

Knowledge, Expertise and Policy in the Examinations crisis in England

Jenny Ozga, Jo-Anne Baird, Luke Saville, University of Oxford, Margaret Arnott, Niclas Hell University of the West of Scotland

Abstract

The Covid 19 Pandemic suspended established practices that, in normal times, are seen as central to the functioning of education systems. For example, in England, school closures led to the cancellation of national examinations in 2020, and their attempted replacement with an algorithmic model. Following public outcry about what were seen as the unjust effects of the application of that model, there was a very public policy reversal, and examination grades were awarded on the basis of moderated teacher assessments or Centre Assessed Grades (CAGs), resulting in substantial grade inflation. This paper draws on research that investigated the actors involved in examinations policy in this period and focuses especially on the sources of expertise and the kinds of knowledge that were mobilised-or not-in the decision to cancel examinations, to develop the algorithm, and to revert to Centre Assessed Grades.

Keywords: Covid 19, Examinations, Policy, Expertise, Knowledge

Introduction

To understand the significance of the policy reversal discussed here, it is first necessary to grasp the scale of examinations, their place in the English system, and political concerns about standards. We stress here that our focus in this paper is not on debating the rights and wrongs of examinations, but rather on the knowledge and expertise available to policymakers in dealing with a crisis that arose when examinations could not take place, in a system that is heavily reliant on them. Qualifications and examinations are fundamental to English education, with some claiming that the English system suffers from ‘diploma disease’ because of its reliance on qualifications (Dore, 1976). A-level examinations, usually taken at age 18, are the main qualification at school-leaving age for entry to employment, or further or higher education. Together with the examinations taken at age 16, the examination boards deal with over 26 million examination scripts and items of coursework. England has a quasi-market among examination boards; five of them are accredited by the government regulator, the Office for Qualifications and Examinations Regulation (Ofqual), to offer A-levels (Opposs et al., 2020). Teachers are involved in the marking of coursework and practical work in normal years and there are systems of moderation in place that require re-marking samples of work.

Qualification reform is frequent in England, however the period before the pandemic was relatively stable, as the UK had a Conservative government in power since 2010-first in coalition, then as the majority party. The Conservatives came to office expressing significant concerns about educational standards. The then Secretary of State for Education, Michael Gove, considered that England was not performing well in international league tables such as OECD’s PISA, that there had been grade inflation in the national qualification results and that there was underhand competition between exam boards on standards (HC 141-I, Appendix 1, 2012). Ofqual introduced a system (comparable outcomes; see Ofqual, 2015) to monitor examination standards between years and across exam boards. The Secretary of State initiated a series of reforms, including the removal of coursework from a large number of A-levels and other qualifications, in order to reduce the amount of cheating and illicit support that students were receiving from parents, teachers and commercial organisations (Meadows and Black, 2018). Coursework is currently only included in creative and practical subjects (art, music, design & technology, physical education, drama, food preparation and nutrition), though there is a certification of practical science at A-level too. Another argument for minimising teacher assessment was that the accountability system for schools produced a conflict of interest for

teachers, with better outcomes reflecting well on schools and affecting teacher careers. Successive Conservative UK governments saw examinations as more rigorous, valid and reliable forms of assessment and also as reducing the assessment burden in comparison with coursework. Standards were more tightly controlled by the regulator from 2010. Using a 'comparable outcomes' statistical technique to predict and monitor standard setting, outcomes largely became stable over time (Taylor and Opposs, 2018). This technique also meant that students' progress between exams taken at age 16 and those taken at 18 became similar between exam boards. This, then, is the policy background to the examinations crisis of 2020, as schools were closed, and examinations cancelled in March of that year.

The Crisis Unfolds

School closures in England were announced by the then Prime Minister of the United Kingdom, Boris Johnson, in the House of Commons on 16th March 2020. Following on rapidly from school closures, the announcement of the cancellation of examinations in England was made, again by the Prime Minister, in the House of Commons on 18th March, five days prior to the announcement of a national lockdown, on 23 March 2020. This was a decision taken centrally: the Department for Education (the UK government department responsible for education in England) was not consulted. The regulatory body responsible for examinations in England, Ofqual, had indicated a preference for holding at least some socially distanced examinations, or for delaying examinations. By the 31st of March the then Secretary of State for Education in England, Gavin Williamson, had instructed Ofqual that a statistical model (algorithm) needed to be developed to substitute for examination grades. This algorithm was to be designed to moderate teacher assessments, which are known to err on the side of generosity (Wikstrom and Wikstrom, 2006). The directive from Williamson to Ofqual also set a clear priority that reflected the Conservative government's pre-occupation with standards: grade inflation was to be avoided: 'Ofqual should ensure, as far as is possible, that qualification standards are maintained and the distribution of grades follows a similar profile to that in previous years' (Williamson, 31st March 2020). We do not imply that attention to the currency of examinations is unwarranted - one of the authors has spent her career researching methods of standard-setting. However, we draw attention to the value-driven choice to continue to prioritise control of outcomes during the pandemic; the policy choices were not found beyond the British Isles during the pandemic (UNESCO, 2020).

