and student support mechanisms.

This conceptual article asks whether quantitative institution level student outcome 'gap' metrics usefully measure educational (in)equity? I use the UK undergraduate degree classification 'awarding gap' as a case study to explore theoretical and statistical value of institution level metrics. I will not consider what actions are appropriate when gap data identifies inequity; my focus is on whether these metrics adequately capture educational inequity in a useful way.

What is educational equity?

While equity is a commonly used term in education, it has multiple interpretations (Brookover and Lezotte, 1981; Espinoza, 2007; Brennan & Naidoo, 2008; Wilson-Strydom, 2015; Naylor and Mifsud, 2020; Edgar, 2022; Levinson et al., 2022). Several authors identify confusion within the educational literature between 'equity' and 'equality', with both terms being used interchangeably or to mean broader concepts such as fairness or social justice (Espinoza, 2007; Brennan & Naidoo, 2008; Edgar, 2022). Table 1 summarises a variety of interpretations of 'equity' within HE. The operationalisation of equity matters, as university leaders make decisions based on improving their metrics (Locke et al., 2008). Here, I will define equality as treating everyone the same independent of circumstance, while equity is giving everyone what they need to succeed in a given environment, which may involve unequal allocation of resources to redress disadvantage (Levinson et al., 2022). Equity therefore requires subjective and potentially controversial judgements to achieve equivalence of outcomes (Espinoza, 2007; Edgar, 2022). Contextualised admissions provide a practical example of equity-driven HE practise, whereby student entry qualifications are judged alongside markers of socioeconomic disadvantage (Boliver et al., 2022). Subjective contextual decisions may radically change composition of the student body (Liu, 2011) so require careful consideration.

One common measure of inequity is the 'achievement gap', a quantitative difference in outcomes between groups. For example, 10% of young people from Australian indigenous communities have a degree compared with 42% from non-indigenous backgrounds, a 32% gap (Lamb, 2020). The outcome measure and demographic groups of interest vary by country, but the 'gap' is widespread. There is increasing use of these measures at policy level (Cagliesi et al., 2023). For example, the UK HE regulator (The Office for Students; OfS) has set sector level expectations around closure of multiple outcome gaps via Key Performance Measures (Office for Students, 2022b).

UK awarding gaps as an outcome metric

Any achievement gap is actively conceptualised through its calculation. Subjective choices made in gap calculations influence meaning and interpretation. I use awarding gaps within UK HE as a case study to explore the utility of achievement gap measures. UK awarding gaps reflect awarding rates of 'good degrees', meaning a 1st class or upper 2nd class (2i)

Table 1 Interpretations of equity within educational literature relevant to HE. Categories based on Brookover and Lezotte ([1981](#)), with the addition of Resourcing. Interpretations drawn from Espinoza ([2007](#)), Naylor and Mifsud ([2020](#)), Cairney and Kippin ([2022](#)), Pitman et al., ([2020](#)) and Levinson et al. ([2022](#)). * indicates the position of UK awarding gaps

Category	Interpretation of equity	Potential measures within HE
Resourcing	• Provision of educational resources	• Financial resource • Staff-to-student ratio • Student satisfaction ratings
Access	• Access to HE • Access to elite institutions or high status disciplines	• Composition of student population • Acceptance rates
Participation	• Individual student experience • Personal/academic growth and development for each learner	• Qualitative measures • Academic grades • Student satisfaction ratings
Outcomes	• Achievement of educational thresholds across groups* • Distribution of educational outcomes across demographic groups	• Course completion rates • Academic grades • Degree classification* • Progression to employment or higher level study • Graduate salaries

degree classification (Fig. 1). A student typically requires an average grade of 70% for 1st and 60% for 2i. This threshold is usually justified on the basis that students with a 1st/2i typically obtain better-paying jobs after graduation (Britton et al., 2022). The UK awarding gap is therefore an outcome measure based on degree classification, constructed by comparing achievement of educational thresholds across groups (Table 1).

Awarding gaps are well documented within UK HE, particularly the racial awarding gap which persists after controlling for entry qualifications (Universities UK and National Union of Students, 2019; AdvanceHE, 2021; Bolton and Lewis, 2023). The awarding gap has become a metric informing UK HE policy and practise (AdvanceHE, 2021; Wong et al., 2021; Cagliesi et al., 2023), although the UK regulator has been inconsistent in whether it focuses on degree completion or classification as an outcome (Office for Students, 2022b). The technical construction of other metrics relevant to UK HE policy has been studied (Tofallis, 2012; Gunn, 2018; Hosier and Hoolash, 2019; Hubbard, 2021). For example, Tofallis (2012) explores the impact of normalisation methods on institutional
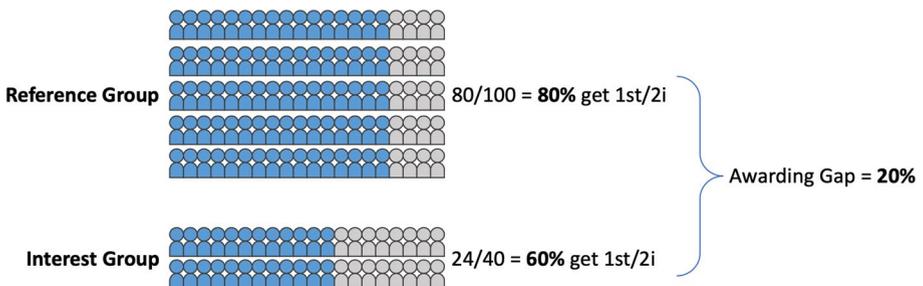
Fig. 1 Calculation method for UK awarding gaps. Awarding gap=% students in the reference group awarded 1st/2i—% students in the interest group awarded 1st/2i., e.g. Interest group might be Black students, white students the reference

league tables. There has been less technical and theoretical critique of UK awarding gap metrics, which this article aims to redress.

Before embarking on critique, I defend awarding gap metrics on a pragmatic basis. They have undoubtedly focussed sector attention on inequitable outcomes (Universities UK and National Union of Students, 2019; Dickenson, 2021). The OfS requires institutions to set Access and Participation Plans addressing inequities including awarding gaps (Office for Students, 2018). Quantitative metrics influence decision-making of university leaders (Hazelkorn, 2007; Locke et al., 2008; Hubbard, 2021), so awarding gap metrics put equity onto the agenda (Hubbard, 2021). A standardised metric allows for systematic evaluation of sector or institution level initiatives (Hazelkorn, 2007). In the UK context, the gap also provides a potential tool for understanding legal issues around discrimination. The value of the awarding gap is therefore perhaps not in its theoretical or technical construction but as a tool to stimulate strategic discussions and actions around educational equity.

Theoretical context

To explore interpretations of equity related to the awarding gap, consider two hypothetical students on the same course. Theo is white, upper middle class and attended private school. His university-educated parents pay for student accommodation close to his department. Theo is confident of getting a high-paying job, so devotes more time to university sports than to study. Ayesha is the child of first-generation immigrants and speaks Urdu at home. The graduate training scheme she aspires to has a minimum entry requirement of a 2i. Ayesha commutes to university via public transport, works part-time and cares for a disabled relative.

Theo and Ayesha both obtain a 2i degree. Is this equitable? Both had equal opportunity to participate in HE and had equal outcomes. However, most would agree that Theo had considerable advantage over Ayesha. How can we understand this in a robust theoretical way? Here I briefly explore three models; Equity Theory (Adams, 1963), Distributive Justice (Rawls, 1971), and the Capability Approach (Nussbaum, 2009, 2011; Sen, 1979, 2009).

