Beyond education by ranking

Let's not return to 'normal'

Phil Taylor

Abstract

The cancellation of public examinations in England during the coronavirus pandemic drew attention to a long-standing educational concern. Grading and ranking students, in various ways, has taken place for many years, but in summer 2020 this process was shared between teachers and, initially, an 'algorithm'. Maintaining standards and consistent grade distributions is a feature of the exam system in 'normal' times. This article considers why exam grades are (roughly) normally distributed, tracing origins of bell-curve thinking, to suggest that we should not be returning to this kind of 'normal'.

Keywords: Examination; assessment; standardisation; evaluation; bell-curve thinking; pandemic

A critical moment of the coronavirus pandemic for education in England was the cancellation of the summer 2020 exams and the decision to calculate GCSE and A-level grades using an 'algorithm'. The government and exams regulator Ofqual stipulated that students should be awarded the grades they were 'most likely to achieve had exams gone ahead', ensuring similar grade distributions to previous years (Ofqual, 2020, p5). Schools provided a 'centre assessment grade' and ranking for each entry, from which exam boards produced 'calculated grades', taking into account past school performance. When calculated A-level results were released on 13 August, nearly 40 per cent of teacher assessments had been downgraded, with historically lower-performing schools disproportionately affected (BBC, 2020). On 17 August, following widespread outcry, the government announced that centre assessment grades would be reinstated and the algorithm for GCSEs grades, due for release the following week, was scrapped.

These unprecedented circumstances focused public attention on the grading algorithm which reflected features of exam standardisation in 'normal' times. The pandemic has re-exposed deep educational inequalities, disproportionately affecting some communities. While many are anxious to 'return to normal', this article suggests that ongoing efforts to maintain consistent grade distributions are partially productive of these inequalities, indicative of systemic incoherence between educational aims and measured outcomes. The reasons why exam grades tend towards normal distribution are discussed, tracing the origins of bell-curve thinking. I also offer reflections on

DOI: 10.3898/FORUM.2021.63.2.02

personal experience, by way of illustration and accounting for my own complicity in this process of grading.

Growing awareness of the normal curve

My age seven school report, kept by my parents, states my 'position in class' as twenty (out of thirty-nine). In the early 1970s, children were ranked in their class and classes were streamed, so my 'position in year group' (twenty-two out of ninety-five) is also reported. 'Average' appears frequently in the accompanying comments; my arithmetic was 'slow', English 'careless and rather inaccurate', writing 'irregular' and composition 'dull'. By the following summer, I was 'improving' in some areas, even 'good' in art and craft, but composition was 'sometimes a little vague'. Despite these small signs of improvement, my 'positions' in class and year had slipped to twenty-four and twenty-five respectively. Next year the report format changed – no rankings and a simple grading rubric (or algorithm), printed on the inside cover, relating to the 'whole of the child's age-group in the school'. The gradings were: A = exceptionally good, B = above average, C = average, D = below average, noting 'plus and minus signs may be added'. Most of my gradings through primary school were just 'above average', and typical comments included 'satisfactory', 'improving', 'slow', 'steady' and 'easily distracted'. I cannot recall what this meant as a child (I was easily distracted), but children make comparisons too. My older siblings gained higher rankings and gradings; I knew my place and my parents' aspirations. I am aware that a story was written for me, which to some extent has transpired. However hard we worked, the simple algorithm, referenced to the age-group, guaranteed a spread of grades and labels.

Around the time I started primary school, Benjamin Bloom (1968) wrote of conditioning to the normal curve and its self-fulfilling prophecies, differentiating students by grading 'even if the differences are trivial in terms of the subject matter' (p2). He reasoned that education as a purposive activity should generate a very different distribution of outcomes, if students learn what is taught. Interestingly, Bloom was arguing for *mastery learning*, believing that the vast majority of students can achieve highly with appropriate teaching and support, regardless of their prior attainments and backgrounds. This view is often espoused by policymakers and educators, and a discourse of mastery has emerged in England in recent years, particularly in mathematics education (Boylan, 2019). However, when students reach high-stakes Key Stage 2 (age eleven) tests, GCSE (age sixteen) and A level (age eighteen) exams, application of the normal curve maintains standards over time. Mastery, measured by exam, is thereby constrained and for some students impossible.