In early April Ofqual issued guidance on moderating procedures for Centre Assessed grades (CAGs) with a deadline for the completion of teacher assessments by May 29th 2020. On 22nd May Ofqual issued a statement on the use of the algorithm, and then in July further material on how it was to be applied. Ofqual tested 11 models using three years' data to predict grading outcomes in 2019 and compared with the actual 2019 results, since they were from a normal, pre-Covid year's examination (Ofqual, 2020). From this modelling, the 'direct centre-level performance' approach was selected because it produced the most accurate predictions. The relationship between pupils' prior attainment at age 16 and outcomes in the A-level subject in question was calculated and it was anticipated that this progress would remain roughly equivalent in 2020. These steps were typical in a normal year under the comparable outcomes method, but in 2020 the statistical moderation was also applied at the level of schools and colleges. Using the prior attainment of pupils entering each A-level in a particular subject in a specific school, Ofqual could predict the outcomes for that group using the national patterns of progress from prior attainment. This was then adjusted, on a school by school (or college) basis for changes in ability between 2019 and 2020; where ability is measured by prior attainment. Depending upon the proportion of students with prior attainment data, the influence of prior attainment upon the algorithm predictions could vary. At an extreme, if there was no prior attainment data, the entire prediction was based upon the outcomes in the school or college in previous years. The higher the proportion of students with prior attainment data, the more that data influenced the prediction. The algorithm predicted marks, not just grades. This allowed for more fine-tuning of the outcomes in each school or college, and therefore at a national level. Teachers' rank

orders for pupils were used in this part of the process. Finally, cut-scores (mark boundaries) at which grades would be awarded were set for each grade and these were used to set standards across all of the exam boards who offered the A-level in that subject.

On the 13th of August the results based on the algorithm were issued to A level and GCSE pupils, and a public outcry ensued. Overall, nearly 36% of grades were lower than the CAG and 3% were down two grades (among the three or four A-Level subjects usually taken) (Kelly 2021:728). Media coverage largely focused on human interest stories of young people's hopes dashed by what they perceived as an unfair reduction in their grades (see, for example, the Guardian 21st February 2021). Parents whose children missed their preferred university place put pressure on their constituency MPs. Media reporting and social media highlighted the number of so called 'outliers', that is, pupils whose algorithmically-derived grades were significantly different from their predicted grades, along with the negative impact of taking account of previous school performance in developing the model, which seemed to penalise high achieving pupils in low performing schools, often situated in areas of social disadvantage. Attempts were made to manage public dissatisfaction through an appeals procedure and on 12th August the Secretary of State announced a 'triple lock' system whereby students would be able to select from a choice of the algorithmically-derived grade, a 'valid' mock examination grade, (that is, a grade obtained in practice examinations where schools had held them under sufficiently rigorous conditions) or an examination to be held in autumn. These plans collapsed in the face of difficulties over, among other things, comparability of mock examinations, and uncertainty about holding examinations in autumn. On 17th August the decision was taken to use Centre Assessed Grades and to abandon the algorithm. As a result, there was a substantial level of grade inflation, with an increase of more than 10 percentage points in the number of top grades awarded (from 25.2% to an estimated 37.7%), the biggest increase for at least 20 years.

This policy reversal in an area that is critical to the futures of young people in terms of employment and apprenticeship, as well as selection for college and university entrance, was very damaging to public confidence (YouGov, 2021). Public trust in examination standards is essential, as they offer politicians a selection mechanism that has a high degree of public acceptance, provided standards are seen to be maintained. However, tensions between policies aimed at improving performance in education had co-existed uneasily and for some time with anxieties about grade inflation, while debates around forms of assessment and differences in performance challenge the prioritisation of conventional examinations and highlight the effects of difference in resources, teacher capacity and home-based support for learning.

Some of those issues became more apparent in England during 2020, as policy makers struggled with the complexity of maintaining examination and assessment systems and standards perceived as trustworthy by the public while also acknowledging the differential impact on pupil learning of prolonged school closures, dependence on home schooling, unequal access to learning support, and isolation during lockdown (Newton, 2021). These factors together challenged assumptions about the fairness of conventional examination forms and threw already existing differences in opportunity between schools and learners into sharp relief, revealing them more clearly to the public. Moreover, the policy turn to an algorithm as the best technical 'fix', contrasts sharply with pupil and public perceptions that young people were being graded by a technology without direct reference to their work, and raises questions about the kinds of knowledge and expertise in play in the policy response to the examinations crisis. We focus here on the findings from our research¹ into the nature of the evidence and knowledge that informed the actors engaged in policy-making for examinations in

¹ **Knowledge use in Assessment policy during Covid-19:** funded by the John Fell Fund University of Oxford, PI Jo-Anne Baird, with Margaret Arnott and Nic Hell, UWS, and Jenny Ozga and Luke Saville, University of Oxford.

England during 2020, exploring the different forms of knowledge drawn upon, the influence of different agencies and actors, their interconnections, and their relationship with political priorities.

Methodology

The research reported here was carried out between May 2021 and March 2022. The overarching project involved research across the United Kingdom, however this paper deals only with England and is confined to reporting on the data directly relevant to the issues of expertise, and the policy and knowledge actors involved in handling the examinations crisis there, with some discussion of the role of politics. Other publications in preparation deal with inter-UK comparison and political leadership. Two main sources of material are drawn on in this discussion: policy texts and interviews with key actors. Data collection was organised in three distinct but overlapping stages: (i) the identification of potentially relevant source material on the broader management of the Covid-19 pandemic by the UK government. This phase of the research enabled us to identify the extent to which education did or did not appear as a significant policy area within the wider context of the developing crisis, and to identify references to educational expertise and knowledge in that broader context. (ii) content analysis of policy texts identified through Stage 1 as relevant to the examinations crisis (iii) analysis of interviews with key actors (n=16) involved in that crisis.

Stage 1 produced a corpus of policy texts relating to overall Covid management. The texts included the published minutes of the Scientific Group for Emergencies (SAGE) and background papers from its subgroups. Content analysis revealed that educational expertise and knowledge were almost totally absent from debates dominated by reliance on health and epidemiological expertise.

Stage 2 focused on the identification and analysis of policy texts where there was a primary concern with education, especially those relating to examination policy in the period from March-August 2020. These texts included Prime Ministerial and Ministerial statements and communications, Ofqual Board Minutes and public documents communicating decisions or consultations referring to examinations. We also drew on retrospective evidence (for example witness statements and oral evidence) contained in various Select Committee reports to the House of Commons. This selection of materials was representative of the main actors-politicians, scientific experts and the regulatory agency (Ofqual)-most directly involved in the decision-making process between March and August 2020. The texts are predominantly recognisable as formal policy texts, though they are expressed in different formats-as public communications, as correspondence and as formal, and sometimes partially-redacted, minutes of proceedings. The texts analysed in Stages 1 and 2 are listed in the Appendix.