Adam's Equity Theory of Motivation (1963) proposes that an individual perceives equity when the ratio of their inputs to outputs is equivalent to that of another person or comparison group (Adams, 1963). It is the perceived ratio of input to output that defines (in)equity, not absolute level. An individual might see it as equitable when a colleague is paid more if they recognise that they also work harder. Adams frames equity as inherently comparative; without knowing the outcomes of a reference group an individual cannot assess fairness of their own outcome. Equity theory also proposes that inequity leads to psychological distress and subsequent behaviour change (Adams, 1963; Davlembayeva & Alamanos, 2022). Ayesha might see Theo achieving equivalent academic grades (output) despite not doing as much studying (input) and may experience anger, frustration or disillusionment as a result. She may study harder to gain academic recognition or may disengage feeling the situation is stacked against her.

Equity might alternatively be considered through the prism of distributive justice (Rawls, 1971, 1985). Through construction of three core principles, Rawls argues that societies may be just even where inequalities exist, providing that all have the ability to benefit from those inequalities and that inequality is not based on factors beyond individual control (the equal opportunity principle). This principle is reflected in many definitions

of educational equity including those of the Organisation for Economic Co-operation and Development (OECD) (Salinas, 2018). Distributive justice-based practises include provision of finance to socioeconomically deprived students to redress disproportionately low enrolment rates (Msigwa, 2016). Distributive justice could potentially be achieved if Ayesha is provided with sufficient resources (e.g. bursaries and additional tuition) to overcome her disadvantage. The resource centric model of Rawls has been widely criticised and expanded on to create more holistic understanding of justice (Cook & Hegtvedt, 1983; Nussbaum, 2009; Wilson-Strydom, 2015). For example, procedural justice describes fairness of decision-making, while interactional justice captures fairness expressed through interpersonal interactions (Colquitt & Greenberg, 2003; Cook & Hegtvedt, 1983; Leventhal, 1980). It has been argued that achieving educational equity requires a distributive justice model rather than procedural justice, particularly in terms of admission to highly selective universities (Boliver et al., 2022).

The Capability Approach moves away from resource-based conceptions of equity, adopting a broader humanistic perspective (Nussbaum, 2011; Sen, 2009; Walker & Unterhalter, 2007; Wilson-Strydom, 2015). This approach centres on individual well-being, defined as "the freedom that a person actually has to do this or be that—things that he or she may value doing or being" (Sen, 2009). It distinguishes between 'capabilities' which are the freedoms or opportunities an individual has to achieve and 'functionings' which are the outcomes actually achieved after individual choices. Sen argues that to assess equity and justice, we should focus not on equal functionings but on equal capabilities. Theo and Ayesha have achieved equivalent functionings (degree classification), but Theo arguably had a higher capability to succeed. The Capability Approach has several advantages over Rawslian models (Vaughan, 2007; Walker & Unterhalter, 2007). It adopts a broader understanding of equity than resource-based models, as capabilities are inherently pluralistic (Flores-Crespo, 2007; Nussbaum, 2011; Robeyns, 2005). The distinction between capabilities and functionings captures choice, and it includes the differing potential that individuals have to convert resources into capabilities via conversion factors (Liu, 2011; Vaughan, 2007; Wilson-Strydom, 2015). Additional primary resources (e.g. bursaries and extra tuition) may not be able to level the playing field between Theo and Ayesha. Even if redistributive finance is available to disadvantaged groups, individuals may still experience barriers accessing that finance, or the finance may not be sufficient to overcome other structural aspects of disadvantage (Msigwa, 2016; Wilson-Strydom, 2015). The Capability Approach also respects diversity of personal and socio-environmental influences via conversion factors and individual choice but does not encapsulate aspects such as procedural justice (Robeyns, 2005).

Theoretical critique of awarding gaps

What do these theoretical ideas add to our understanding of the awarding gap? First, the gap imposes a unidimensional quantitative definition of success, which has been criticised for centring the outcomes of white men (Shukla et al., 2022). Nussbaum notes that people "cannot without distortion be reduced to a single scale" (Nussbaum, 2011, p. 19). The gap fails to recognise the plurality of equity within HE (Table 1). Can an institution be equitable if it has no ethnicity awarding gap but Black students routinely experience discrimination? Students are likely to have a more qualitative interpretation of equity including

procedural and interactional justice not encapsulated by the gap (Struyven et al., 2005; Burger, 2017).

The Capability Approach exposes significant flaws of the gap in describing (in)equity. Theo and Ayesha's achievements (functionings) are seen as equivalent when calculating the gap, but their differential capability to achieve is not reflected. The gap also ignores individual and socioenvironmental influences expressed through personal choice (Robeyns, 2005). Theo's social privilege gives him the freedom to choose sport over studying. Ayesha does not have that luxury. The reductive nature of the metric means that institutions with a gap of 0% may falsely conclude that there is no structural disadvantage. Eliminating the gap is often assumed to mean that equity has been achieved, but this ignores the fact that the gap measures functioning, not capability. Disadvantaged students may have lower outcomes or have to expend more effort to achieve the same outcomes as more advantaged peers. Both situations are inequitable. Some argue gaps reinforce a deficit model of educational outcomes for minoritised groups or prevent recognition of 'educational debt' (Ladson-Billings, 2006) owed to students facing structural disadvantage (Gutiérrez, 2008; Shukla et al., 2022). To genuinely address inequity, we need to understand the conversion factors or barriers that influence student success, from personal beliefs to societal conditions (Wilson-Strydom, 2015). These factors will be highly individualised and intersectional (Crenshaw, 1989). Ayesha's situation and needs may differ considerably from other Asian students in her institution, but the group-based nature of the 'Asian awarding gap' may assume a uniformity of experience.

While the Capability Approach has value, it focuses attention on the student rather than the institution and does not capture procedural justice (Robeyns, 2005). Institutional factors such as course design, assessment weightings, fairness of marking procedures, degree classification algorithms, or biases of staff are not easily captured. The university plays an active role in achieving (in)equity, reflected in the use of 'awarding gap' language rather than 'attainment gap' (Joseph-Salisbury, 2020). It is possible to imagine the university adopting redistributive 'contextual awarding', mirroring distributive justice-based contextual admissions (Boliver et al., 2022). However, this would be highly controversial in UK HE which prioritises student meritocracy and quality standards as a mechanism to achieve procedural equality. It should also be noted that UK HE degree classifications are criterion referenced (QAA, 2019). This limits the value of input-to-output-based models, as degree classifications reflect academic standards not just effort.

The 2i/1st threshold requires particular scrutiny. This high standard differs from most threshold-based interpretations of equity which typically use (inter)nationally defined minimum benchmarks (e.g. literacy standards) (Levinson et al., 2022). The connection between the labour market and the 1st/2i threshold centres economic concerns rather than personal development and growth, distancing it from the philosophy of the capabilities approach. The economic lens also changes the theoretical position of the gap. Even though Ayesha values her 2i, she does so through the lens of employability. Her degree classification is perhaps better understood as a conversion factor towards employment rather than a functioning valued in its own right (Flores-Crespo, 2007). The awarding gap therefore centres on a sector and institutional definition of success that may not align with outcomes valued by individual students and employers. A majority of graduate employers no longer require a 1st/2i (Institute of Student Employment, 2023). The high threshold also puts closing awarding gaps in conflict with the regulator's parallel demand to tackle so-called 'grade inflation' (Bachan, 2017). Interestingly, the UK regulator has recently changed the threshold for its Key Performance Measure related to Black awarding gap to only consider 1st class degrees (Office for Students, 2022d), the implications of which are as yet unclear.

