Article

Using School-Level Data to Investigate the Impact of a One-to-One Mathematics Teaching Resource in English Primary Schools

Chris Rolph [ID]

Nottingham Institute of Education, Nottingham Trent University, Nottingham NG1 4FQ, UK;
chris.rolph@ntu.ac.uk

Abstract: This research investigates the potential for a one-to-one coaching tool used by adults other than teachers to be able to deliver greater mathematics progress for primary school children without adding significantly to school costs. Plus 1 and Power of 2 (+1 and Po2) are workbooks designed to be used by adults other than teachers working on a one-to-one basis with children to develop numeracy skills. This quantitative study seeks to examine the impact of +1 and Po2 by considering performance data aggregated at the school level. The attainment of children at 1071 English schools which use the +1 and Po2 products was compared with that of peers in other schools using a quasi-experimental research design based on England's national performance measures. