



Burgess, S., & Thomson, D. (2019). The impact of the Wolf Reforms on education outcomes for lower attaining pupils. *British Educational Research Journal*, 45(3), 592-621. <https://doi.org/10.1002/berj.3515>

Peer reviewed version

License (if available):
Other

Link to published version (if available):
[10.1002/berj.3515](https://doi.org/10.1002/berj.3515)

[Link to publication record in Explore Bristol Research](#)
PDF-document

This is the accepted author manuscript (AAM). The final published version (version of record) is available online via Wiley at <https://doi.org/10.1002/berj.3515>. Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available: <http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

The impact of the Wolf Reforms on education outcomes for lower attaining pupils

August 2018

Abstract

Between 2004 and 2012, there had been a significant rise in so-called 'equivalent' qualifications taken by young people in England in the final years of their compulsory schooling. These were qualifications other than GCSEs that were approved for young people under the age of 16 under Section 96 of the Learning and Skills Act 2000. This rise was checked by the Wolf Review of vocational education, which led to wholesale changes in the set of qualifications that schools offered to pupils, and which primarily affected lower-attaining students. We quantify the impact of this reform on these pupils on the qualifications they entered and what they achieved. We do so by comparison to previous national cohorts of pupils unaffected by the reform, including the estimation of counterfactual outcomes for the group of pupils most likely to be affected. These pupils tended to enter fewer qualifications overall after the reform than their predecessors, but with a higher fraction of GCSEs. Age-16 attainment fell, with a lower percentage achieving level 2 of the NQF. This finding is repeated in the post-16 outcome measures, which were stable throughout this period. There is no evidence from the attainment data that these reforms helped low-attaining pupils.

Keywords: Wolf Report, School Accountability, England, Pre-16 Qualifications

We are very grateful to the Department for Education for access to the National Pupil Database.

1. Introduction

The education system in England has always struggled to find the best policies for pupils at the lower end of the ability range. At different times, a variety of different policies have been suggested and tried. Should vocational subjects be taught in school alongside academic subjects or in other institutions? Should we have a sharper academic/vocational divide as some European countries do? What are the best qualifications for pupils who are less likely to prosper on more university-oriented courses? How should we encourage schools to offer these and pupils to take them? And, coincidentally or not, England has a long tail of pupils finishing education with low skills (Kuczera et al, 2016, as part of the OECD Skills Study). This long tail costs the economy a great deal and contributes to inequality and poverty. It is now recognised as a matter of serious policy concern.

In this paper, we study the aftermath of a recent significant change to the landscape of qualifications designed specifically for the lower end of the pupil attainment range. These are the so-called 'equivalent' qualifications at age 16 – qualifications other than GCSEs that are approved for young people under the age of 16. As part of the process of encouraging schools to offer them, they were scored to determine their 'equivalence' to a GCSE for inclusion in the school Performance Tables, the main channel of school accountability in England; for example, a pass in a BTEC National Diploma was considered to be equivalent to grade C passes in 4 GCSEs. However, it was argued that some schools began entering pupils in these equivalent qualifications for whom they were not designed, as a cheap way of boosting Performance Tables rankings. In 2011, the government commissioned Professor Alison Wolf to review vocational education in England for 14-19 year olds, and her recommendations to reduce the value of some qualifications, dropping some entirely from Performance Tables, were swiftly enacted in full.

We examine what happened to the first cohort of pupils after this reform had been implemented, and focus on the group that were the intended beneficiaries of the equivalent qualifications, predominantly lower ability pupils. There are two component parts to the outcome, determined by different players. The range of qualifications offered to a pupil with a particular academic profile is largely under the control of the school, so the role of schools in offering a portfolio of qualificationsⁱ is our first focus. We look at which schools were offering a lot of equivalent qualifications, and then explore their heterogeneous reactions to the Wolf reforms: to what degree they switched pupils to the more academic GCSEs, to Wolf-approved equivalent qualifications, or to nothing. The second component is how well the pupils do on the new portfolio of qualifications they are entered for. The combination of these two factors determines the overall outcome for the pupils most affected by the reform. We define these below as the Wolf-Relevant Group (WRG), pupils who would otherwise have taken many of the now-disqualified equivalent qualifications.

In principle, the ideal way of evaluating this change in qualification structure would be to estimate the impact on lifetime earnings or wellbeing. Obviously, we cannot wait that long, so the usual shortcut is to estimate the impact on educational attainment as a good predictor of those long-run outcomes. This approach is complicated in this case, as the reform is predicated on the idea that some of the qualifications are in fact valueless. We deal with this by focussing on the pupils' post-16 attainment as a summary measure of the overall outcome. This is relatively stable in perceived value over the period of the study. The overall outcome could be positive or negative: first, if these pupils are switched to other, better qualifications once the valueless ones are identified, then they would be better off; second, if these qualifications are not replaced by anything else, or are replaced by

harder qualifications that the pupils fail, they would have fewer qualifications overall and be worse off.

We find that following implementation of the recommendations of the Wolf Report, lower-attaining pupils are taking fewer qualifications overall. They are tending to enter more academic qualifications (GCSEs), but as grades achieved in academic qualifications tend to be lower than those achieved in the non-academic equivalents pursued previously, there has been a drop in overall headline attainment. In terms of our summary measure of post-16 outcomes, there does not appear to be much significant change. There is certainly no evidence from the stable post-16 attainment outcomes to suggest that the Wolf reforms significantly helped low-attaining pupils. Indeed, a lower percentage of focus pupils had achieved level 2 by age 18. Clearly, a full evaluation of the reforms should also consider early labour market outcomes, but the initial impact on their attainment is not promising.

The paper is organised as follows: in the next section we briefly revisit the background to the Wolf reform and its content. Then we describe the data and our methodology. We present our results in Section 4, and offer some broader conclusions in the final section.

2. Background to the Wolf reform

By international standards, England is unusual in that its system of high-stakes accountability based on examination results gives the government an “exceptionally firm hand” and leads to a “teaching agenda dominated by what is required for examinations success” (Creese and Isaacs, 2016).