Future directions: theoretical

The awarding gap implicitly draws on aspects from multiple theoretical models but also has significant limitations. To move on from 'gap gazing', there is a need for awarding gap models to have a predictive component (Gutiérrez, 2008). There have been quantitative analyses of student outcomes by sociodemographic and institutional factors (AdvanceHE, 2021; Office for Students, 2022c), but these are limited by reliance on demographic factors (e.g. gender), rather than underlying explanatory drivers (e.g. conscious/unconscious bias and motivation). Without clear theoretical models underpinning the awarding gap, it is challenging to make testable predictions about underlying causes, hampering identification of effective interventions (Brennan & Naidoo, 2008).

I present potential awarding gap models aligned with Equity Theory and the Capability Approach in Fig. 2. These include both student factors (e.g. prior learning and personal circumstances) and institutional factors (e.g. assessment modes and degree classification algorithms). These models potentially allow for identification of causative factors. For example, students undertaking significant amounts of paid work may have different input-to-output ratios than for peers. Appropriate actions might therefore be financial support or block timetabling to minimise the number of classes missed. By more clearly articulating the theoretical basis underpinning the gap, it should be easier to identify evidence-based interventions.

Technical critiques of awarding gaps

I now turn to technical concerns. The critiques below are not an endorsement of quantitative metrics as the 'correct' approach but a recognition that they influence decision-making in practise and therefore their construction and interpretation require scrutiny.

Quantitative metrics should accurately capture underlying data. In the UK, degree classification and gap data are publicly available, but numerical grades underlying these are not published. Without having both, it is impossible to assess the technical accuracy of the gap. I therefore construct a model university with individual student grade data and calculated gaps (see Supplementary information). The model allows real-world situations to be simulated while providing a simplified system through which to explore accuracy of gap metrics.

The simulated institution has 1000 students graduating, split between three faculties (Table 2). Three demographic factors are represented; race (white, Black, and Asian[1]), ability (no disability and disabled) and age (young and mature). For each demographic and faculty combination, the number of students, mean and standard deviation of grades and difference in mean grade between groups are set to replicate real-world scenarios. Grade data are randomly generated within these parameters for each student, and population level awarding gaps calculated. The relationship between demographic and grades is verified via linear models. The institution has a 16% racial awarding gap, 11% age-related gap and 4% disability gap (Table 2), typical of recent UK gaps (Office for Students., 2020).

---

[1] These are the most commonly used racial categories used in UK HE, which may not be relevant in other international contexts where e.g. indigenous students are the interest group. These categories also reflect the social construction of race, which has a history inherently linked to racism and oppression which should be acknowledged.
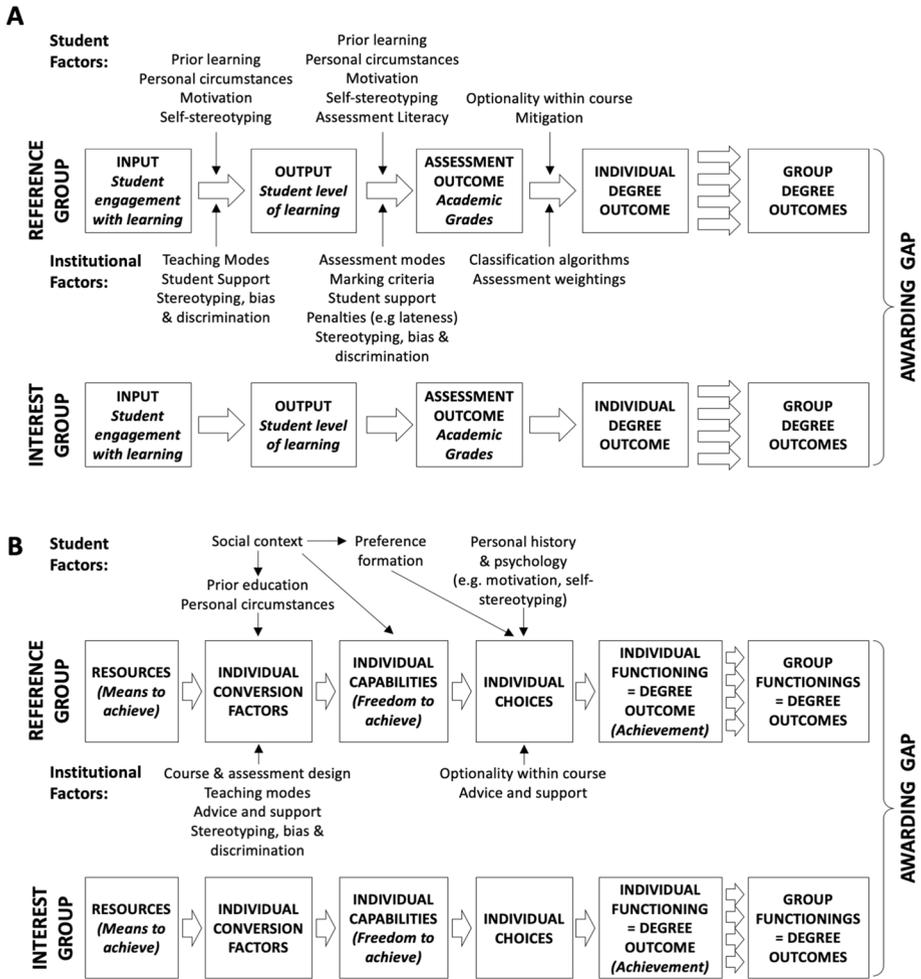
Fig. 2 Potential awarding gap models aligned to **A** Equity Theory of Motivation and **B** the Capability Approach. Note that final degree outcomes are calculated from multiple assessment grades. B based on Robeyns (2005)

We cannot consider whether awarding gap metrics are useful without defining who they are useful to. I therefore establish multiple stakeholders through which to assess the usefulness of gaps:

1. **The regulator.** The regulator assesses data at national and institutional level and defines key performance measures to provoke behavioural change from institutions. I assume the regulator acts in good faith and genuinely wants equity of outcomes
2. **The institution.** The institution awards degrees and is overseen by the regulator. Senior leaders make strategic decisions based on their visions and values, key performance indicators, and external pressures including the regulator. I establish two extreme positions an institution might take but acknowledge most institutions will exist somewhere between these:

Table 2 Overview of awarding gaps within the simulated University. Negative gaps indicate the group of interest has higher outcomes than the reference group. *BAME*, Black, Asian, Minority Ethnic

Demographic	Outcomes	Faculty A	Faculty B	Faculty C	Institution
All students	*n*	330	290	380	1000
	% 1st/2i	94%	79%	61%	77%
'BAME'	White *n*	240	132	267	639
	White % 1st/2i	95%	95%	65%	83%
	'BAME' *n*	90	158	113	361
	'BAME' % 1st/2i	89%	66%	50%	67%
	'BAME' awarding gap %	6%	29%	15%	16%
Asian	Asian *n*	44	89	61	194
	Asian % 1st/2i	89%	76%	56%	73%
	Asian awarding gap %	6%	19%	9%	10%
Black	Black *n*	46	69	52	167
	Black % 1st/2i	89%	52%	44%	60%
	Black awarding gap %	6%	43%	21%	23%
Disability	No disability *n*	238	241	286	765
	Disabled *n*	92	49	94	235
	No disability % 1st/2i	94%	80%	63%	78%
	Disabled % 1st/2i	93%	76%	54%	74%
	Disability awarding gap %	1%	4%	9%	4%
Age	Young *n*	277	239	117	633
	Mature *n*	53	51	263	367
	Young % 1st/2i	94%	77%	58%	81%
	Mature % 1st/2i	94%	88%	62%	70%
	Mature Awarding Gap %	0%	-11%	-4%	11%

a   **The socially just institution.** This university has established equity as a core principle and wants to ensure that all students are able to realise their full potential. Leaders act to meet targets established by the regulator, but this is secondary to the broader principle of equity embedded in the culture of the institution

b   **The cynical institution.** This university does not fundamentally care about equity of outcomes but recognises the importance of meeting targets established by the regulator. Leaders act to meet key performance measures in the most expedient way possible