Many complex personal and social factors contribute to exam performance, apart from subject learning at school, including personality traits (Borghans *et al.*, 2016), aspirations (Khattab, 2015), test anxiety and stress (Room and Soane, 2019), socio-economic status (Morris *et al.*, 2016), and 'occasion-related factors', such as revised topics coming up, tiredness, room conditions and family circumstances (Baird and Black, 2013, p15). I experienced severe hay fever as a teenager and most of my O- and A-level exams took place during the pollen season. I recall several exam-hall occasions, with eyes itching, nose streaming and packs of tissues, or jaded from antihistamines. Without over-attributing my mixed exam outcomes to allergy, I gained better grades in one taken early and two re-sat at other times of year. Contingencies and 'luck of the draw' partly define the system, where 'we accept that how a person performs on the occasion of the examination is assessed, rather than believing in an underlying person's ability that can be objectively measured' (Baird and Black, 2013, p16). However, perhaps this is not accepted by all.

As a secondary school teacher in the 1980s I was only vaguely aware of national grade distributions. Like most teachers I knew, I focused on enabling my own students to learn and enjoy the subject, leaving exam preparation until the end of the course. I recall a collective yet unspoken commitment to teaching as well as we could, but responsibility for success rested with students themselves. This changed with the introduction of exam league tables and school inspections in 1992 for monitoring and accountability of teachers and schools (Mansell, 2007). As a school leader in the 1990s, when comparing students' results with national figures I noticed minimal year-on-year changes in grade proportions in most subjects (less than 2 per cent). In local authority advisory roles during the 2000s, I became more aware of predictive models for projecting outcomes from one stage of schooling to the next (Taylor, 2011) and the pressures on schools to improve league-table rankings. Annual increases, and later decline, in headline age-sixteen thresholds (percentage 5+ A* to C grades) resulted from the range of 'equivalent' qualifications allowed (Taylor, 2015). For a time, after moving into higher education, I neglected these issues until a colleague mentioned subject-grade distributions and I looked more closely.

Why are grade distributions roughly 'normal'?

Grade distribution graphs rarely appear in published results, but readily available tabulated figures, when plotted, are usually bell-shaped. The graph below (Fig.1) shows grade distributions in GCSE mathematics since 2017, when 9 to 1 gradings were introduced (replacing A* to G). The impact of the 2020 debacle (darker columns) skews the roughly normal distribution upwards by less than 2 per cent on any grade. Parents are informed that grade 4 is a 'standard pass', 'a credible achievement' and 'a passport to future study and employment', however if this 'minimum level' is not reached in English and/or maths then the subject(s) must be retaken during post-sixteen study (DfE, 2019,