School Performance Tables, often pejoratively labelled “league tables” had first been published in 1992 as an attempt to introduce market-based accountability to raise standards in schools (Allen et al, 2014). Initially, they provided a range of summary information about attainment and the characteristics of the pupil intake (DfE, 1994). Towards the end of the 1990s, measures of progress were introduced, which controlled for attainment at an earlier point in time (DfEE, 1998). Published performance indicators (and others) were used during school inspections by the Office for Standards in Education (OFSTED), the schools inspectorate.

The positive and negative incentives introduced by Performance Tables have been the subject of research for some time (e.g. West and Pennell, 2000). However, the Coalition Government of 2010 to 2015 was the first to use the machinery of Performance Tables to incentivise the take-up of specific qualifications by pupils beyond the so-called ‘basics’ of English and mathematics (Parameshwaran and Thomson, 2015). This was first attempted through the introduction of the English Baccalaureate (EBacc) in January 2011 as a means of encouraging a more traditional curriculum in schools (DfE, 2010). The EBacc is attained by studying GCSEs in certain subjects: English, mathematics, two sciences, a humanities subject, and a foreign language. The percentage of pupils achieving at least a Grade C in six EBacc subjects is now included in school Performance Tables. Incentives to enter pupils for the EBacc subjects were later strengthened by the introduction of Progress 8 as the headline measure of school accountability in 2016 (DfE, 2017a).

Following *Curriculum 2000*, and from around 2004 onwards, there had been a significant rise in so-called ‘equivalent’ qualifications taken by young people in England in the final years of their compulsory schooling. These were qualifications other than GCSEs that were approved for young people under the age of 16 under Section 96 of the Learning and Skills Act 2000. Through a system of

points scores and qualification 'sizes', the equivalence of each of these qualifications to GCSEs was determined for Performance Tables purposes. These allowed some pupils, particularly middle attainers, to achieve level 2 of the National Qualifications Framework (NQF), the equivalent of 5 A*-C passes at GCSE and therefore progress to level 3 study post-16 (Hodgson & Spours, 2014).

Concerns were expressed (noted in Wolf, 2011) that some schools were inappropriately entering some pupils in these equivalent qualifications to boost league table rankings rather than improve pupils' further learning and employment prospects. This was subsequently supported by research conducted on behalf of the Department for Education by the Institute for Fiscal Studies (Jin et al, 2011). Studying core subjects such as languages, English, maths and science is related to long-term economic outcomes (Iannelli, 2013).

In response to these concerns, Professor Alison Wolf was commissioned to review the provision of vocational education between the ages of 14 and 19. Reporting a matter of months later in March 2011, the foreword by the then secretary of state for education Michael Gove set the tone:

"She starts by confronting us with some stark truths. Far too many 14-16 year olds are doing courses with little or no value because performance tables incentivise schools to offer these inadequate qualifications." (Wolf, 2011, p. 4)

The Department for Education (DfE) accepted the 27 recommendations of the report in full. The recommendations worked through revising school accountability – that is, by ruling some qualifications as ineligible for the school performance tables:

"1. The DfE should distinguish clearly between those qualifications, both vocational and academic, which can contribute to performance indicators at Key Stage 4, and those which cannot. The decision criteria should be explicit and public. They will include considerations of depth and breadth (including consultation with/endorsement by relevant outside bodies), but also assessment and verification arrangements which ensure that national standards are applied to all candidates.

3. Non-GCSE/IGCSE qualifications from the approved list (recommendation 1 above) should make a limited contribution to an individual pupil's score on any performance measures that use accumulated and averaged point scores. This will safeguard pupils' access to a common general core as a basis for progression. At the same time, any point-based measures should also be structured so that schools do not have a strong incentive to pile up huge numbers of qualifications per pupil, and therefore are free to offer all pupils practical and vocational courses as part of their programme." (Wolf, 2011, p. 13)

The decisions on which qualifications would be retained was done on a bureaucratic rather than market return basis. The latter approach would have been appealing as it would have been based on evidence that the qualifications truly were valued by employers. But data were not available to provide reliable estimates of such rates of return; consequently, qualifications were deemed acceptable or not by committee. The DfE published technical guidance (DfE, 2015) for awarding organisations setting out the requirements against which qualifications would be deemed suitable for inclusion in Performance Tables:

"From 2014, the Key Stage 4 Performance Tables will be restricted to qualifications that are high quality, rigorous and enable progression to a

range of study and employment opportunities for the **majority of pupils**”
(our emphasis).

Awarding organisations were invited to propose existing qualifications for approval. These were reviewed by officials against the criteria specified in the technical guidance, resulting in a list of approved qualifications (DfE, 2012a) that would be counted in 2014 Performance Tables. The technical guidance also clarified that: qualifications could only be counted if they were at least the size of a GCSE; and, only a maximum of two qualifications per pupil that were not GCSEs, established international GCSEs (iGCSEs) or AS levels could be included in the headline 5 A*-C including English and maths measure.

So although the Wolf Report was ostensibly concerned with vocational qualifications, the response to its recommendations led to a whole suite of qualifications no longer being counted in Performance Tables. While most of these were vocational qualifications many were not. For instance, the hitherto popular short course GCSE in religious studies, equivalent to half a GCSE, was removed.

It should be emphasised here that schools were not ‘banned’ from entering pupils in these qualifications, simply that they were ineligible for Performance Tables purposes. According to the technical guidance accompanying the publication of the 2014 list of approved qualifications:

“Schools may offer qualifications that are approved for teaching pre-16 but are not included in the performance tables, and are encouraged to do so where they judge this to be in the best interests of a particular pupil.”

3. Data and Methodology

We describe the data used, our characterisation of different qualifications, and our method for identifying the group of pupils of most interest.

a) Data: The National Pupil Database

We use the National Pupil Database (NPD; DfE, 2017). This covers all pupils in all state schools in England, and is now available for over a decade. We focus on the three years around the reform, 2012 and 2013 from the pre-Wolf regime and 2014 when the recommendations for reform of Performance Tables came into effectⁱⁱ.