3.   **The local area education lead.** This person is responsible for closing awarding gaps identified in their area. They may be a programme director or another local leadership role. They have multiple competing priorities and limited capacity, so value efficiency and appropriately targeted information

4.   **The student from a disadvantaged group**. I assume the student is aware of institutional awarding gaps relating to their socio-demographic group via a student union campaign. Students will have differing perceptions of the awarding gap (Wong et al., 2021). Some students may have a secondary activist interest in fairness of awarding, but my student is only interested in fairness of their own outcomes

Technical critique 1: awarding gaps are an incomplete metric of equitable outcomes

The construction of any metric will influence its meaning and interpretation (Tofallis, 2012). Awarding gaps do not capture resource, access or participation-based conceptions of equity (Table 1). The most obvious criticism of the UK awarding gap calculation is that it only includes students who graduate (Table 3).

This tension between different outcome measures can potentially be exploited. The cynical university may game the metrics (Oravec, 2019), perhaps by encouraging students from disadvantaged groups who are unlikely to achieve a 1st/2i to discontinue studying. To ensure institutions are not rewarded for unethical behaviour, the regulator must adopt multiple complementary measures of equity and ensure institutions cannot pick and choose metrics.

Technical critique 2: awarding gap metrics oversimplify distributions by using a binary threshold model

The 1st/2i threshold-based gap has significant consequences for data interpretation data. The operating assumption is that an awarding gap reflects a difference in average marks (Fig. 3A, Scenario A). However, my simulated data demonstrates there may be little relationship between underlying grade distributions and calculated awarding gaps (Scenarios B, C and D). Scenario B illustrates that there can be no awarding gap as defined by the 2i threshold when there is a significant difference in mean marks. Scenarios C and D illustrate that there can be large threshold-based gaps with no statistically significant difference in marks. This may lead to a disconnect between assumptions of stakeholders and actual data. The most serious implication is for institutions whose data mirrored scenario B. The institution and regulator would not identify an awarding gap and therefore fail to act despite significant inequalities.

Within the simulated university, the 60% threshold model complicates the relationship between gap and underlying distributions (Fig. 3B). Faculties A and C have equivalent differences in mean marks between Black and white students (3%), but the awarding gap in C is much larger ($A = 6\%$, $C = 21\%$). The mean mark in Faculty C is close to the 60% threshold, so relatively small differences in average marks give large changes in proportions meeting the threshold, whereas only students at the tail of the distribution affect the gap for Faculty A (Fig. 3).

Table 3 Effect of non-completion on awarding gap metric

Outcome	Reference group	Interest group
Students starting course	100	100
Students graduating	80	40
% completion gap	40%	
Students graduating with 1st/2i	60	30
% awarded 1st/2i	75%	75%
% awarding gap	0%	
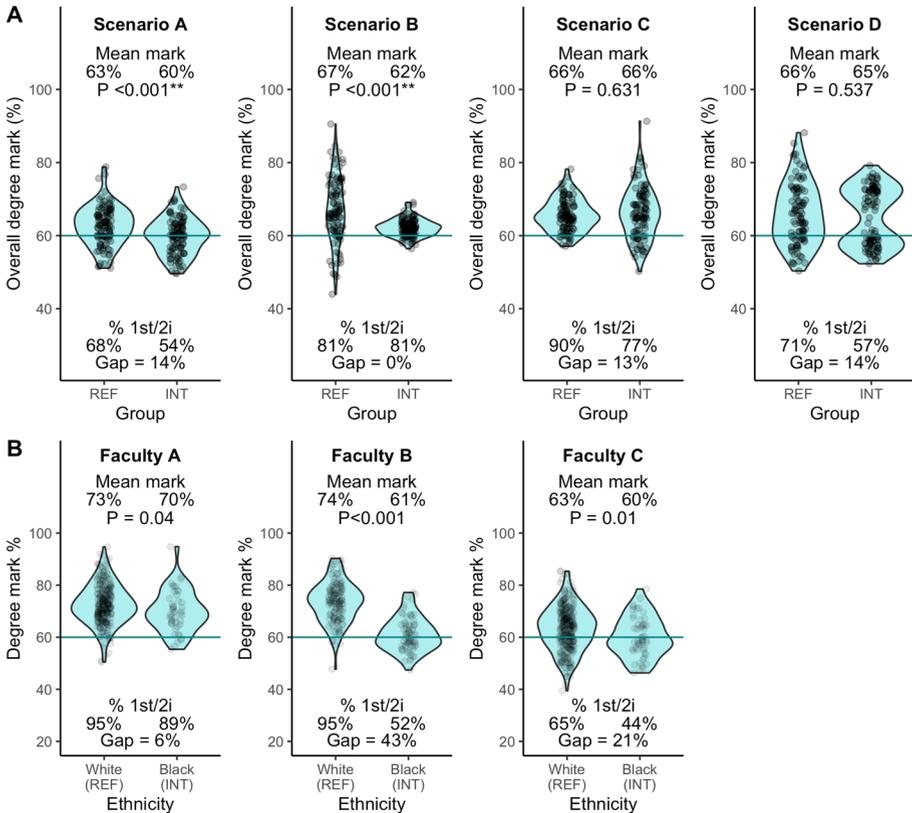
Fig. 3 Simulated data illustrates the (lack of) relationship between average marks and awarding gaps. **A** Hypothetical scenarios to illustrate the relationship between mark distributions and awarding gaps. **B** Black awarding gaps by faculty for the simulated university. REF, reference group; INT, interest group. Points represent individual students ($n = 100$ in each group), blue shapes the underlying distribution. Horizontal line indicates 60% threshold for a 2i. P values give results of Mann–Whitney tests for differences, ** indicates statistically significant difference

Technical critique 3: binary comparisons obscure subgroup differences

Gap-based measures rely on binary comparisons between groups. Typically, the group of interest is a population of disadvantaged students (e.g. disabled students), with anyone else (e.g. non-disabled students) defined as the reference. It is important to recognise that categorisation is highly subjective and potentially politically motivated. Pairwise classifications also create false binary models of society. For example, binary gender categories fail to recognise and respect trans and non-binary identities (Goldberg, 2018). Reliance on crude racial categories defined by white Europeans in educational policies can be viewed as a tool of white supremacy (Gillborn, 2007), and there can be significant tension between ethnic, racial, and national classifications (Arday & Mirza, 2018). Other factors such as socioeconomic status do not easily align with binary models.