Stages 1 and 2 enabled identification of key issues for investigation through the stage 3 interviews with a representative sample of the key agencies and actors involved in the examinations process. These interviews were carried out between November 2021 and March 2022 with DfE officials, members of the Ofqual Board and Standards Advisory Group, Examination Boards, Teacher Unions, and assessment experts. The interviews were recorded. Informants were anonymised, identified only by their roles (for example- Regulatory Actor indicates a member of Ofqual, GA a central government actor, PA an actor from a professional organisation, AE an assessment expert). The data collected through the interviews in stage 3 was also used to check and elaborate the analysis of stage 2 material.

A political discourse analysis approach to the policy texts and interview data was adopted (Wodak 2014) in order to identify the interaction of politics and processes of knowledge production and use, with a focus on the power relations among actors, and their capacity to influence one another and the context in which they do policy work.

Findings

School closure and examination cancellation

Our analysis of Stage 1 data makes it very clear that the decision to close schools was a political decision. Public health concerns were paramount at this time, and it was health expertise that shaped policy for closure and subsequent examination cancellation. Discussion of school closures focused on evidence about their potential contribution to the reduction of the peak of Covid infections. The discussions throughout March are almost entirely about schools as sites of transmission, though there is a note on 17th March that SAGE discussed behavioural science considerations on school closures.

In late April there is a wider discussion of possible ways of managing school-re-opening through better ventilation, distancing and cutting down on pupil movement between classes. On 29th April, in a discussion of ethical issues raised by the pandemic, and in a wider discussion of protecting public institutions, behavioural science expertise is drawn on in a discussion of children's well-being. In summary, the SAGE minutes and those of associated sub-groups show that the main focus was on learning from epidemiological and modelling expertise about the effects of closures on transmission. There is also a reference (SO61:57) to 'the shock to education' and its serious implications for lives and outcomes in a discussion of the risks in keeping schools closed, which seems to draw on psychological and sociology of public health expertise. Retrospective investigation of the DfE's response to the crisis suggests a lack of planning and a failure to broaden the expertise available to it: for example, the Public Accounts Committee of the House of Commons in a report on the DfE's support for children's education during the pandemic (PAC240) is highly critical of the Department and castigates it for its absence of both forward planning and learning from experience.

Our interviews with senior officials in the Department emphasise that decisions were taken at speed and under pressure-once the Prime Minister and Cabinet had decided on school closure:

'It all happened very quickly-there was no pandemic plan, we were doing a lot from scratch-everything was decided within 48 hours. There was no outside input, no discussion outside the Department at all. After the school closure decision then cancelling examinations was seen as inevitable' (GA1)

There was a debate about the possibility of socially distanced examinations, using large buildings, to keep some examinations going, but:

'....we decided we were not doing that...it was not a sophisticated, consultative or engaged [discussion]. No planning had been done on how to make it work, there was no time. Nobody knew how long schools would be closed for-we argued about it but were not confident that we could do it. It was not really credible that we could-so we cut our losses, this was the least worst option'. (GA1)

At this early stage in the crisis, then, the decision to close schools was taken on the basis of public health concerns, drawing on epidemiological and modelling expertise from government chief scientific officers and the wider scientific community. Educational expertise is not visible in our data.

The algorithm

Moving on to the development of the algorithm, here the focus is on the expertise drawn upon in designing a substitute for examinations by the main actors-politicians, scientific experts and the regulatory agency (Ofqual)-most directly involved in the decision-making process between March and August 2020. The Government had announced the decision to cancel examinations without further consultation with Ofqual. Under normal circumstances, Ofqual would have been consulted to a greater extent. Once examinations had been cancelled, the Secretary of State for Education wrote to Ofqual, indicating that students should be issued with 'calculated results' based upon teachers'

judgments, supplemented by a range of other evidence, and, as indicated above, that grade inflation should be minimised.

Ofqual-and the wider assessment community-understood that teacher assessments produce higher grades than examinations, so it was also understood that the moderation required a statistical intervention. Equally, production of what was termed 'Centre Assessed Grades' required teacher judgments to rank-order learners within schools and colleges. Ofqual consulted widely on the details of the Centre Assessed Grades procedures within the education sector, receiving over 12,000 responses from teachers, students, parents, schools and local authorities, teacher unions and exam boards, with a smaller number of responses from higher education institutions (Ofqual 2020) and also created an External Advisory Group on Examination Grading. The composition of the Advisory Group represented the 'statistical and assessment communities', who brought 'particular expertise and/or sector representation' and representatives of the Department for Education. This Group advised on grading of the summer qualifications generally, including the Centre Assessed Grades procedures and the selection of the algorithm. Exam board statisticians were heavily involved in the modelling which was conducted to select the algorithm. The Advisory Group members included Ofsted, UCAS, academics (including Baird), a representative of the Treasury, a Headteacher, an exam board researcher and independent consultants. In developing the Centre Assessed Grades, education practitioners' expertise was especially sought, and it formed part of the evidence drawn upon in selection of the algorithm. The Royal Statistical Society offered assistance to Ofqual, but declined to participate when Ofqual asked them to sign a confidentiality agreement -now standard practice. Although a later exchange implied that problems with the algorithm were due to a lack of statistical expertise in the expert group, most members of that group were appointed precisely because they had the expertise necessary to understand the complexities of the assessment data and modelling in detail.

The use of educational expertise in developing a substitute mechanism for examinations is evident from the consultation documents produced by Ofqual, (Ofqual 2020 a,b,c,d) including the responses to that consultation and the subsequent document on exceptional arrangements for GCSE and A level examination. Some of that evidence illustrated concerns about Ofqual's standardisation model, including issues that did, indeed, become highly problematic when the results were published, for example the use of historic data, or where there were 'outliers' - high achievers in historically low performing schools, or small specialist providers with highly variable year on year cohorts.

We draw here on analysis of these documents, including analysis of oral and written submissions to the House of Commons Select Committee enquiry into the impact of Covid on Education and Children's services (HoC 617, July 2020). In addition to the oral evidence from the Minister and representatives of Ofqual, a wide range of witnesses provided written evidence. These submissions were mostly from organisations representing young people (for example YoungMinds), from institutions (for example the Russell Group, Education Unions, the Education Endowment Foundation) but there was also a smaller number of research-based submissions from individual academics, drawing attention to issues raised by the handling of the crisis, for example in transition to work and higher education, or to the impact on social mobility. We also draw here on the interview data from stage 3 of the research.

This analysis highlights themes relating to knowledge, expertise and policy, which may be summarised as (i) the selective expertise drawn upon (ii) increased politicisation of policy (iii) increased concern with and reliance on technocratic solutions, (iv) the degree of independence of the regulatory agency (Ofqual). We consider each of these in turn.

Selective Expertise The analysis highlights a continuing sense of crisis and uncertainty, which restricted the knowledge and expertise available to policy makers. Our informants talk about how the policy process *'kind of closed-down, not panic but, we've got to get on, we've got to come up with something'* (RA 3).

'It was a quick decision-we couldn't do exams so what were we going to do? -This was also a quick decision without proper engagement outside government..... the fact that the policy debate was going on behind closed doors was the problem, rather than that there wasn't any expertise in the department.'[GA1]

A related issue was that of infrastructure, and the Department's capacity to communicate and potentially learn from developments and practices 'on the ground' in schools and localities. The centralisation of oversight of schools following academisation and the reduction of local authority controls and responsibilities that characterise system development in England in recent years (Lawn 2014) had significantly increased the role of national politicians:

'there was kind of a closing down of the policy thing, rather than an opening out at the crucial stage, so local authorities and schools all felt under-represented in the discussions' (PA1)

In addition to absence of communication infrastructure that could support knowledge flows within the system, the 'closing' of the policy process extended to the design and development of the algorithm. Although, as indicated above, Ofqual consulted widely on the design of the algorithm and CAGs, the Standards Advisory Group (SAG), which included academic experts on assessment, did not meet during the period of developing the algorithm:

'I think the Standards Advisory Group was quite marginal in the whole process' (RA 2)

'I had kept all my copies of my papers and my contemporaneous notes of those meetings, and I looked back and I couldn't see anything about any of those, certainly cancellation of exams, centre assessed grades, nothing, nothing on the algorithm. So as far as I am concerned, SAG had absolutely no involvement in any of that.' (AE1)

Instead, the External Advisory Group on Examination Grading had representation from members of the SAG, as well as a wider membership including statisticians, analysts from Teacher Tapp, the University Admission Service (UCAS), the Inspectorate (Ofsted), a headteacher and the economic adviser to the Chancellor of the Exchequer.

Increased politicisation of policy

Political priorities narrowed the parameters within which policy for replacing examinations was developed. The key political priority was the development of a substitute for examinations that avoided grade inflation. Preventing grade inflation-and avoiding any reduction in standards-was a central preoccupation of the UK government:

'The biggest problem of whole thing-which should have been open for choice.....was that the no 1 priority for ministers was -ironically as it turned out-to avoid grade inflation.if avoiding grade inflation is your principal aim, then you end up with an algorithm-you've made the choice that avoiding grade inflation is your most important priority.'(GA1)

'I think right from the start- I look back even to notes from meetings in March, we were very much on tracks, I felt, you know the direction had been set from the Secretary of State, we've got to give some grades with no grade inflation, or minimal grade inflation.....there perhaps should have been more

scrutiny of what it was we are trying to achieveI think that perhaps it all got onto... a narrow pathway.... at that stage it was just about can we make this work, rather than is this the right thing to be doing? (RA3).

Technocratic Solutionism

The development of the algorithm became the pre-occupation of the Ofqual board members and Ofqual's technical experts in the period of intense work between April and July 2020. Considerable effort went into estimating what grades people would have achieved if they had sat examinations in as technically defensible and precise a way as possible. Our analysis highlights the key role played by Ofqual's technical working groups while also underlining their experience and expertise:

'...these are people who are very used to dealing with the mechanisms for the comparable outcomes process, by which prediction matrices, expected distributions are created. So we've got all the most technical people from the boards..... and so as well as having people who know a lot about statistics and psychometrics they know about the detail of the exams, which, as it became increasingly clear, is incredibly big, when you're trying to do this, this thing by algorithm. I think there was a very high level of technical expertise both at Ofqual and in the [examination] boards involved in this..... so perhaps really the issue is whether the criteria for success for the algorithm.... were with hindsight the right ones'. (RA 3)

That final sentence highlights the extent to which finalising the algorithm became the key focus for those engaged in finding a policy solution to the cancellation of examinations. Our informants stress the high levels of commitment to testing the model, and also indicate, as the quotation above hints, that they may have been captured by it-impressed by its elegance and sophistication, by its capacity to offer an objective and valid solution to the policy problem. This is a specific example of what Selwyn (2018) calls statistical 'solutionism' in education, where the use of algorithms to predict or manage issues of access and distribution inhibits or obscures some fundamental questions about purpose.