Data on pupils from NPD, includes enrolment history, demographic characteristics, and a full history of test scores at ages 6, 10 and 15 including all outcomes in approved qualifications at Key Stage 4. Demographic characteristics include gender, ethnicity, month of birth, whether the pupil has a statement of special educational needs, and the standard measure of poverty used (eligibility for free school meals). We use the standard Key Stage 2 tests (KS2) taken at age 10 at the end of primary school in maths, English and Science to capture prior attainment. Key Stage 2 tests are externally set and marked. Most General qualifications, such as GCSEs, taken during the period would have been composed of several externally marked papers. Some may have had internally-marked controlled assessments and coursework subject to external moderation from the awarding organisation through a process of sampling. Many popular vocational qualifications were achieved by completing assignments without any examinations.

We limit our analysis to pupils attending state-funded mainstream schools, excluding pupils attending special schools, independent (private) schools and other settings. We exclude a very small number of schools for which we do not have three years of KS4 data, such as recently opened schools whose first intake reached the end of Key Stage 4 in 2014.

We also observe the post-16 study options of our pre- and post-Wolf cohorts. These have been constructed using data on learning aims available from the National Pupil Database (for those in schools) and from the Individualised Learner Record (for those in Colleges and other funded training providers).

Using both data sources, we identify each pupil's:

- a) Highest National Qualifications Framework (NQF) level of study observed in the year after KS4 and
- b) Highest NQF level of *sustained* study observed in the year after KS4

Sustained study relates to learning aims on which a pupil was enrolled for at least 180 days continuously.

We also calculate summary measures of attainment at ages 17 and 18 using data from:

- Key Stage 5 Performance Tables (which collects data on the achievements of 17-19 year olds in NQF level 3 qualifications)
- Key Stage 4 Performance Tables (the PT process collects data on the achievements of 17 year olds at level 2 and below in approved qualifications)
- Individualised Learner Record (which may contain additional records not contained in the above, e.g. related to apprenticeships or from smaller awarding organisations)
- The Level 2 and 3 by age 19 dataset, a composite dataset compiled from the above sources plus awarding body data.

b) Characterising Qualifications

Following the Wolf Review, qualifications can be classified into three broad categories:

- Academic qualifications, including GCSEs, AS level and level 1/ level 2 international certificates (**Academic, shortened to ACA**)
- Other types of applied and vocational qualification, such as BTEC and vocational GCSEs, which can be counted in Performance Tables (**PT eligible, PTE**)
- Qualifications that are approved for use pre-16 under Section 96 of the Learning and Skills Act 2000 but which no longer counted in Performance Tables from 2014 onwards (**PT ineligible, PTI**)

We sometimes use the expression “non-GCSEs” to describe PTE+PTI combined.

We then apply our three broad categories of qualifications. This is straightforward for 2014 but less so for 2012 and 2013. Firstly, many international GCSEs were approved for pre-16 use in 2012 and 2013 but not in 2014. We therefore treat them as academic qualifications in the years in which they were approved. Secondly, some legacy non-academic qualifications were succeeded by replacements that could be counted in Performance Tables in 2014. We treat these as PT ineligible qualifications for the purposes of this analysis. This may under-estimate numbers of PT eligible qualifications in 2012 and 2013.

Our primary focus is the set of PT ineligible qualifications. In 2012 and 2013 this included almost 4000 individual qualifications entered by pupils reaching the end of secondary education although many were entered by just a handful. A list of the most common are presented at Appendix 1, the majority of which were at level 2 of the National Qualifications Framework, equivalent to grades A*-C at GCSE.

Substantial numbers of pupils took these qualifications. Pupils who reached the end of Key Stage 4 in state-funded mainstream schools in 2012, whose qualification pathways would have already been established by the time the Wolf Report was published, entered some 1.5 million qualifications (equivalent to 1.3 million GCSEs) that were not on the list of approved qualifications for 2014. The average number of total qualifications (measured in GCSE equivalents) entered per pupil reached a peak of 11.8 in 2012, since when it has begun to decline (Parameshwaran & Thomson, 2015). Of these 11.8 qualifications, 2.4 were in qualifications that would be ineligible for Performance Tables in 2014. These qualifications accounted for at least half of the total qualification portfolio of around 55 thousand pupils, roughly 10% of the national cohort.

Further changes were made to Performance Tables calculations, such as no qualification was to be counted as more than one GCSE. Previously under the 'old rules' some qualifications counted as up to 4 GCSEs. The reasons why were not always clear: many qualifications formerly equivalent to 4 GCSEs in Performance Tables were designed to be completed in around 360 guided learning hours (equivalent to 3 GCSEs). In other words, the size of a qualification in GCSE equivalents did not necessarily reflect its recommended length of teaching time.

In addition, 2014 saw the phased introduction of a 'first entry' rule. Whereas previously a pupil's best result in a qualification was counted if taken more than once, henceforth only the first entry counts. This move was designed to counter a tendency for some schools to repeatedly enter some pupils, particularly in English and mathematics (Taylor, 2016), in order to "bank" a grade C, the key threshold in the prevailing accountability framework.

Finally, the number of non-GCSE qualifications that could be included in each pupil's set of results was capped at two.

All of these reforms taken together had a significant effect on qualification entries across the entire national cohortⁱⁱⁱ. This effect can be seen by applying both the 2013 and 2014 Performance Tables rules to the 2014 data (Table 1). Entries fall from 10.9 to 9.1 overall, and from 10.4 to 8.2 for pupils eligible for Free School Meals element of the Pupil Premium (FSM6). Table 1 also shows the effect of the various rule changes on pupil attainment measured by average point scores. The impact for English and maths GCSEs is modest: the difference in 0.6 points observed for English is equivalent to one-tenth of a grade. However, the average total points score under the 2014 rules is lower by 75.3 points. Given the difference in entries under both rule systems of 1.8, the drop of 75.3 points is equivalent to 41.8 points per entry, slightly higher than a C grade in a GCSE (40 points).