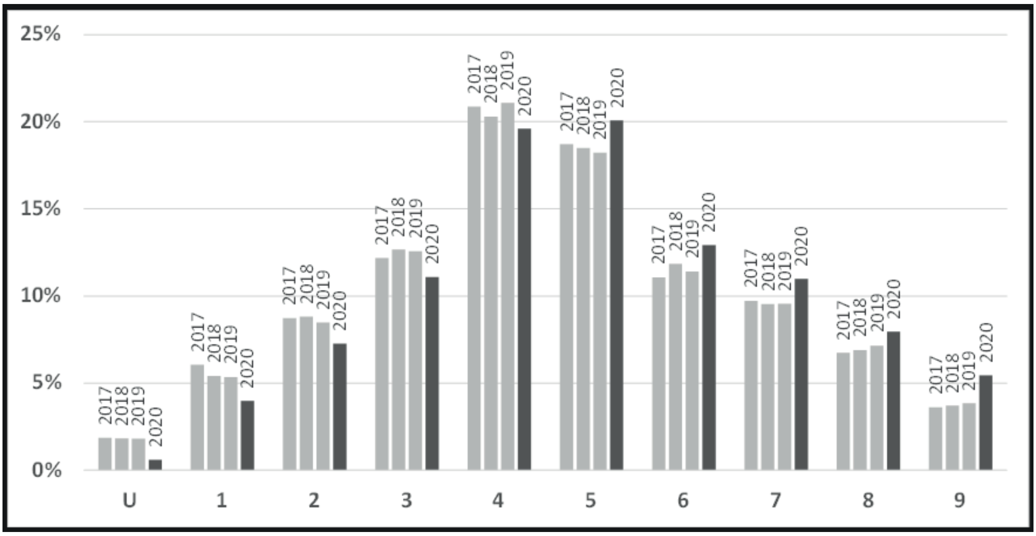

Figure 1. Mathematics GCSE grade distribution 2017-2020 (per cent excluding absences)
Source: https://www.gov.uk/government/collections/statistics-gcses-key-stage-4

p2). Unsurprisingly, among the nearly one-third of students (around 180,000 each year) who do not attain at least grade 4, many are demotivated (Higton *et al.*, 2017). While it is claimed that GCSEs are not norm-referenced, a minimum expectation itself establishes a norm (Fendler and Muzaffar, 2008).

Public examinations are reasonably expected to assess students on curriculum content covered during the preceding period of study, not against each other, without presumption over grade distributions (Baird and Black, 2013). This implies criterion-referencing, whereby students' demonstrable knowledge and skills are defined in absolute terms, not relative to other candidates (Stringer, 2012). GCSE and A-level exams have never been wholly or strongly criterion-referenced, due to variations in perceived difficulty and demands for comparable standards over time; nor have they been entirely norm-referenced, maintaining fixed year-on-year grade distributions (Newton, 2020, 2021). Instead, as Newton explains:

> We might refer to the overarching principle that has always underpinned the maintenance of exam standards in England as attainment-referencing. This involves: ranking students in terms of their overall level of attainment in a subject area, on the basis of the marks they achieve in their subject exam; and then classifying students in terms of whether their level of attainment is high enough to be awarded a particular grade (Newton, 2021, p14).

Attainment-referencing weakens links to published grading criteria (Newton, 2020) and is 'compensatory', such that 'better performance in one area can compensate for poorer performance in another' (Jadhav, 2017, np). This combines the quality of response to each test item and its perceived difficulty, along with any errors introduced in the process (Stringer, 2012; Newton, 2020). Summating marks on many individual questions across several exam papers is essentially an averaging process, increasing the likelihood that total scores are normally distributed (Dudley-Marling, 2020, p203). This is reinforced by setting grade boundaries, which in England 'has always relied upon a judicious balance of expert judgement of performance evidence and statistical expectations of cohort attainment' (Newton, 2020, p11). Examiner judgement is usually exercised only at 'judgemental grades' (7, 4 and 1 for GCSE; A and E for AS/A level), with the rest 'set arithmetically' (Ofqual, 2020, p295). For GCSE, statistical expectations are derived from Key Stage 2 tests as the 'best predictor', combined with grade distributions from the previous year's exam to construct 'prediction matrices' (Newton, 2020, p12). This invokes 'the similar cohort adage' that 'if the cohort hasn't changed much, then don't expect the pass rate to change much either', also known as 'the curve method', to establish 'statistically expected boundaries (SEBs)' (*ibid*, p11). Maintaining year-on-year grade distributions has also been termed 'cohort-referencing' (Stringer, 2012) and 'the principle of comparable outcomes' (Ofqual, 2017, p4). Whatever the nomenclature, maintaining standards means that exam results are not only determined by students' subject knowledge and skills, plus performance on the day. There is circularity in using past performance to shape current outcomes after pronouncing it the best predictor, and assumptions of normally distributed underlying general ability are explicit in some exam modelling (e.g. Benton, 2018).