The final theme that emerges across our data concerns the **Nature of regulation** and the independence of Ofqual. The official literature stresses the independence from government of the regulatory agency. Ofqual describes the relationship on its website as follows: 'We're independent of government and report directly to Parliament.' That characterisation of parliament as protecting Ofqual's independence was an issue raised in our interviews, where the need for an 'arm's length' relationship was acknowledged, but concerns were expressed about the communication problems this presented in a crisis:

'...whether and how children should be doing exams-that is a decision that has to be made by ministers in a much more direct way -government needs to decide what to do. It's looking to Ofqual for expertise but they were supposedly outside...in this situation there was no separation-separation was not a plausible situation-not real-another example of how a set of arrangements and policies designed for a steady state world don't work in crisis.' (GA 1)

Moreover the Secretary of State's letter to Ofqual (Williamson 31st March 2022) is, in formal terms, a 'direction'-it makes clear that the approach to be developed by the regulator was to be standardised across centres and Ofqual was further directed to ensure that grade inflation was minimised. As indicated above, political preoccupations with grade inflation shaped the response, while responsibility for designing the procedures for adjusting grade distribution was placed with Ofqual. Some informants described Ofqual's independence as 'technical', pointing to its funding through government, that its membership is appointed by government, that the chief regulator is appointed by the minister, and can be removed or 'life made very difficult for them'.

Discussion

We have highlighted four aspects of the examinations crisis in England in 2020, which we summarise as the selective expertise, the increased politicisation of policy, the increased concern with and reliance on technocratic solutions, and the nature of regulatory activity. These are issues that appear in the wider research and literature on contemporary education, in England and elsewhere. For example, the increased reliance on information, especially data-based information, in the governing of education (Addey et al 2017) is well documented, while, at the same time, independent research in education, conventionally based in universities, is subject to increased policy steering through targeted funding and the setting of policy priorities for research (Marques et al 2017). Datafication reduces complex problems to statistical calculations. Model building may be seductive and persuasive but disconnected from real world effects. Algorithms are ubiquitous; indeed some current scholarship call ours 'the algorithmic society' (Schiulenburg and Peeters 2021) and highlights their use in policy areas such as health care, criminal justice, tax administration and education. Algorithms present a new paradigm of governance as increasingly rational. As Schiulenburg and Peeters suggest, algorithmic governance designs out human agency in decision-making, and changes decision making from a matter of individual or collective judgement to one of classification (op cit: 2) The expertise in use changes: the 'discretionary space' of expertise shifts to IT professionals and data analysts. In the specific case of assessment policy in England in 2020, the consequences of the results for individuals were obscured by system-level, data-based thinking. Accustomed to thinking of assessment at an aggregate level, application of an algorithm was seen by some interviewees as part of the 'natural order' and the values underlying the algorithm were under-examined until there was a public backlash.

The politicisation of policy is also recognised in the wider literature, the relationship between knowledge, expertise and policy making is understood to be tense, across a range of policy fields (see, for example Stone 2013). Pressures are heightened in crisis, where the demand for scientific and technical expertise and for experts to inform, justify or legitimise policy responses is increased (Stone 2019). Problems arise when specialists are asked to deal with intractable and complex 'wicked' problems, as they tend to fall back on established procedures, and interpret new, uncertain situations as extensions of earlier 'tame' problems (Grundmann 2018:380). Political pressures increase as governments seek to reassure publics that the crisis is manageable and that solutions can be found. The result, as Weible et al (2020:7) argued, is 'a simultaneous increased reliance on scientific and technical experts and politicization of scientific and technical information.' Furthermore, as scientific and technical experts help to inform and legitimise decisions, they may also obscure political responsibility for policy responses and outcomes. The algorithm was developed by the regulator to address the remit set by the Secretary of State to control grade inflation (Williamson's letter of 31st March 2020). Thus, there was a political envelope within which the regulators operated and the expertise that was selected was within this scope. The decision to control grade inflation was influenced by political and ideological considerations; other policy options were available. When the algorithm was questioned, the debates were political in nature; not about assessment, education or science-based forms of knowledge. Had a wider range of expertise been drawn upon in assessment policy-making, it is possible that the grade inflation which resulted in 44% of the candidates being awarded a grade A could have been avoided. Positioning teachers as the upholders of educational standards and involving them through social moderation (Klenowski and Wyatt-Smith, 2014) might have resulted in lower rates of inflation.

The ambiguous relations between government and regulatory agencies are also well-documented. Although regulation means different things in different contexts, and varies from hands on, to hands off activities by political leaders, these agencies possess considerable expertise: as Majone suggests, 'regulation is not achieved simply by passing a law, but requires detailed knowledge of, and intimate

involvement with, the regulated activity' (Majone, 1994, 81). It is on the basis of this 'intimate involvement' that regulatory agencies are given considerable discretion and operate at 'arms-length' from political control. It is also suggested that, given the enormous expansion in information available, and its complexity, policy-makers delegate responsibility to regulatory agencies to solve the problem of their lack of expertise in increasingly complex areas, and promote the operation of the agencies as a technocratic solution to intractable social and political problems. Thus regulatory agencies are expanding the scope of their operations, and have considerable discretion in reconciling often conflicting mandates (Sparrow, 2000:5). Moreover, because regulatory objectives are typically defined rather narrowly, discussion is limited to quite technical issues (Koop and Lodge 2020). Senior appointments to regulatory agencies are made by the Government and in the case of Ofqual and Ofsted must be ratified by the Education Select Committee. Alignment of appointees with the Government's values is advantageous, with many regulatory board level positions and significant policy reviews, for example of the Academies policy (Larsen, Bunting, and Armstrong. 2011) being carried out by individuals whose expertise derives from organisations affiliated with NPM marketisation policies such as Multi-Academy Trusts, Teach First or consultancies such as Capita.

The political nature of the final outcome suggests that the government was swayed primarily by public discontent with the consequences of the direction to prevent grade inflation, in a context where the unfairness of technical adjustment to grade distribution was much more apparent than in 'normal' times. It is also clear that the regulatory agency, Ofqual, was initially positioned by the Secretary of State to take responsibility for the perceived unfairness of the statistical adjustment. The analysis of documents reveals very considerable tension in the 'zone of discretion' of expertise (Majone 1994) between the agency and the Department. At the same time, Ofqual's expertise perhaps encouraged a focus on the statistically best adjustment to the distribution, somewhat removed from understanding of how this adjustment would be perceived by the public, and treated by the media, including social media. More generally, this initial exploration of the examinations crisis of 2020 suggests that there is a gap in public understanding of assessment practices, including the routine activities of standardisation of examinations performance, and this connects to the absence of informed, expert voices and debate about assessment more widely. Examinations, it might be argued, have become the absolute priority area of education policy and practice, yet their assumptions and technologies are rarely discussed or questioned. They dominate English education policy, they fuel political scepticism about alternative approaches to assessment, and they are constructed, in political discourse, as objective, reliable and 'real' indicators of educational worth. In the emergency caused by the pandemic, the political impetus from the Government was to control grade inflation, rather than, for example, to focus on ensuring that the progression from school to employment or university went as smoothly as possible. An ideological focus upon standards and the rationing of resources through selection closed down the agenda.