TABLE 1 ABOUT HERE

c) Identifying the focus group of pupils

While obviously the new rules on qualifications no longer considered suitable applied to everyone, they were only strongly relevant to a minority, to pupils taking a lot of those qualifications. We therefore focus on that group and document the impact of the reform on the pupils it most strongly affected. We identify the segment of the pupil population most affected by the Wolf reforms as pupils who, in the absence of the reforms, would otherwise have been entered for a substantial number of PT ineligible qualifications. We denote this as the Wolf-Relevant Group (WRG). Our

approach is to define this group pre-reform, and estimate the pupil characteristics associated with it. We can then predict which pupils would have been in the group post-reform.

We define this group as pupils who were both:

- Entered for qualifications classified as PT ineligible in 2014 that were equivalent to more than 3 GCSEs; and
- Entered for the equivalent of fewer than 8 academic (ACA) and PT-eligible (PTE) qualifications in total

Using the 2012 cohort of pupils, we estimate the likelihood of being in WRG as a function of prior attainment, FSM eligibility, SEN, gender, ethnicity and school fixed effects (details provided at Appendix 3). Predictions from this model are then applied to the 2013 and 2014 cohorts of pupils. The use of school fixed effects implies that if no pupils were in the target group in 2012 at a school then pupils have a 0% probability of being in the target group in 2012 and in subsequent years. We use this model to estimate relative likelihoods of pupils being in WRG and then calibrate the cutoff into a binary WRG indicator to match the same percentage as in 2012, 13%.

The key factors raising the likelihood of being in WRG are poverty – 20% of pupils eligible for the Pupil Premium are WRG – and low prior attainment. In 2012, we correctly predict 53% of those in WRG to be in, and of those not in WRG we correctly predicted 93%. Appendix Figure S1 shows, the modelled distribution has slightly lower prior attainment than the actual distribution. Over half of pupils in the modelled distribution in 2012 were in the lowest two deciles of Key Stage 2 prior attainment (Figure 1). In other words, pupils in WRG tended to have low prior attainment but this was not universally the case.

FIGURE 1 HERE

4. Results

We first describe the variation between schools in the types of qualifications in which they entered pupils in 2012. We examine the association between the characteristics of schools and the mean number of PTI qualifications entered per pupil. Second, we examine the responses of schools to the reforms in 2014. We consider the heterogeneity in schools' responses based upon the percentage of pupils at each school who are in WRG. Third, we explore the impact of the reforms on the pupils in WRG, their qualification entries, and the results they achieve. Finally, we estimate the impact on their post-16 attainment as an overall summary measure of the reform.

a) The qualifications entered by schools in 2012

We begin by examining the variation between schools in PTI qualifications entered in 2012 prior to the reforms. We then examine the characteristics of schools which were associated with entry in PTIs before examining the variation between schools in percentages of pupils in the Wolf-Relevant Group.

Figure 2 shows how schools varied in the mean number of PTI qualifications entered by pupils in 2012. Almost 80 schools entered pupils for at least 5 PTI qualifications (measured in GCSE

equivalents) on average. Entry in non-GCSEs was negatively correlated with entry in GCSE and PTE qualifications combined ($r=-0.4$).

FIGURE 2 HERE

We then regress the school-level means of the number of entries in PTI qualifications in 2012 on a range of school-level characteristics and previous performance data. These include measures of the composition of the pupil body, school governance, Ofsted inspection judgments (at the start of the 2011/12 academic year), region and the number of schools within a 1 kilometre radius, an indicator of local competition. The main effects from the regression are shown in Table 2.

TABLE 2 ABOUT HERE

The mean number of non-GCSEs entered per school is positively correlated with school-level disadvantage (%FSM) and negatively correlated with mean KS2 of the intake and the percentage of pupils with English as an additional language (%EAL). It was also negatively correlated with performance three years earlier, suggesting that- for some schools at least- entering more PTI qualifications could be a response to previous poor performance.

In addition, schools in the North East and sponsored academies were more likely to enter pupils for larger numbers of non-GCSE qualifications, all other things being equal. There did not appear to be any association with increased competition, defined by the number of schools within a 1km radius.

Wolf-Relevant Group

There was marked variation between schools in the percentage of pupils in WRG (Table 3). In 2012, there were over 1200 schools with none whatsoever, 40% of all schools in our analysis. By contrast, there were 119 schools at which more than 80% of pupils were WRG. Schools in the latter group tended to have pupil intakes with lower prior (KS2) attainment than other schools.

TABLE 3 ABOUT HERE

Confirming the findings of Jin et al (2011), the higher propensity to enter pupils in 2012 for relatively large numbers of the qualifications that subsequently became PT ineligible appears to be associated to some extent with an increased risk of falling below the government's floor standards for secondary schools (Table 4). Although expected progress indicators in English and maths featured in the definition for floor standards (DfE, 2013), it was largely driven by falling below 40% of pupils achieving 5 or more A*-C grades including English and maths (AC5EM). Schools with larger proportions of pupils in WRG were more likely to fall below this threshold in 2012.

TABLE 4 ABOUT HERE

But the threat of floor standards cannot be considered the sole motivation for the qualification entrance policies of schools with lots of pupils in WRG. The floor standard was, effectively, based upon achieving grade C or above in English and maths. The percentage of pupils who did so but failed to achieve another 3 grade C passes (or equivalent) in other subjects was rather small- less than half of one percent of pupils in 2012. In 2012, over half of the schools with 80% or more of WRG pupils achieved 50% or more A*-C grades in both English and maths.