It is worth contemplating the likely standardisation response to more students attaining sufficient total marks to reach the crucial grade 4 boundary. Concern over grade inflation prompts a 'tendency to rely heavily upon statistical expectations and to place correspondingly less confidence in examiner judgement' (Newton, 2021, p21). Thus, although exams ostensibly assess grasp of curriculum content, statistical expectations utilising prior attainment patterns would establish the 'correct' grade boundaries. The statement that 'we would not expect a cohort of learners to become *substantially* better at learning from just one year to the next' (Newton, 2020, p11, original emphasis) might surprise students, teachers and school leaders, who collectively strive for exactly that. It implies that exams do not measure what students have learned, but how good they are at learning, and that their efforts would either be fruitless or cancel out across the whole cohort as they compete for the available grades. Predicting future outcomes from prior attainment, then measuring to what extent expectations are exceeded or undershot, is a zero-sum value-added process, with roughly equal numbers of winners and losers at student and school levels, rendering system-wide improvement impossible (Taylor, 2015). Such a system guarantees failure for some, naturalises ranking, and perpetuates beliefs in elitism and 'bell-curve thinking' (Fendler and Muzzafar, 2008, p82; Dorling, 2015, p46).

Bell-curve thinking

Bell-shaped or normal distributions arise from observing certain real-life measures, including heights and weights of populations, that cluster around the mean and tail-off either side. They originate from plotting long-run results of coin-tossing games and the spread of errors in astronomical measurements during the 1700s and 1800s (Hacking, 1990, p106). The best estimate of a measurable *physical* quantity was found to be the mean of many measurements, with inevitable errors, hence the normal distribution was known as the 'error curve' (MacKenzie, 1981, p56). Crucially, the error curve was appropriated by social scientists, notably Adolphe Quetelet and Francis Galton, to further their interests in the 'average man' and eugenics respectively (MacKenzie, 1981; Hacking, 1990). Quetelet turned physical measurement into 'a theory of measuring ideal or abstract properties of a population' such that 'they became real quantities' (Hacking, 1990, p108). Galton argued by analogy that hitherto unmeasurable qualities, 'such as cleverness, morality, wit, and civility', could be counted and standardised (Fendler and Muzzafar, 2008, p74). By 1869, Galton (cited by MacKenzie, 1981, p57) had decided that:

> there must be a fairly constant average mental capacity in the inhabitants of the British Isles, and that deviations from that average – upwards towards genius, and downwards towards stupidity – must follow the law that governs deviations from all true averages.

Treating intangible qualities as physical quantities, or *reification*, is central to debates over intelligence and cognitive ability. Bell-curve thinking promotes 'IQism', or belief in the intelligence quotient (IQ) as a valid measurement scale for ranking general ability (Dorling, 2015, p51). The modern-day equivalents, cognitive ability tests, are still deployed for school entrance exams, grouping students and predicting their GCSE and A-level outcomes. The first IQ tests, designed by Binet in the early 1900s, aimed to identify those needing additional support, expressly not to reify, measure and rank innate intelligence (Gould, 1996, p185). However, IQ testing was taken up by those seeking heredity in human qualities, including the UK eugenics movement founded by Galton and involving statisticians Pearson, Spearman, Fisher and Burt (MacKenzie, 1981). Cyril Burt was instrumental in promoting intelligence as fixed and innate, ushering in the 11-plus exam for secondary school selection in the 1950s and 1960s (Dorling and Tomlinson, 2019). (I did not take the 11-plus as it was phased-out where I lived several years earlier.)