Towards a Conclusion

Policy responses to the pandemic offer insights into the relationship between knowledge and policy, a relationship under particular stress in a period of crisis, when knowledge is uncertain and political pressures are intense (Ansell et al 2020). In the case of England, strong interventions in the policy field of education during the pandemic were enacted against the backdrop of a history of continuous 'reform' of education in line with new public management (NPM) practices and neo-liberal principles (Broadbent 2020, Greany, 2020). These are reflected in the design of competition between schools and the encouragement of private actors in new governing networks. Simultaneous decentralisation and recentralisation of education is visible from the 1980s onwards, (Lawn 2013) for example in the Academies programme and in the erosion of the powers of local government, at the same time as a complex data-based national regime of performance assessment, curriculum control and inspection, was created (Ball 2018). Critics have identified a dangerous diffusion of responsibility, the hollowing

out of state capacities and reliance on corporate outsourcing or regulatory agencies as problematic consequences of this shift (Ball 2021; Schillemans et al, 2020). These reforms, and especially the privatisation and decentralisation that accompanied the implementation of New Public Management, blurred lines of accountability, making it hard to see clearly who is accountable for what. At the same time, political accountability from governments to the electorate and through elected local representatives with responsibility for provision is reduced or obscured (Ozga 2020). That fractured and fragmented system frames our discussion of examinations policy and connects to the questions about the interrelationship of knowledge and governing that underpin our interests here. In particular, we note the growth of reliance on knowledge expressed as data, its increased availability and complexity, indeed a trend towards reliance on data collection and analysis as *the* lens through which the system is understood and managed, along with a growth of 'in house' expertise. Both trends support the increased steering of knowledge production to meet policy requirements, and its translation into 'actionable' knowledge (Grundmann and Stehr 2012, Grundmann 2017, 2018).

Other scholars have analysed the crisis drawing on resources from political science. For example Cairney (2020) and Kippin and Cairney (2021) highlight the lack of capacity of British policy makers, their inability to control their policy environments and their limited knowledge. We suggest that sources of advice and actionable knowledge in education policy in England were restricted during the exams crisis because of the preoccupation with avoiding grade inflation, the dominance of technical solutionism, and the consequent privileging of statistical expertise, and of those analysts fluent in the technology that supports such activity. Other forms of educational expertise, including professional and research-based academic knowledge, have been displaced, so that the sources of advice and information on which government could draw was limited. Centralisation of policy making, combined with crisis, accentuated the tendency to make policy behind closed doors. The material presented in our analysis of the examinations crisis of summer 2020 in England reveals the absence of input from a wide range of educational expertise, including on the societal consequences of school closure, on transitions, on pedagogy, on alternative forms of assessment. These exclusions were, we suggest, unfortunate, as they are forms of knowledge that make evident the centrality of the school as an institution to social, as well as economic life, and thus illuminate its social significance, which extends beyond certification and selection.

Appendix: Policy Texts

Evidence considered by the Scientific Advisory Group for Emergencies has been published within the gov.uk collection 'Scientific evidence supporting the government response to coronavirus (COVID-19)'. (URL: Scientific evidence supporting the government response to coronavirus (COVID-19) - GOV.UK (www.gov.uk) date accessed: 26/06/2021).

Gavin Williamson's Directive to Ofqual of 31st March 2020

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/877611/Letter_from_Secretary_of_State_for_Education_to_Sally_Collier.pdf

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/879627/Exceptional_arrangements_for_exam_grading_and_assessment_in_2020.pdf

Education Select Committee (2020) The impact of Covid-19 on education and children's services. House of Commons, HC 254. Wednesday 2 September.
<https://committees.parliament.uk/oralevidence/790/pdf/>

Prime Minister's statement on coronavirus (COVID-19): 19 March 2020,
www.gov.uk/government/speeches/pmstatement-on-coronavirus-19-march-2020

Prime Minister's statement on coronavirus (COVID-19): 30 April 2020,
www.gov.uk/government/news/primeministers-statement-on-coronavirus-covid-19-30-april-2020

Science and Technology Select Committee, Oral evidence, HC 136, 13 May 2020, Q 514,
<https://committees.parliament.uk/oralevidence/380/pdf/>

Science and Technology Select Committee (2021) The UK response to covid-19: use of scientific advice HoC 136

Education Select Committee, Oral evidence: The impact of Covid-19 on education and children's services, HC 254, 2 September 2020, Q 948, <https://committees.parliament.uk/oralevidence/790/pdf>

Education Select Committee, Oral evidence: The Impact of Covid-19 on education and children's services, HC 254, 2 September 2020, Q 948, <https://committees.parliament.uk/oralevidence/790/pdf>

Education Select Committee, Oral evidence: The impact of Covid-19 on education and children's services, HC 254, 9 March 2021, Q 1257, <https://committees.parliament.uk/oralevidence/1837/pdf>

Education Select Committee (2020) The impact of Covid-19 on education and children's services. House of Commons, HC 254. Wednesday 2 September.
<https://committees.parliament.uk/oralevidence/790/pdf/>

Public Accounts Committee support for children's education during the pandemic House of Commons (PAC240)

Department for Education, 'Triple lock for students ahead of A level and GCSE results', GOV.UK, 12 August 2020,

www.gov.uk/government/news/triple-lock-for-students-ahead-of-a-level-and-gcse-results