One other explanation may be seeking enhancement to league table rankings and, concomitantly, the judgment of outstanding (or even good) in Ofsted inspection. Ofsted now admit that schools with lower attaining intakes are less likely to be judged good or outstanding^{iv}. Some schools may

have pursued more favourable inspection outcomes by demonstrating good rates of progress from Key Stage 2 to Key Stage 4. According to the Department for Education's main value added performance indicator, based on attainment in pupils' best 8 GCSEs (or equivalents), schools with a high proportion of pupils in the Wolf-relevant group achieved a value added score 0.12 standard deviations above the national average in 2012 (Table 5).

However, this advantage was not present in the value added indicators for English and mathematics, suggesting that the enhanced level of performance on the main value added indicator was at least in part related to the nature of qualifications offered by schools with a high percentage of pupils in WRG.

TABLE 5 ABOUT HERE

The percentage of WRG pupils at each school in 2012 was also correlated ($r=0.5$) with multiple entry in GCSE mathematics, another method used by some schools to improve performance measures that was tackled in the Government's response to the Wolf Review (Taylor, 2016).

b) Schools' responses to Wolf reforms, 2012 to 2014

We now turn to how schools responded to the Wolf reforms in terms of the types of qualifications they entered pupils in. We then examine how schools with the largest proportions of pupils in WRG responded, and how this varied by pupil prior attainment.

Changes in qualifications entered

The response of schools to the reforms has been to enter pupils in more academic qualifications at the expense of non-GCSE qualifications. Figure 3 shows the mean number of non-GCSEs (i.e. PTE+PTI qualifications combined) at each school in both 2012 and 2014. With some exceptions, the 2014 school means are lower than the 2012 school means, particularly among the schools with higher means in 2012. This contrasts with an increase in academic qualifications entered over the same period (Figure 4).

FIGURE 3 HERE

FIGURE 4 HERE

There was an inverse correlation ($r=-0.56$) between the change in the mean number of non-GCSE qualifications entered and the mean number of academic qualifications entered (Figure 5). On the whole, schools which tended to enter pupils in fewer non-GCSE qualifications also tended to increase the number of academic qualifications entered. Using the prevailing system of measuring qualification size in GCSE equivalents, the increase in academic qualifications was smaller than the decrease in non-GCSE qualifications, but this may simply reflect the fact that the prevailing qualification sizes did not adequately reflect the amount of curriculum time required to deliver them.

FIGURE 5 HERE

Schools with large percentages of pupils in WRG

In 2012, there were 232 schools with at least 60% of pupils in the Wolf-Relevant Group (Table 3). For the most part, their response to the Wolf reforms was similar to other schools. The mean number of entries per pupil fell from 12.7 in 2012 to 11.2 in 2014 although the mean number of GCSEs entered increased from 5.8 to 7.2 (Table 6). We observe schools moving away from PTI qualifications after 2012, mainly into PTE qualifications in 2013 and then more towards GCSEs in 2014. On the whole, such schools still tended to enter pupils for more PTE and PTI qualifications than other schools.

TABLE 6 ABOUT HERE

Some heterogeneity in school responses can be observed. Some schools switched to predominately entering pupils for GCSEs and other academic qualifications, whereas others made greater use of vocational qualifications. Those schools which entered pupils for a higher fraction of academic qualifications typically entered pupils for fewer qualifications overall (Figure 6).

FIGURE 6 HERE

We then examine the types of qualifications entered in a finer level of detail (Table 7). Overall, there was a 12% decrease in qualifications entered measured by GCSE equivalents but only a 5% drop in qualifications measured by counting individual qualifications. There was an increase in academic entries in both EBacc and non-EBacc subjects. There is also some evidence of schools switching from PTI BTEC/OCR qualifications to PTE versions. Entries in PTI qualifications at level 1 and below fell at a slower rate than entries in PTI qualifications at level 2, suggesting that some schools at least continued to offer them to lower attaining pupils.

TABLE 7 ABOUT HERE

Within the 232 schools, the impact of the Wolf reforms varied according to pupil prior attainment. In Figure 7, pupils have been grouped into deciles based on their Key Stage 2 (age 11) mean test score. In general, numbers of entries in GCSEs increased across the prior attainment distribution but more so in the middle. Whereas in 2012, non-GCSE (PTE+PTI) qualifications accounted for 40% of the qualifications entered by pupils with the highest levels of prior attainment, by 2014 this figure had almost halved.

FIGURE 7 HERE

c) The outcome for WRG pupils at age 16, Key Stage 4

We now turn to pupils in the Wolf Relevant Group as determined above. Firstly, we examine changes in numbers of qualifications entered in 2012, 2013 and 2014 at a broad level before secondly considering a) a more granular classification of qualifications and b) subjects entered. Thirdly, we show the impact on headline attainment for the group.

Number of qualifications entered by WRG

Overall, the mean number of qualifications (in GCSE equivalents) entered by the WRG group fell from 11.6 to 10.6 (Table 11). However, this masks an increase, from 4.9 to 6.2, in the number of academic qualifications entered. The mean number of entries in PTI qualifications fell from 4.6 to 1.4.

This switch resulted in a dip in the APS per entry from 36.7 to 35.1 (equivalent to over a quarter of a grade in a GCSE) and in the percentage of entries passed at grade C and above (or equivalent) from

63% to 56%. There was little overall change in attainment in academic qualifications despite the increase in entries. This is perhaps to be expected, given the method of comparable outcomes (Benton, 2016) employed by Ofqual and awarding bodies to ensure consistency in grading from year to year.

TABLE 8 ABOUT HERE

As we showed (Figure 1), the WRG tend to be located in the lower deciles of prior attainment. Nonetheless, the overall increase in entries in academic and PTE qualifications is broadly consistent across the prior attainment range (Figure 8). One slight exception is that the increase in academic entries among the lowest decile (0.9) is slightly lower than the average for deciles 2-10 (1.4). By 2014, pupils in the lowest decile were the most likely to be entered for PTI qualifications. Far larger decreases in mean entries in these qualifications can be observed among pupils with higher levels of prior attainment.