Accepting the need for classification, implementation is challenging. For example, the UK definition of 'disabled' includes multiple conditions (e.g. wheelchair use, blindness,

autism, and dyslexia) (Bolton & Hubble, 2021; Disabled Students, 2022). Different disability groups have different student outcomes, so aggregation to 'disabled' obscures important subgroup differences (Office for Students, 2022c). Some students administratively classed as 'disabled' (e.g. neurodivergent students) may not identify as such (Shattuck et al., 2014). Categorisation also relies on formal declaration of data at an appropriate point in the student life cycle. A 'disability' category will under-represent disabled students if it does not capture those diagnosed during their studies and excludes those who choose not to declare their condition. Monitoring of gaps by sexual orientation or gender identity is likely to fail as individual status may be fluid and not declared (Rankin, 2006).

Categorisation issues are present within the simulated institution. There is a 16% 'BAME' (Black, Asian, Minority Ethnic) gap. However, consistent with UK data (AdvanceHE, 2021), gaps are larger for Black students (23%) than for Asian students (10%). The cynical university might report its more favourable 'BAME' gap and focus actions on Asian students while doing nothing to address outcomes for Black students. It is now considered poor practise to use 'BAME' within UK policy (Race Disparity Unit, 2022), but using subgroups is still problematic. For example, 'Asian' includes South Asian (Indian, Pakistani and Bangladeshi) and East Asian (Chinese and Korean) students whose outcomes are not equivalent (AdvanceHE, 2021; Office for Students, 2022c). Deciding on the most appropriate level of granularity requires consideration of both quantitative differences between groups and differential causal factors, and there is a trade-off between demographic resolution and sample size (see critique 6). Moving towards more granular categories means more nuance is required in defining the reference group. Should the reference group for Black Caribbean students be white students, all other students, all other non-White students or Black African students? Changing the reference group will change the size of the awarding gap reported, so it needs careful consideration.

Technical critique 4: single demographic based metrics obscure intersectional effects

Construction of pairwise gaps obscures intersectional interactions between demographic groups. In the simulated university, Black-disabled students have a larger awarding gap (31%) than might be expected from combining the Black (23%) and disabled gaps (4%) (Table 4).

The socially just institution wants to identify intersectional inequity, but the limitations of the pairwise metric require additional data analysis. However, this intersectional analysis may push statistical limits of data due to small cohorts (Hubbard, 2021).

Technical critique 5: institution level metrics obscure effects within institutions

Nationally, the size of the awarding gap differs by academic discipline, with smaller gaps in science, engineering and technology programmes (AdvanceHE, 2021). Granularity within the institution may therefore be important, particularly when interventions are likely to be made at local level. Two potential relationships between institution and local level gaps exist:

Within-area gaps: Within each area of the university, there is differential awarding between the reference and interest group. For example, the institutional Black awarding gap (23%) is mirrored in all three faculties (Table 2), although Faculty B has a larger gap (43%) than A (6%) or C (21%).

Table 4 Interaction between Black and disability gaps in the simulated university

Indicator	Demographic group	Race	Disability	Race & disability
Number of students	Reference	639 (White)	765 (Non-disabled)	505 (White & Non-disabled)
	Interest	167 (Black)	235 (Disabled)	54 (Black & Disabled)
% Awarded 1st/2i	Reference	83% (White)	78% (Non-disabled)	83% (White & Non-disabled)
	Interest	60% (Black)	74% (Disabled)	52% (Black & Disabled)
% Awarding Gap		**23%**	**4%**	**31%**

Between-area gaps: One area of the institution has both a higher proportion of students from an disadvantaged group and lower awarding rates than other areas, driving the institutional gap. In the simulated institution, there is an 11% mature student gap, but this is not seen within faculties. Faculty C produces the institutional gap, having the majority of mature students and low awarding rates for all students (Table 2).

With only institutional data available, local leaders may be ignorant of which model applies or the scale of gaps in their area. Without identifying the low awarding rate in Faculty C, the socially just institution could allocate significant resources to support mature students across the institution incorrectly assumed to be underperforming. The education lead for Faculty B will be unaware of the magnitude of their Black awarding gap, so may not act with the urgency required. The institutional metric is unhelpful at local level in both cases and potentially distracts from the most unequal outcomes. Statistically, this may represent an example of Simpson's paradox, whereby a trend observed within groups disappears or is reversed when those groups are aggregated together due to the influence of a covariate (Samuels, 1993).

Technical critique 6: awarding gaps lack statistical power, particularly for small cohorts

The awarding gap is presented as a percentage value with little context, making interpretation challenging. Is an awarding gap of 5% significant, or is this relatively favourable? Several other UK metrics are contextualised calculated via a *z*-score[2] (Gunn, 2018). The gap model also loses statistical power with very small cohorts (Hubbard, 2021). The UK regulator acknowledges this by redacting data for small cohorts (Office for Students, 2022a), but this effect is not necessarily appreciated within smaller institutions, those analysing data at local level or considering intersectional effects. Without acknowledging statistical realities, the gap opens itself up to criticism. Staff may be quick to point out statistical limitations of the gap (particularly in quantitative disciplines), distracting from addressing underlying inequity through questioning validity rather than focussing on action.

---

[2] A *z*-score describes the difference between an observation and a population, normalised to the population standard deviation; a *z*-score of 0 indicates outcomes are equivalent, while a *z*-score of 1 indicates the two differ by one standard deviation. Positive scores indicate the observation is higher than the population, and negative scores indicate it is lower.

Future directions: technical

In this article, I have identified a number of flaws in the use of institution level achievement gap metrics, including the perspectives of multiple stakeholders (Table 5). A number to summarise educational equity is of most value to the regulator for assessing fairness at scale and is arguably of least use to the student. The cynical university may 'game' the metrics, chasing or even manipulating the numbers rather than using them as a tool to genuinely tackle inequality (Oravec, 2019). The regulator must therefore play an active role in ensuring the metrics accurately capture equity to build trust and allow the identification of effective actions.

On the basis of the critiques above, I make the following technical recommendations for future development of the awarding gap. These primarily concern policy makers and the regulator, as their conceptions of the gap shape sector activity.

1. Move away from threshold metrics to statistical models that reflect differences in distributions
2. Develop measures that capture multiple outcomes, including non-completion
3. Develop metrics centring on the individual rather than the group, which would allow for intersectional analysis. For example, some institutions have developed 'Value Added Scores' for each student, with outcomes compared against predictions to identify unexplained areas of low awarding (Office for Students, 2022e)
4. To develop standard protocols for internal data analysis e.g. level of resolution, how to handle increased statistical noise of very small sample sizes

An illustration of a potential z-score based approach to address recommendations 1 and 2 is presented as Fig. 4. I simulate a second institution with six subjects, with data aggregated over five years to increase sample sizes (details in Supplementary Information). An outcome z-score compares grades in the subject area to the institutional average. An equity z-score compares outcomes of the interest group with that of the reference group for each subject. This partially mirrors the 'indicator' (outcome) and 'split indicator' (equity) structure used by the UK regulator in assessing other student outcomes (Office for Students, 2022b). Subject 1 has equitable grade outcomes (grade equity $z = -0.13$), but subject 2 has significant inequity (grade equity $z = -1.87$). This method captures inequity at any area of the distribution. For example, subject 3 has no awarding gap (-0.5%) using the 1st/2i threshold model, but a grade equity z-score of -1.2 indicates significant inequity. The use of z-scores also allows for composite metrics to be calculated (Song et al., 2013). Here, I create overall outcome and equity scores for each subject by adding together the respective scores for completion and grades. For example, subject 5 has low outcomes for all students (outcome $z = -2.17$), but outcomes are equitable (equity $z = -0.24$), while subject 6 has low and inequitable outcomes (outcome $z = -1.67$, equity $z = -4.57$). My example is at institution level, but it could equally be applied across the sector, enabling the regulator to identify institutions with particularly poor outcome and/or equity scores (AdvanceHE, 2021).