Ofqual Board Minutes 30th January, Ofqual Emergency Board Minutes 20th March 2020, 22nd April, 13th May, 20th May, 3rd June, 17th June, 24th June, 30th June, 09 July, 23rd July, 29th July, 04 August, 12th August, 13, 14, 15, 16 August, 18th August, 23rd August available at <https://www.gov.uk/government/publications/ofqual-board-minutes-for-2020>

Ofqual (2020a) 'How GCSEs, AS & A levels will be awarded in summer 2020', GOV.UK, 3 April 2020, www.gov.uk/government/news/how-gcses-as-a-levels-will-be-awarded-in-summer-2020

Ofqual (2020b) Consultation decisions: Exceptional arrangements for exam grading and assessment in 2020, Ofqual, May 2020
pp. 8-10
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/887048/Summer_2020_Awarding_GCSEs_A_levels_-_Consultation_decisions_22MAY2020.pdf

(Ofqual 2020c) Ofqual Board Briefing, 11 August 2020,
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/928302/20200811_1245__Ofqual_board_briefing.pdf

(Ofqual 2020d)
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/886555/Analysis_of_consultation_responses_21MAY2020.pdf

References

- Addey, C. Sellar, S. Steiner-Khamsi, G. Lingard, B. and Verger, A. (2017) The rise of international large-scale assessments and rationales for participation, *Compare: A Journal of Comparative and International Education*, 47:3, 434-452
- Ansell, C. Eva Sørensen & Jacob Torfing (2020): The COVID-19 pandemic as a game changer for public administration and leadership? The need for robust governance responses to turbulent problems, *Public Management Review*, DOI: 10.1080/14719037.2020.1820272
- Ball, S.J. (2018) Commercialising education: profiting from reform, *Journal of Education Policy*, 33:5, 587-589, DOI: [10.1080/02680939.2018.1467599](https://doi.org/10.1080/02680939.2018.1467599)
- Ball, S. J. (2021). Response: policy? Policy research? How absurd?. *Critical Studies in Education*, 62 (3). doi:10.1080/17508487.2021.1924214
- Broadbent, J. (2020), "The response to Covid-19 in England: political accountability and loss of trust", *Journal of Accounting & Organizational Change*, Vol. 16 No. 4, pp. 527-532. <https://doi.org/10.1108/JAOC-07-2020-0093>
- Cairney, P. (2021) The UK government's COVID-19 policy: assessing evidence-informed policy analysis in real time. *British Politics* **16**, 90–116. <https://doi.org/10.1057/s41293-020-00150-8>
- Dore, R. (1976). *The Diploma Disease: Education, Qualification and Development*. London: George Allen & Unwin.
- Education Select Committee (2020) The impact of Covid-19 on education and children's services. House of Commons, HC 254. Wednesday 2 September. <https://committees.parliament.uk/oralevidence/790/pdf/>
- Greany, T. (2020): Place-based governance and leadership in decentralised school systems: evidence from England, *Journal of Education Policy*, DOI: 10.1080/02680939.2020.1792554
- Grundmann, R. (2017) The Problem of Expertise in Knowledge Societies. *Minerva* **55**, 25–48 (2017). <https://doi.org/10.1007/s11024-016-9308-7>
- Grundmann, R. (2018) The Rightful Place of Expertise, *Social Epistemology*, 32:6, 372-386, DOI: [10.1080/02691728.2018.1546347](https://doi.org/10.1080/02691728.2018.1546347)
- Grundmann R and Stehr, N (2012) The Power of Scientific Knowledge Cambridge, Cambridge University Press
- Guardian (2020) The government's secret science group has a shocking lack of expertise. 27 April. <https://www.theguardian.com/commentisfree/2020/apr/27/gaps-sage-scientific-body-scientists-medical>
- Guardian (2021) <https://www.theguardian.com/education/2021/feb/18/the-student-and-the-algorithm-how-the-exam-results-fiasco-threatened-one-pupils-future>

House of Commons Education Select Committee (2012) *The administration of examinations for 15 – 19 year olds in England*. HC 141-I.

<https://publications.parliament.uk/pa/cm201213/cmselect/cmeduc/679/67904.htm>

Koop, C. & Lodge, M. (2020) British Economic Regulators in an age of Politicisation: from the responsible to the responsive regulatory state?, *Journal of European Public Policy*, 27:11, 1612-1635, DOI: [10.1080/13501763.2020.1817127](https://doi.org/10.1080/13501763.2020.1817127)

Kelly, A (2021) A tale of two algorithms: The appeal and repeal of calculated grades systems in England and Ireland in 2020 *British Educational Research Journal* Vol. 47, No. 3, June 2021, pp. 725–741 DOI: [10.1002/berj.3705](https://doi.org/10.1002/berj.3705)

Kippin S and Cairney, P (2021) The COVID-19 exams fiasco across the UK: four nations and two windows of opportunity *British Politics* <https://doi.org/10.1057/s41293-021-00162-y>

Klenowski, V. and Wyatt-Smith, C. (2014) *Assessment for Education. Standards, Judgement and Moderation*. Sage: London.

Larsen, J., V. Bunting, and D. Armstrong. 2011. "What Works? An Empirical Assessment of Strengths, Weaknesses and Next Steps for the Academies Initiative." In *The State and Education Policy*, edited by H. M. Gunter, 105–119. London: Continuum.