FIGURE 8 HERE

One might have expected *ab initio* the PTI qualifications to be replaced by PTE analogues. The reason why this did not occur is likely to be due to the changes to the way PTE qualifications were valued in 2014 Performance Tables. Firstly, a maximum of 2 non-academic qualifications per pupil could be counted in headline performance measures. Secondly, no qualification would be counted as more than one GCSE and this may have acted as a disincentive to deliver longer courses. Finally, a raft of hitherto popular qualifications counted as less than one GCSE in size were henceforth no longer counted. These included short course GCSEs (mostly in religious studies and citizenship), and basic and functional skills in literacy and numeracy.

Qualifications and subjects entered

We now examine specific qualifications and subjects entered in slightly greater detail. We recapitulate here that for comparative purposes we have used the pre-2014 qualification sizes in our analysis. The 209 thousand entries in PTE qualifications in 2014 observed in Table 8 were counted as 109 thousand entries in 2014 Performance Tables using the post-Wolf qualification sizes.

Expanding the broad classification of qualifications used so far (e.g. Table 8), we observe that the most popular type of PTI qualifications entered in 2012 by WRG pupils were level 2 BTEC and OCR National qualifications (Table 9). These are sector specific vocationally-related qualifications awarded by two of the major awarding organisations; Edexcel in the case of BTEC and OCR in the case of OCR Nationals. Entries in PTI versions fell from 192 thousand GCSE equivalents in 2012 to 31 thousand in 2014. By contrast, entries in PTE versions increased, though by a smaller margin, from 129 thousand to 193 thousand. Entries in academic qualifications, in both EBacc and non-EBacc subjects also increased.

There were also smaller reductions in hitherto less popular PTI qualifications, particularly Key Skills, Functional Skills and Basic Skills. For the most part, these were taken in literacy and numeracy although they were available for ICT and work-related skills (e.g. problem solving). Key Skills in particular came in for stringent criticism in the Wolf Report (Wolf, 2011, p.84). The associated dip in APS per entry in these qualifications is largely the result of the lower-scoring Functional Skills composing a much greater proportion of the entries in 2014 (75% in 2014 compared to 46% in 2012).

TABLE 9 ABOUT HERE

The switch to academic qualifications was broadly even between Ebacc and non-Ebacc subjects. The Ebacc itself does not seem to have been a major policy driver. 4% of the WRG group were entered for all six subjects of the Ebacc in 2012 but this only increased to 12% in 2014.

The broad changes in attainment and entries by qualification type observed in Table 9 can be expanded to look at individual subjects (Table 10). In some cases, we see that a decline in entrants was offset by an increase in entrants for analogous qualifications that were counted in 2014. This was particularly true in sports studies and ICT/computer use. But we also observe a decline in 'large' qualifications (equivalent to >2 GCSEs), perhaps in response to changes to rules for Performance Tables calculations as a result of which they are now counted as equivalent to a single GCSE. For example, entries in applied sciences have also declined although there is also evidence of schools switching from PTI qualifications equivalent to 2 GCSEs to the BTEC first award (PTE) equivalent to a single GCSE. Overall, the proportion of pupils entered in applied sciences began to decline from 2013 having risen markedly during the previous 8 years, particularly for pupils eligible for free school meals. This is the corollary of schools tending to revert to GCSEs in science (Parameshwaran & Thomson, 2015).

We also observe a large proportion of pupils switching from GCSE short course religious studies (PTI) to the full course (academic). There is also a switch from the single GCSE in combined English towards twin GCSEs in language and literature. The former would not be available for teaching from September 2015. There were small increases in entry rates in Ebacc subjects such as geography, history, French and Spanish.

There were no clear systematic changes in attainment in individual subjects. In some, such as English language and geography, attainment increased as the entry rate increased. However, the inverse was true in other subjects such as religious studies and English literature.

TABLE 10 ABOUT HERE

Impact on headline attainment

The overall impact on attainment for the Wolf Relevant Group was a fall in the percentage of pupils achieving level 2 of the National Qualifications Framework (5 or more A*-C grades at GCSE or equivalent) from 72% in both 2012 and 2013 to 61% in 2014 (Table 11). This can be attributed to schools tending to enter pupils for fewer large non-GCSE qualifications and instead pursuing GCSEs, which tended to accrue fewer points in Performance Tables calculations.

However, there was no change (27%) in the percentage achieving this threshold including grade C passes in English and maths. This would tend to suggest that there was no overall change in the attainment profile of this group of pupils and that the fall in the percentage achieving 5 A*-C was due to material changes in qualifications entered.

Attainment in English rose slightly, by one point (a sixth of grade), between 2013 and 2014 although attainment in maths fell by a slightly larger amount. This is likely to be due to the decision taken at the same time to count a pupil's first result in English and maths to avoid multiple resits.

We can only speculate here why the APS dipped in maths. Although we still report the best result for each pupil in Table 11 (Performance Tables in 2014 reported the 'first result'), the accounting change appears to have encouraged schools to enter pupils only once in year 11 (rather than twice or more as in previous years). Table 12 shows that the mean number of entries per pupil in GCSE maths fell from 2.25 among the 2013 cohort to 1.66 among the 2014 cohort.

TABLE 11 ABOUT HERE

d) Outcomes post-16

Finally, we examine the study options and attainment of pupils in the Wolf Relevant Group in the two years following completion of compulsory schooling.

The fall in the percentage of WRG pupils achieving NQF level 2 (5 or more A*-C grades at GCSE or equivalent) noted above would be most concerning if their post-16 study options were restricted as a result. To check this, we constructed post-16 study profiles using data from NPD (post-16 learning aims) for those who stay in the schools system and ILR (learning aims) for those who move into the Further Education (College) sector.

We identify each pupil's

- a) highest level of study observed in the year after KS4 and
- b) highest level of *sustained* study observed in the year after KS4

Sustained study relates to learning aims on which a pupil was enrolled for at least 180 days continuously.