Table 5 Summary of institution level awarding gap critiques from the perspectives of stakeholders

	Regulator	Socially just Institution	Cynical institution	Local education lead	Student from disadvantaged group
Critique 1: Incomplete measure of student outcomes	May fail to identify inequity in other outcomes, resulting in sector wide inaction	May overlook significant inequity in other outcomes unless instructed to by regulator	Will not consider inequity in other outcomes unless instructed to by regulator	Modest gaps might distract from other more significant differences in outcomes	Has false sense of confidence in institution if gaps are modest but other outcomes differ
Critique 2: Binary threshold oversimplifies data	May fail to identify inequity in outcomes not captured by binary threshold, e.g. may ignore inequity in top/bottom of distribution	Potential tension between addressing inequity of outcomes and improving the threshold metric may undermine institutional strategy	Will not act to address inequity unless it is captured by the binary threshold, e.g. may ignore inequity in top/bottom of distribution	More likely to be looking at local assessment marks than classification outcomes, so lack of alignment may cause confusion	May become demotivated if peers have higher outcomes but told there is no gap or blame themselves for lower outcomes
Critique 3: No consideration of subgroup effects	May fail to identify inequity in subgroup outcomes, resulting in sector wide inaction	Cannot identify most significant inequity without additional internal data analysis. May fail to identify subgroups with greatest inequity	Only acts on gaps specified by regulator. May focus on subgroups that influence the metric rather than those with greatest inequity	Cannot identify most significant inequity without additional internal data analysis	Personal identity likely to be at subgroup level and may not align with administrative classification
Critique 4: Lack of intersectionality	May fail to identify inequity in intersectional outcomes, resulting in sector wide inaction	Cannot identify most significant intersectional inequity without additional internal data analysis	Only acts on gaps specified by the regulator so does not consider intersectionality	Cannot identify most significant intersectional inequity without additional internal data analysis	Personal experience likely to be intersectional and nuanced in ways not captured by metric
Critique 5: No consideration of within institution effects	Only interested in institutional level data and would argue institutions are responsible for internal analysis	May focus efforts in the wrong area, e.g. target specific departments with large gaps rather than low awarding levels elsewhere	Lacks information about where to most expediently target interventions	Will not know magnitude of gaps in their area so may be unaware of issues, or act unnecessarily when local outcomes are equitable	May incorrectly assume inequity of own outcomes from institution data even though local outcomes are equitable or even higher

Table 5 (continued)

	Regulator	Socially just Institution	Cynical institution	Local education lead	Student from disadvantaged group
Critique 6: Limited statistical power	Sector level data has lowest statistical noise—may fail to realise the limitations of data at provider/local level	Colleagues may use statistical noise as an argument to push back against acting, undermining efforts to reduce inequity	May use lack of statistical power as a justification for not taking action	Local level data likely to be statistically noisy due to small cohorts, so difficult to interpret and act appropriately	May assume large gaps represent significant inequity rather than statistical noise
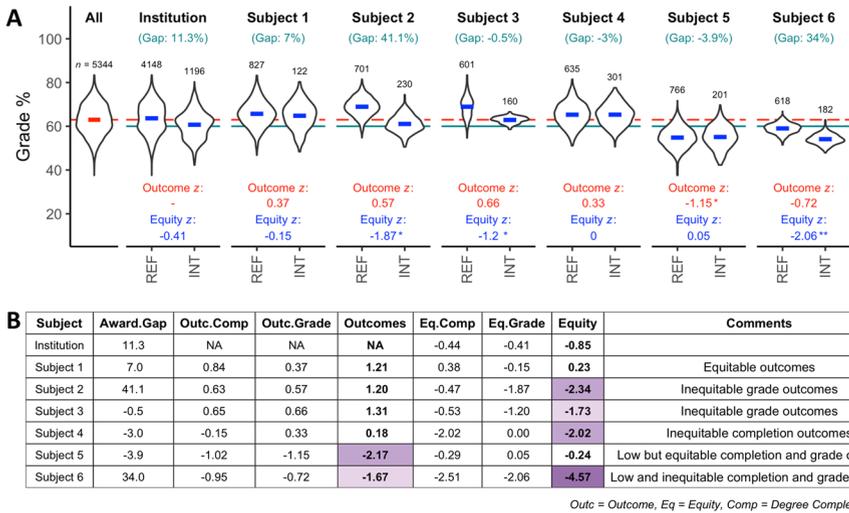
Fig. 4 Potential *Z*-score-based equity metrics. **A** *Z*-score method is applied to degree grades. Red line = institution mean, blue = REF group mean for each subject area, green line = 2i threshold. Number above shapes indicates aggregated 5-year cohort size. **B** Illustration of composite outcome and equity *z*-scores for each subject area. Composite metrics are the sum of the two measures. Shaded cells indicate subjects with significant negative composite *z*-scores; darker shading indicates more significant differences

Conclusion

The adoption of standardised awarding gap metrics in the UK has undoubtedly shone a light on systematic inequalities and prompted sector and institutional action (Universities UK and National Union of Students, 2019; Dickenson, 2021). However, if the sector is to use equity metrics, it has an ethical responsibility to use measures that are robust, accurate, meaningful to multiple stakeholders, and that cannot easily be gamed. I have demonstrated that the current UK awarding gap metrics are both theoretically and technically reductive. At best, the awarding gap gives an incomplete picture that requires significant additional institution/local level data gathering and analysis to identify the largest inequities. At worst, the threshold-based awarding gap is an active distraction from inequity. However, flaws in the metric do not mean that inequity does not exist. The sector needs to move beyond the flawed awarding gap model, devising robust metrics developed with multiple stakeholders that accurately identify both magnitude and cause of inequity to inform appropriate and effective action.

References

Adams, J. S. (1963). Towards an understanding of inequity. *Journal of Abnormal and Social Psychology, 67*(5), 422–436. https://doi.org/10.1037/h0040968

AdvanceHE (2021) *Ethnicity awarding gaps in UK higher education in 2019/20*. AdvanceHE. Available at: https://www.advance-he.ac.uk/knowledge-hub/ethnicity-awarding-gaps-uk-higher-education-201920 (Accessed: 2 July 2022).

Arday, J., & Mirza, H. S. (2018). In J. Arday & H. S. Mirza (Eds.), *Dismantling race in higher education: Racism, whiteness and decolonising the academy* (1st ed.). Springer International Publishing. https://doi.org/10.1007/978-3-319-60261-5

Bachan, R. (2017). Grade inflation in UK higher education. *Studies in Higher Education, 42*(8), 1580–1600. https://doi.org/10.1080/03075079.2015.1019450

Bertolin, J., & McCowan, T. (2022). The persistence of inequity in Brazilian higher education: Background data and student performance. In O. Tavares, C. Sá, C. Sin, & A. Amaral (Eds.), *Equity policies in global higher education: Reducing inequality and increasing participation and attainment* (pp. 71–88). Springer International Publishing. https://doi.org/10.1007/978-3-030-69691-7_4

Boliver, V., et al. (2022). Reconceptualising fair access to highly academically selective universities. *Higher Education, 84*(1), 85–100. https://doi.org/10.1007/s10734-021-00755-y

Bolton, P. and Hubble, S. (2021) *Briefing paper: Support for disabled students in higher education in England*. 8716. House of Commons Library. Available at: https://commonslibrary.parliament.uk/research-briefings/cbp-8716/ (Accessed: 2 July 2022).