Lawn, M (2013) A Systemless System: Designing the Disarticulation of English State Education. *European Educational Research Journal*. 2013;12(2):231-241. doi:[10.2304/eej.2013.12.2.231](https://doi.org/10.2304/eej.2013.12.2.231)

Marques, M. Powell, J. Zapp, M. Biesta, G.(2017) How does research evaluation impact educational research? Exploring intended and unintended consequences of research assessment in the United Kingdom, 1986–2014 *European Educational Research Journal* Vol. 16(6) 820–842

Majone, G. (1994) The rise of the regulatory state in Europe, *West European Politics*, 17:3, 77-101, DOI: [10.1080/01402389408425031](https://doi.org/10.1080/01402389408425031)

Meadows, M. and Black, B. (2018) Teachers' experience of and attitudes toward activities to maximise qualification results in England, *Oxford Review of Education*, 44:5, 563-580, DOI: [10.1080/03054985.2018.1500355](https://doi.org/10.1080/03054985.2018.1500355)

Newton, P.E. (2021). Learning during the pandemic: the context for assessments in summer 2021. Report 1 of 5 on learning during the 2020 coronavirus (COVID-19) pandemic. Ofqual/21/6803/1. July. <https://www.gov.uk/government/publications/learning-during-the-pandemic>

Ofqual. (2015). Setting GCSE , AS and A level grade standards in summer 2014 and 2015. Coventry, UK: Office of Qualifications and Examinations Regulation. Retrieved from <https://www.gov.uk/government/publications/setting-gcse-and-a-level-grade-standards-in-summer-2014-and-2015>

Ofqual (2020). Awarding GCSE, AS, A level, advanced extension awards and extended project qualifications in summer 2020: interim report. 13 August. <https://www.gov.uk/government/publications/awarding-gcse-as-a-levels-in-summer-2020-interim-report>

Ofqual (2020a) 'How GCSEs, AS & A levels will be awarded in summer 2020', GOV.UK, 3 April 2020, www.gov.uk/government/news/how-gcses-as-a-levels-will-be-awarded-in-summer-2020

Ofqual (2020b) Consultation decisions: Exceptional arrangements for exam grading and assessment in 2020, Ofqual, May 2020

pp. 8-10

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/887048/Summer_2020_Awarding_GCSEs_A_levels_-_Consultation_decisions_22MAY2020.pdf

(Ofqual 2020c) Ofqual Board Briefing, 11 August 2020,

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/928302/20200811_1245__Ofqual_board_briefing.pdf

(Ofqual 2020d)

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/886555/Analysis_of_consultation_responses_21MAY2020.pdf

Opposs, D. (2020) The impact of the coronavirus outbreak on exams around the world. Ofqual blog. 22 May. <https://ofqual.blog.gov.uk/author/dennis-opposs/>

Opposs, D. Baird, J., Chankseliani, M., Stobart, G., Kaushik, A. McManus, H. and Johnson, D. (2020) Governance structure and standard setting in educational assessment, *Assessment in Education: Principles, Policy & Practice*, 27:2, 192-214, DOI: 10.1080/0969594X.2020.1730766

Ozga, Jenny (2020) The politics of accountability. *Journal of Educational Change* 21, 19–35 <https://doi.org/10.1007/s10833-019-09354-2>

Science and Technology Committee (2021) The UK response to covid-19: use of scientific advice HoC 136

Royal Statistical Society (2020) RSS statement of grade adjustment in UK examinations in 2020 <https://rss.org.uk/news-publications/2020general-news/rss>

Schillemans, T. Overman, S. Fawcett, P. Flinders, M. Fredriksson, M. Laegrid, P. Maggeti, M. Papadopoulos, Y. Rubecksen, K. Hellebo R. Houlberg, L. Salomonsen, H. Smullen, A. Wood, M (2020) Understanding Felt Accountability: the institutional antecedents of the felt accountability of agency-CEOs to central government *Governance* 34 (3) 893-916 <https://doi.org/10.1111/gove.12547>

Selwyn, N. (2018). Technology as a focus of education policy. In R. Papa, & S. W. J. Armfield (Eds.), The Wiley Handbook of Educational Policy (pp. 459-477). John Wiley & Sons.

Schuilenburg M. and Peeters, R. (2021) *The Algorithmic Society* London Routledge

Sky News (2020) Coronavirus: SAGE membership revealed after criticism over transparency. 4 May. <https://news.sky.com/story/coronavirus-sage-membership-revealed-after-criticism-over-transparency-11982941>

Sparrow, M (2000) The Regulatory Craft Washington, Brookings Institution Press

Stone D (2013) Knowledge Actors and Transnational Governance: The Private-Public Policy Nexus in the Global Agora Palgrave Macmillan

Stone, D (2019) Making Global Policy Cambridge, Cambridge University Press

Taylor, R. and Opposs, D. (2018). Standard setting in England: A levels. Chapter 6 in Baird, J., Isaacs, T., Opposs, D. and Gray, L. (Editors) *Examination standards: How measures and meanings differ around the world*. UCL IOE Press.

UNESCO (2020) COVID-19. A glance of national coping strategies on high-stakes examinations and assessments. Working Document. 11

April. https://en.unesco.org/sites/default/files/unesco_review_of_high-stakes_exams_and_assessments_during_covid-19_en.pdf

Weale S, 'Pressure grows on ministers over England A-level results "mess"', The Guardian, 11 August 2020, www.theguardian.com/education

Wikström, C. and Wikström, M. (2006) Grade inflation and school competition: an empirical analysis based on the Swedish upper secondary schools. *Economics of Education Review*, 24, 3, 309 – 322.

Weible, C et al. (2020) Covid-19 and the Policy Sciences: initial responses and reactions. *Policy Science*

Williamson, G (2020) Direction issued to the Chief Regulator of Ofqual DIRECTION UNDER S 129(6) OF THE APPRENTICESHIPS, SKILLS, CHILDREN AND LEARNING ACT 2009
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/877611/Letter_from_Secretary_of_State_for_Education_to_Sally_Collier.pdf

Wodak, R (2014) Political Discourse Analysis: distinguishing between frontstage and backstage contexts in Wodak R. and Flowerdew J. (eds) Discourse in Context London Bloomsbury

YouGov (2021). Perceptions of AS and A levels, GCSEs, and Applied General qualifications in England – Wave 19. Ofqual/21/6777/1. April.
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/980782/Perceptions_Survey_Wave_19_-_Report.pdf