Results for WRG are presented in Table 13 and Table 14. The 2013 KS4 cohort was the first to be affected by the raising of the participation (RoPA) age to 17 (DfE, 2012b). The 2014 cohort was the first to be affected by the raising of the participation age to 18. We do not consider the post-16 study profiles of the 2012 cohort as they were unaffected by RoPA.

TABLE 13 ABOUT HERE

TABLE 14 ABOUT HERE

Compared to the 2013 cohort of WRG pupils, the post-16 study options of the 2014 cohort were broadly similar, for both any post-16 study and sustained post-16 study. A slightly lower proportion of the latter was not observed in education and a slightly higher proportion was studying at NQF level 2. Similar proportions were observed to be studying at level 3 compared to the 2013 cohort.

Overall, attainment in qualifications taken in the year following Key Stage 4 among the WRG for both the pre and post Wolf cohorts was fairly similar although the post Wolf cohort (KS4 Year 2014) was slightly less likely to achieve level 2. In other words, the fact that more of them were observed to be studying at level 2 (Table 14) did not materialise into stronger attainment in the year following Key Stage 4 (Table 15). In fact, just 62% of the 2014 WRG cohort who were studying at level 2 went on to achieve level 2 compared to 74% among the 2013 cohort. Overall though, attainment in the year following Key Stage 4 was broadly similar between the two groups.

TABLE 15 ABOUT HERE

Finally, we turn to pupils' levels of cumulative attainment by age 18, two years after Key Stage 4. Although the 2014 WRG cohort made very similar post-16 choices to their predecessors, proportionally fewer of them had achieved NQF level 2 by age 18 (75% compared to 82%). The 11 percentage point difference in level 2 attainment observed at the end of Key Stage 4 (Table 11) narrowed slightly by age 18 but remained at almost 7 percentage points (Table 16).

TABLE 16 ABOUT HERE

5. Conclusion

The Wolf Review of vocational education led to wholesale changes in the set of qualifications that schools offered to pupils. The policy was implemented by assigning zero value to the ineligible qualifications in school performance tables, thus working through schools' strong focus on this accountability framework. The pupils typically affected were those following a less academic curriculum. We have quantified the impact of this reform on these low-attaining pupils, focussing on the types of qualifications offered and the scores on those qualifications.

Pre-reform, the data show wide variation among schools in the use of the PTI (performance-table ineligible) qualifications, and correspondingly wide variation in the use of GCSEs. Each school could determine a set of qualifications to offer its pupil body, balancing to some extent the needs of learners (progression, motivation) with the demands of external forces (league tables, Ofsted) to boost institutional performance.

The response of schools to the de-valuing of these in performance table terms was also heterogeneous, but on average they substantially reduced PTI entries, and instead offered more GCSEs and more PT-eligible non-GCSEs. However, the increase in academic qualifications was lower than the fall in the non-academic ones so that overall, pupils in 2014 tended to enter fewer qualifications than their predecessors in 2012 measured using GCSE equivalents but this included a higher fraction of GCSEs. While there is some evidence that schools are continuing to make use of PTI qualifications where they feel it is appropriate to do so, entries in such qualifications have fallen markedly overall.

In terms of attainment, the focus group of pupils took more GCSEs but scored about the same or slightly worse on them; for example, their average point score (APS) per entry on GCSE English and Maths fell slightly 32.4 to 32.3. The score on other (non-Ebacc) GCSEs, the academic qualifications that increased the most, fell from 32.5 to 32.2. That said, we should perhaps expect little change in attainment given that the prevailing system of comparable outcomes used by Ofqual and the awarding bodies result in similar distributions of GCSE outcomes conditional on prior (Key Stage 2) attainment from year to year.

Similarly, the PTE qualifications that increased the most, Level 2 BTEC/ OCR National, also saw a fall in APS per entry from 43.3 to 43.1. The overall change in attainment for the focus group was a fall in the percentage of pupils achieving level 2 of the National Qualifications Framework from 72% in 2012 and 2013 to 61% in 2014. This was driven by change in qualifications entered rather than grades achieved.

Of course, it is difficult to judge the meaning of this outcome; if the previous qualifications were inadequate as the Wolf Report contended, then perhaps there is nothing real lost in this apparent fall in attainment.

We therefore turn to the more stable post-16 outcome measures to reach an overall judgement. The fall in the percentage achieving Level 2 would be concerning if this resulted in their post-16 study options being restricted. In fact, post-16 study choices at age 17 among the focus group of pupils was similar pre and post reform, although the post-reform cohort was slightly less likely to achieve level 2 by age 17. The pre- and post-reform difference in level 2 attainment at age 16 closed slightly by age 18 but remained at almost 7 percentage points. Therefore we conclude that the reforms have not delivered any benefit to this group of pupils in the short term.

Finally, we turn to two broader issues that our study bears on. First, these results attest again to the power of the school accountability system. With no changes in the underlying nature of the qualifications, simply assigning some of them a zero 'price' in the school performance tables led to very substantial changes in their use. It could be argued that the Wolf Review gave schools new information on the true nature of these qualifications, but this seems unlikely given that schools had been actively working with them for a considerable time.

Second, the sudden and exogenous change in the value of a qualification allows us to learn a little more about the drivers of school behaviour. We have documented heterogeneous responses – some schools stick with largely a non-academic approach, and simply switch from ineligible to eligible qualifications. Others drop the ineligible qualifications and start pursuing a more academic, GCSE-heavy curriculum. Doing one or the other appears to be idiosyncratic and not associated with a range of school performance and demographic characteristics.

The Wolf reforms significantly changed the landscape of qualifications in England, and certainly had effects on the portfolios of qualifications offered by schools. There is no evidence from the attainment data so far that these reforms have helped low-attaining pupils. Indeed, proportionally fewer young people now appear to be entering the labour market with level 2 qualifications. Furthermore, the ineligible qualifications previously taken by lower attaining pupils may have conferred additional but unobserved benefits in terms of work-related and independent learning skills. Labour market data on employment and earnings may therefore reveal more about the impact of reform on affected pupils.