Bolton, P., & Lewis, J. (2023). *Equality of access and outcomes in higher education in England*. House of Commons Library. https://researchbriefings.files.parliament.uk/documents/CBP-9195/CBP-9195.pdf

Brennan, J., & Naidoo, R. (2008). Higher education and the achievement (and/or prevention) of equity and social justice. *Higher Education, 56*(3), 287–302. https://doi.org/10.1007/s10734-008-9127-3

Britton, J., Walker, I., Waltmann, B., & Zhu, Y. (2022). *How much does it pay to get good grades at university?* Institute for Fiscal Studies. https://ifs.org.uk/sites/default/files/output_url_files/How-much-does-it-pay-to-get-good-grades-at-university.pdf

Brookover, W. B., & Lezotte, L. (1981). Educational equity: a democratic principle at a crossroads. *The Urban Review, 13*(2), 65–71. https://doi.org/10.1007/BF01956008

Burger, R. (2017). Student perceptions of the fairness of grading procedures: A multilevel investigation of the role of the academic environment. *Higher education, 74*(2), 301–320. https://doi.org/10.1007/s10734-016-0049-1

Cabral-Gouveia, C., Menezes, I., & Neves, T. (2023). Educational strategies to reduce the achievement gap: a systematic review. *Frontiers in Education, 8*. https://doi.org/10.3389/feduc.2023.1155741

Cagliesi, M. G., Hawkes, D., & Smith, S. (2023). Narrowing awarding gaps: the contributory role of policy and assessment type. *Studies in Higher Education*, 1–13. https://doi.org/10.1080/03075079.2023.2209597

Cairney, P., & Kippin, S. (2022). The future of education equity policy in a COVID-19 world: a qualitative systematic review of lessons from education policymaking [version 2; Peer review: 2 approved]. *Open Research Europe*, *1*, 78. https://open-research-europe.ec.europa.eu/articles/1-78

Colquitt, J. A., & Greenberg, J. (2003). Organizational justice: a fair assessment of the state of the literature. In J. Greenberg (Ed.), *Organizational behavior: the state of the science* (2nd ed., pp. 165–210). Lawrence Erlbaum Associates Publishers.

Cook, K. S., & Hegtvedt, K. A. (1983). Distributive justice, equity, and equality. *Annual review of sociology, 9*(1), 217–241. https://doi.org/10.1146/annurev.so.09.080183.001245

Crenshaw, K. (1989) 'Demarginalizing the intersection of race and sex: A black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics', *The University of Chicago legal forum*, 1989(1), p. 8. Available at: https://chicagounbound.uchicago.edu/uclf/vol1989/iss1/8/ (Accessed: 17 February 2020).

Davlembayeva, D. and Alamanos, E. (2022) *Equity theory*, *TheoryHub*. Available at: https://open.ncl.ac.uk/theories/5/equity-theory/ (Accessed: 10 May 2023).

Dickenson, J. (2021) *Here's what happened when I read every Access and Participation Plan*, *WonkHE*. Available at: https://wonkhe.com/blogs/heres-what-happened-when-i-read-every-university-access-and-participation-plan-in-england/ (Accessed: 6 December 2023).

Disabled Students UK (2022) *Going back is not a choice*. Disabled Students UK. Available at: https://disabledstudents.co.uk/not-a-choice/ (Accessed: 31 March 2022).

Edgar, S. (2022). The tricky concept of "Educational Equity" – In search of conceptual clarity. *Scottish Educational Review, 54*(1), 3–25. https://doi.org/10.1163/27730840-54010002

Espinoza, O. (2007). Solving the equity–equality conceptual dilemma: a new model for analysis of the educational process. *Educational Research*, 49(4), 343–363. https://doi.org/10.1080/0013188070 1717198

Flores-Crespo, P. (2007). Situating education in the human capabilities approach. In M. Walker & E. Unterhalter (Eds.), *Amartya Sen's Capability Approach and Social Justice in Education* (pp. 45–66). Palgrave Macmillan.

Gillborn, D. (2007) 'Education policy as an act of white supremacy: Whiteness, critical race theory and education reform', *Journal of Education Policy*, 20(4), pp. 485–505. Available at: https://www.tandfonline.com/doi/abs/https://doi.org/10.1080/02680930500132346

Goldberg, A.E. (2018) *Transgender students in higher education*. Williams Institute, UCLA. Available at: https://williamsinstitute.law.ucla.edu/publications/trans-students-higher-education/ (Accessed: 4 August 2023).

Gunn, A. (2018). Metrics and methodologies for measuring teaching quality in higher education: Developing the Teaching Excellence Framework (TEF). *Educational Review, 70*(2), 129–148. https://doi.org/10.1080/00131911.2017.1410106

Gutiérrez, R. (2008). Research commentary: A gap-gazing fetish in mathematics education? Problematizing research on the achievement gap. *Journal for Research in Mathematics Education, 39*(4), 357–364. https://doi.org/10.5951/jresematheduc.39.4.0357

Hazelkorn, E. (2007). The impact of league tables and ranking systems on higher education decision making. *Higher Education Management and Policy*, *19*(2), 1–24. https://www.oecd-ilibrary.org/content/paper/hemp-v19-art12-en

Hosier, M., & Hoolash, B. K. A. (2019). The effect of methodological variations on university rankings and associated decision-making and policy. *Studies in Higher Education*, 44(1), 201–214. https://doi.org/10.1080/03075079.2017.1356282

Hubbard, K. (2021) 'Using data-driven approaches to address systematic awarding gaps', *Doing Equity and Diversity for Success in Higher* [Preprint]. Available at: https://link.springer.com/chapter/https://doi.org/10.1007/978-3-030-65668-3_16.

Institute of Student Employment (2023) *ISE recruitment survey 2023*. Institute of student employment. Available at: https://ise.org.uk/page/ISE_Recruitment_Survey_2023&VID (Accessed: 8 December 2023).

Joseph-Salisbury, R. *et al.* (2020) *Race and ethnicity in British sociology*. British Sociological Association. Available at: https://www.britsoc.co.uk/media/25345/bsa_race_and_ethnicity_in_british_sociology_report.pdf.

Ladson-Billings, G. (2006). From the achievement gap to the education debt: Understanding achievement in U.S. schools. *Educational researcher, 35*(7), 3–12. https://doi.org/10.3102/0013189X035007003

Lamb, S. *et al.* (2020) *Educational opportunity in Australia: Who succeeds and who misses out?* Centre for International Research on Education Systems, Victoria University, for the Mitchell Institute: Melbourne. Available at: https://www.vu.edu.au/sites/default/files/educational-opportunity-in-australia-2020.pdf.