While the selective breeding proposed by eugenics is from a bygone era, its legacy survives in IQism, selective schooling, classifying people and the 'structuring of inferiority' (Dorling and Tomlinson, 2016, p73). Saini (2019) charts how beliefs in racial superiority and inferiority are entwined with biological determinism, despite compelling evidence and experience of their social construction. A prominent example is *The Bell Curve* (Herrnstein and Murray, 1994), which claims that differences in average IQ between ethnic groups are partly explained by inherited intelligence. Gould (1996) exposes the lack of justification for this claim, and 'deep and instructive fallacies' in the reification of IQ to a 'unitary, linearly ranked, innate, and minimally alterable intelligence' (p26). A common factor, presumed to be general intelligence (Spearman's g), is derived from mathematical constructs (factor analysis) of test score correlations (Gould, 1996). It is fictional and self-defining as a hypothesised cause of test score variation that, if existing at all, may not have quantitative measurable form (Michell, 2020). As Michell explains, 'test items, being irrelevant to people's life circumstances, may be engaged by different motives within different people' (p7). The common factor associated with IQ tests could just as easily be chosen as a measure of how bothered/motivated/affected/*(insert adjective here)* people are by these tests.

Assumptions of innate and fixed general intelligence are further undermined by the 'Flynn effect', tracking increasing IQ scores over time (Dorling and Tomlinson, 2016, p73), as well as concerted efforts to develop it, exemplified by services purporting to prepare children for the 11+ exam. Heritability adds to confusion, estimating percentage genetic contributions to *differences between individuals* in attributes such as IQ, but meaningless for particular individuals, as nature and nurture 'cannot be disentangled in any person's life history' (Rose, 2014, p2). Counterintuitively, heritability depends on environment, which must be experienced identically for all differences between individuals to be explained genetically (Rose, 2014). Variation in the many socio-economic and cultural factors comprising environment reduces heritability, particularly among those most disadvantaged (*ibid*). As Gould concluded twenty-five years ago:

> We pass through this world but once. Few tragedies can be more extensive than the stunting of life, few injustices deeper than the denial of an opportunity to strive or even to hope, by a limit imposed from without, but falsely identified as lying within (Gould, 1996, p50).

Finally, real-life phenomena for which observations are not normally distributed include earthquakes, wealth, market fluctuations, heart rhythms, election results and learning events (Davis and Sumara, 2010, p40). Many of these follow power-law distributions, characterised by many events or occurrences of small magnitude and very few with large effect (*ibid*). Learning, as a complex adaptive process, is such a phenomenon, whereby knowledge and skills are accumulated and applied in many small steps over time, punctuated by occasional highly transformative experiences. Complexity thinking looks beyond the predictive determinism of bell-curve thinking that defines 'normal' and measures deviation from it, embracing 'a counter-normative sensibility, whereby diversity is understood to be inevitable and necessary ... the source of a system's flexible responsiveness – its intelligence, as it were' (Davis and Sumara, 2010, p50).

Assessment beyond the normal curve

The continuation of the pandemic has again led to the cancellation of exams in summer 2021. To avoid repeating the 2020 debacle, government has instructed schools to assign grades, based only on course content covered (Ofqual, 2021). To assist teachers, exam boards are producing topic-based test materials and additional grade descriptors for each subject, signalling a welcome shift towards criterion-referencing. There is no proposed standardisation algorithm, however a form of cohort-referencing is retained by requiring schools to check their grade distributions against past performance trends, subject to internal and external quality assurance.

Longer-term, Rethinking Assessment (www.rethinkingassessment.com) is campaigning for the end of GCSEs and seeking alternatives. What should be avoided is replacing one system of scoring, ranking and classifying students with another. Resisting and rejecting bell-curve thinking and conditioning to the 'normal' would be more ambitious, abandoning the pretence that exam scores and rankings are also valid measurements of teacher and school effectiveness. A starting point is to acknowledge that 'public examinations are not objective measures of a pre-existing physical reality: they are socially constructed' (Baird and Black, 2013, p14).

A more generative approach is offered by 'relational evaluation', placing relationships and values at the heart of engagement in learning, supported by dialogue, appreciation and co-inquiry (Gill and Gergen, 2020). This prospect might be too radical and idealistic for some educators, particularly those who seek to prepare students for a competitive world in which they have already succeeded. However, by sticking to 'normal' we risk perpetuating pervasive and divisive inequalities that have become even more conspicuous during the pandemic.