References

- Allen, R., Burgess, S. and McKenna, L. (2014) *School Performance and Parental Choice of School: Secondary Data Analysis*. Department for Education Research Report. Available online at: [https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/275938/RR310 - School performance and parental choice of school.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/275938/RR310_-_School_performance_and_parental_choice_of_school.pdf) (accessed 10th August 2018).
- Benton, T. (2016) *Comparable Outcomes: Scourge or Scapegoat?* Cambridge Assessment Research Report. Available online at: <http://www.cambridgeassessment.org.uk/Images/346267-comparable-outcomes-scourge-or-scapegoat-.pdf> (accessed 10th August 2018).
- Creese, B. and Isaacs, T. (2016) International instructional systems: How England measures up, *The Curriculum Journal*, 27(1), 151-165.
- DfE (Department for Education) (1994) *School Performance Tables*. Available online at: http://webarchive.nationalarchives.gov.uk/20150812155602/http://www.education.gov.uk/schools/performance/archive/schools_94/sec1.shtml (accessed 10th August 2018).
- DfE (Department for Education) (2010) *The Importance of Teaching: The Schools White Paper 2010*. Available online at: www.gov.uk/government/uploads/system/uploads/attachment_data/file/175429/CM-7980.pdf (accessed 10th August 2018).
- DfE (Department for Education) (2012a) *Key Stage 4 Performance Tables: Inclusion of 14 to 16 qualifications in 2014*. Available online at: [https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/411159/Archived-KS4 performance tables - inclusion of 14-16 qualifications in 2014.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/411159/Archived-KS4_performance_tables_-_inclusion_of_14-16_qualifications_in_2014.pdf) (accessed 10th August 2018).
- DfE (Department for Education) (2012b) *Raising the Participation Age*. Available online at: <https://www.gov.uk/government/collections/raising-the-participation-age>
- DfE (Department for Education) (2013) *GCSE and Equivalent Results in England 2011/12 (revised)*. Available online at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/219341/sfr02_202013.pdf (accessed 10th August 2018).
- DfE (Department for Education) (2015) *Qualifications for 14-16 Year Olds and Performance Tables*. Available online at: <http://dera.ioe.ac.uk/14743/1/2015%20Performance%20Tables%20-%20Technical%20Guidance.pdf> (accessed 10th August 2018).
- DfE (Department for Education) (2017a) *GCSE and Equivalent Results in England 2015/16 (revised)*. Available online at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/584473/SFR03_2017.pdf (accessed 26th November 2018).
- DfE (Department for Education) (2017b) *National Pupil Database: User Guide and Supporting Information*. Available online at: <https://www.gov.uk/government/publications/national-pupil-database-user-guide-and-supporting-information> (accessed 10th August 2018).

DfEE (Department for Education and Employment) (1998) *Key Stage 3 to GCSE/ GNVQ Pilot*.

Available online at:

http://webarchive.nationalarchives.gov.uk/20150812113656/http://www.education.gov.uk/schools/performance/archive/vap_98.shtml (accessed 10th August 2018).

Hodgson, A. and Spours, K. (2014) Middle attainers and 14-19 progression in England: half-served by New Labour and now overlooked by the Coalition? *British Educational Research Journal*, 40(3), 467-482.

Iannelli, C. (2013) The role of the school curriculum in social mobility, *British Journal of Sociology of Education* 34 (5-6), 907-928.

Jin, W., Muriel A. and Sibieta, L. (2011) *Subject and course choices at ages 14 and 16 amongst young people in England: insights from behavioural economics*. Department for Education Research Report RR160. Available online at:

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/182677/DFE-RR160.pdf

Kuczera, M., Field, S. and Windisch, H. C. (2016) *Building Skills for All: A Review of England*. Available online at: <https://www.oecd.org/unitedkingdom/building-skills-for-all-review-of-england.pdf> (accessed 10th August 2018).

Parameshwaran, M. and Thomson, D. (2015) The impact of accountability reforms on the Key Stage 4 curriculum: How have changes to school and college Performance Tables affected pupil access to qualifications and subjects in secondary schools in England? *London Review of Education* 13(2), 157-173.

Taylor, R. C. (2016) The effects of accountability measures in English secondary schools: Early and multiple entry to GCSE mathematics assessments, *Oxford Review of Education*, 42(6), 629-645.

West, A. and Pennell, H. (2000) *Publishing school examination results in England: incentives and consequences*, *Educational Studies*, 26(4), 423-436.

Wolf, A. (2011) *Review of Vocational Education – The Wolf Report*. Available online at:

https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/180504/DFE-00031-2011.pdf (accessed 10th August 2018).

ⁱ Of course, pupils and their families choose schools, but the timing of our study focussing soon after the reform means that the vast majority of pupils in the study had already chosen schools before the revised qualification offers were needed or known.

ⁱⁱ Adding more recent years into the analysis would be of dubious value because of further subsequent changes in the definition of qualifications.

ⁱⁱⁱ However, the impact of the changes may have been even greater. The “Wolf” changes were announced prior to the 2014 cohort beginning Key Stage 4 in September 2012 and the “first entry rules” were announced the following September. This may have influenced the behaviour of schools, with regard to both which qualifications to enter pupils for and when to enter them. Between 2013 and 2014, the percentage of pupils in state-funded mainstream schools achieving 5 or more A*-C grades (or equivalent) including GCSE English and maths fell from 61.9% to 61.6% based on the 2013 rules. This checked a previously monotonically increasing trend in an indicator which had risen 10 percentage points since 2009.

^{iv} This admission was made in correspondence by Ofsted’s Chief Statistician. A copy can be viewed at <https://www.dropbox.com/s/ucls5vl23e6n8f0/Letter%20from%20Robert%20Pike%20Ofsted%27s%20Chief%20Statistician%20and%20Deputy%20Director%20Data%20and%20Insight%20230615%20%282%29.pdf?dl=0> , retrieved 14th August 2018.