Leventhal, G. S. (1980). What should be done with equity theory? In K. J. Gergen, M. S. Greenberg, & R. H. Willis (Eds.), *Social exchange: Advances in theory and research* (pp. 27–55). Springer US. https://doi.org/10.1007/978-1-4613-3087-5_2

Levinson, M., Geron, T., & Brighouse, H. (2022). Conceptions of educational equity. *AERA Open, 8*, 23328584221121344. https://doi.org/10.1177/23328584221121344

Li, I. W., & Jackson, D. (2023). 'Influence of entry pathway and equity group status on retention and the student experience in higher education', *Higher Education* [Preprint]. *Available at*. https://doi.org/10.1007/s10734-023-01070-4

Liu, A. (2011). Unraveling the myth of meritocracy within the context of US higher education. *Higher Education, 62*(4), 383–397. https://doi.org/10.1007/s10734-010-9394-7

Locke, W., Verbik, L., Richardson, J. T. E., & King, R. (2008). Counting what is measured or measuring what counts? League tables and their impact on higher education institutions in England. http://oro.open.ac.uk/11799/

Msigwa, F. M. (2016). Widening participation in higher education: A social justice analysis of student loans in Tanzania. *Higher Education, 72*(4), 541–556. https://doi.org/10.1007/s10734-016-0037-5

Naylor, R., & Mifsud, N. (2020). Towards a structural inequality framework for student retention and success. *Higher Education Research & Development, 39*(2), 259–272. https://doi.org/10.1080/07294360.2019.1670143

Nussbaum, M. C. (2009). *Frontiers of justice: Disability, nationality, species membership*. Belknap Press. https://doi.org/10.4159/9780674041578

Nussbaum, M. C. (2011). *Creating capabilities: The human development approach*. Harvard University Press.

Office for Students. (2018). Access and participation plans. Office for Students. https://www.officeforstudents.org.uk/advice-and-guidance/promoting-equal-opportunities/access-and-participation-plans/. Accessed 15 Jul 2022

Office for Students (2022a) *Access and participation data resources: Supporting data user guide*, *Office for Students*. Available at: https://www.officeforstudents.org.uk/media/3c114aea-fc36-44da-9e59-394efa3c3057/access-and-participation-data-supporting-data-user-guide-2022.pdf (Accessed: 11 April 2022).

Office for Students (2022b) *Condition B3: Student outcomes*. Office for Students. Available at: https://www.officeforstudents.org.uk/publications/student-outcomes-and-teaching-excellence-consultations/student-outcomes/.

Office for Students (2022c) *Exploring student outcomes: Differences in continuation, completion and progression between students at English higher education providers*. Office for Students. Available at: https://www.officeforstudents.org.uk/media/6c99753c-2ccf-4dc6-93e1-cc712a987b43/exploring-student-outcomes.pdf.

Office for Students (2022d) 'KPM 7: Degree attainment by ethnicity - Office for students'. Available at: https://www.officeforstudents.org.uk/about/key-performance-measures/kpm-7-degree-attainment-by-ethnicity/ (Accessed: 19 May 2023).

Office for Students (2022e) *Using the value added score to challenge the BME attainment gap – BME Attainment Gap*, *OfS*. Available at: https://closingtheattainmentgap.co.uk/value-added/ (Accessed: 10 August 2023).

Office for Students (2020) *Access and participation data dashboard - Office for students*. Office for Students. Available at: https://www.officeforstudents.org.uk/data-and-analysis/access-and-participation-data-dashboard/ (Accessed: 31 March 2022).

Oravec, J. A. (2019). The "dark side" of academics? Emerging issues in the gaming and manipulation of metrics in higher education. *The review of higher education, 42*(3), 859–877. https://doi.org/10.1353/rhe.2019.0022

Pitman, T., Edwards, D., Zhang, L.-C., Koshy, P., & McMillan, J. (2020). Constructing a ranking of higher education institutions based on equity: Is it possible or desirable? *Higher Education, 80*(4), 605–624. https://doi.org/10.1007/s10734-019-00487-0

QAA (2019) *Annex D: Outcome classification descriptions for FHEQ Level 6 and FQHEIS Level 10 degrees*. QAA. Available at: https://www.qaa.ac.uk/docs/qaa/quality-code/annex-d-outcome-classification-descriptions-for-fheq-level-6-and-fqheis-level-10-degrees.pdf.

Race Disparity Unit (2022) *Why we no longer use the term 'BAME' in government*. Available at: https://equalities.blog.gov.uk/2022/04/07/why-we-no-longer-use-the-term-bame-in-government/ (Accessed: 4 August 2023).

Rankin, S. R. (2006). LGBTQA students on campus: Is higher education making the grade? *Journal of Gay & Lesbian Issues in Education, 3*(2–3), 111–117. https://doi.org/10.1300/J367v03n02_11

Rawls, J. (1971). *A theory of justice* (Original). UK: Harvard University Press. https://doi.org/10.2307/j.ctvjf9z6v

Rawls, J. (1985) 'Justice as fairness: Political not metaphysical', *Philosophy and Public Affairs*, 14(3), pp. 223–251. Available at: https://philpapers.org/rec/RAWJAF.

Robeyns, I. (2005). The capability approach: A theoretical survey. *Journal of Human Development, 6*(1), 93–117. https://doi.org/10.1080/146498805200034266

Salinas, D. (2018). Can equity in education foster social mobility? Organisation for Economic Co-Operation and Development (OECD). *Available at*. https://doi.org/10.1787/ca611a13-en

Samuels, M. L. (1993). Simpson's paradox and related phenomena. *Journal of the American Statistical Association, 88*(421), 81–88. https://doi.org/10.1080/01621459.1993.10594297

Sen, A. (1979). Equality of what? In McMurrin (Ed.), *Tanner lectures on human values* (pp. 197–220). Cambridge University Press UK.

Sen, A. (2009) *The idea of justice*. 1st edn. Belknap Pr.

Shattuck, P. T., et al. (2014). Disability identification and self-efficacy among college students on the autism spectrum. *Autism research and treatment, 2014*, 924182. https://doi.org/10.1155/2014/924182

Shukla, S. Y., et al. (2022). Reframing educational outcomes: Moving beyond achievement gaps. *CBE life sciences education, 21*(2), 2. https://doi.org/10.1187/cbe.21-05-0130

Song, M.-K., et al. (2013). Composite variables: When and how. *Nursing Research, 62*(1), 45–49. https://doi.org/10.1097/NNR.0b013e3182741948

Struyven, K., Dochy, F., & Janssens, S. (2005). Students' perceptions about evaluation and assessment in higher education: A review1. *Assessment and Evaluation in Higher Education, 30*(4), 325–341. https://doi.org/10.1080/02602930500099102

Tofallis, C. (2012). A different approach to university rankings. *Higher Education, 63*(1), 1–18. https://doi.org/10.1007/s10734-011-9417-z

United Nations (2015) *Goal 4, UN department of economic and social affairs*. Available at: https://sdgs.un.org/goals/goal4 (Accessed: 1 June 2023).

Universities UK and National Union of Students (2019) *Black, Asian and Minority Ethnic student attainment at UK universities: #closingthegap*. Universities UK; National Union of Students. Available at: https://www.universitiesuk.ac.uk/policy-and-analysis/reports/Documents/2019/bame-student-attainment-uk-universities-closing-the-gap.pdf.

Vaughan, R. (2007). Measuring capabilities: An example from girls' schooling. In M. Walker & E. Unterhalter (Eds.), *Amartya Sen's Capability Approach and Social Justice in Education* (pp. 109–130). Palgrave Macmillan.

Walker, M., & Unterhalter, E. (2007). The capability approach: Its potential for work in education. In M. Walker & E. Unterhalter (Eds.), *Amartya Sen's Capability Approach and Social Justice in Education* (pp. 1–18). Palgrave Macmillan.

Wilson-Strydom, M. (2015). University access and theories of social justice: Contributions of the capabilities approach. *Higher education, 69*(1), 143–155. https://doi.org/10.1007/s10734-014-9766-5

Wong, B., ElMorally, R., & Copsey-Blake, M. (2021). "Fair and square": What do students think about the ethnicity degree awarding gap? *Journal of Further and Higher Education, 45*(8), 1147–1161. https://doi.org/10.1080/0309877X.2021.1932773