References:

Baird, J. A. and Black, P. (2013) 'Test theories, educational priorities and reliability of public examinations in England', *Research Papers in Education*, vol. 28, no. 1, 5–21. DOI: 10.1080/02671522.2012.754224.

BBC (2020) *A-levels and GCSEs: How did the exam algorithm work?*, available at: https://www.bbc.co.uk/news/explainers-53807730 (accessed 31 March 2021).

Benton, T. (2018) 'How many students will achieve straight grade 9s in reformed GCSEs?', *Research Matters*, no. 25, 28–36, available at: https://www.cambridgeassessment.org.uk/Images/476536-how-many-students-will-get-straight-grade-9s-in-reformed-gcses-.pdf (accessed 31 March 2021).

Bloom, B. S. (1968) 'Learning for Mastery', *Evaluation Comment*, Center for the Study of Evaluation of Instructional Programs, University of California at Los Angeles., vol. 1, no. 2, 1–12.

Borghans, L., Golsteyn, B. H. H., Heckman, J. J. and Humphries, J. E. (2016) 'What grades and achievement tests measure', *Proceedings of the National Academy of Sciences*, vol. 113, no. 47, 13354–13359. DOI: 10.1073/pnas.1601135113.

Boylan, M. (2019) 'Remastering mathematics: Mastery, remixes and mash ups', *Mathematics Teaching*, no. 266, pp. 14–18, available at: https://www.atm.org.uk/write/MediaUploads/Journals/MT266/MT26605.pdf (accessed 31 March 2021).

Davis, B. and Sumara, D. (2010) 'Decentralizations and Redistributions: A Complex Reading of Normality', Dudley-Marling, C. and Gurn, A. (eds), *The myth of the normal curve*, Disability studies in education, New York: Peter Lang, 39–51.

DfE (2019) *GCSE factsheet for parents*, Department for Education, available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/800507/GCSE_factsheet_for_parents__final_.pdf (accessed 31 March 2021).

Dorling, D. (2015) *Injustice: why social inequality still persists*, Revised ed., Bristol: Policy Press.

Dorling, D. and Tomlinson, S. (2016) 'The Creation of Inequality: Myths of Potential and Ability', *Journal for Critical Education Policy Studies*, vol. 14, no. 3, 56–79, available at: http://www.jceps.com/wp-content/uploads/2016/12/14-3-3.pdf.

Dudley-Marling, C. (2020) 'The Tyranny of the Normal Curve: How the "Bell Curve" Corrupts Educational Research and Practice', Allen, D. M. and Howell, J. W. (eds), *Groupthink in Science: Greed, Pathological Altruism, Ideology, Competition, and Culture*, Cham: Springer International Publishing. DOI: 10.1007/978-3-030-36822-7.

Fendler, L. and Muzaffar, I. (2008) 'The History of the Bell Curve: Sorting and the Idea of Normal', *Educational Theory*, vol. 58, no. 1, 63–82. DOI: 10.1111/j.1741-5446.2007.0276.x.

Gill, S. and Gergen, K. (2020) 'Educational Evaluation: A Relational Perspective', McNamee, S., Gergen, M. M., Camargo-Borges, C., and Rasera, E. F. (eds), *The Sage handbook of social constructionist practice*, 1st edition., Thousand Oaks: SAGE Inc, 402–411.

Gould, S. J. (1996) *The mismeasure of man*, Rev. and expanded., New York: Norton.

Hacking, I. (1990) *The taming of chance*, Ideas in context, Cambridge [England]; New York: Cambridge University Press.

Herrnstein, R. J. and Murray, C. A. (1994) *The bell curve: intelligence and class structure in American life*, New York: Free Press.

Higton, J., Archer, R., Dalby, D., Robinson, S., Birkin, G., Stutz, A., Smith, R. and Duckworth, V. (2018) *Effective practice in the delivery and teaching of English and Mathematics to 16-18 year olds.*, London, Department for Education.

Jadhav, C. (2017) *Mythbusting: 3 Common Misconceptions*, available at: https://ofqual.blog.gov.uk/2017/03/17/mythbusting-3-common-misconceptions/ (accessed 31 March 2021).

Khattab, N. (2015) 'Students' aspirations, expectations and school achievement:

what really matters?', *British Educational Research Journal*, vol. 41, no. 5, 731–748. DOI: 10.1002/berj.3171.

MacKenzie, D. A. (1981) *Statistics in Britain, 1865-1930: the social construction of scientific knowledge*, Edinburgh: Edinburgh University Press.

Mansell, W. (2007) *Education by Numbers: The Tyranny of Testing*, London: Politico's Publishing Ltd.

Michell, J. (2020) 'The fashionable scientific fraud: Collingwood's critique of psychometrics', *History of the Human Sciences*, vol. 33, no. 2, 3–21. DOI: 10.1177/0952695119872638.

Morris, T., Dorling, D. and Davey Smith, G. (2016) 'How well can we predict educational outcomes? Examining the roles of cognitive ability and social position in educational attainment', *Contemporary Social Science*, vol. 11, no. 2–3, 154–168. DOI: 10.1080/21582041.2016.1138502.

Newton, P. E. (2020) *Maintaining Standards - During normal times and when qualifications are reformed*, Ofqual, available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/936340/Maintaining_Standards.pdf (accessed 30 March 2021).

Newton, P. E. (2021) 'Demythologising A level Exam Standards', *Research Papers in Education*, 1–32. DOI: 10.1080/02671522.2020.1870543.

Ofqual (2020) *Awarding GCSE, AS, A level, advanced extension awards and extended project qualifications in summer 2020: interim report*, Ofqual, available at: https://www.gov.uk/government/publications/awarding-gcse-as-a-levels-in-summer-2020-interim-report (accessed 30 March 2021).

Ofqual (2021) *Decisions on how GCSE, AS and A level grades will be determined in summer 2021*, Ofqual, available at: https://www.gov.uk/government/consultations/consultation-on-how-gcse-as-and-a-level-grades-should-be-awarded-in-summer-2021 (accessed 30 March 2021).

Roome, T. and Soan, C. A. (2019) 'GCSE exam stress: student perceptions of the effects on wellbeing and performance', *Pastoral Care in Education*, vol. 37, no. 4, 297–315. DOI: 10.1080/02643944.2019.1665091.

Saini, A. (2020) *Superior: the return of race science*, London: 4th Estate Ltd.

Stringer, N. S. (2012) 'Setting and maintaining GCSE and GCE grading standards: the case for contextualised cohort-referencing', *Research Papers in Education*, vol. 27, no. 5, 535–554. DOI: 10.1080/02671522.2011.580364.

Taylor, P. (2011) 'All above average: expectations and the use of prior attainment data to estimate future attainment in pupil target-setting.', *Online Education Research Journal (OERJ)*, available at: http://community.dur.ac.uk/p.b.tymms/oerj/publications/18.pdf.

Taylor, P. (2015) 'All Above Average: secondary school improvement as an impossible endeavour', *FORUM*, vol. 57, no. 2, 239-250. DOI: 10.15730/forum.2015.57.2.239.

Phil Taylor's work in higher education for the last ten years has been dedicated to supporting the professional growth, organisational development and self-determined learning of teachers and educational leaders, particularly through practice-based inquiry and further study. Pursuing a lifelong career in education, he taught maths, computing and ICT for fourteen years in two large London comprehensive schools, then settled in the Midlands and worked for eight years in advisory roles for two local authorities, before moving into higher education. Currently he is Assistant Professor in Educational Leadership and Management, at the University of Nottingham.

Phil.Taylor@nottingham.ac.uk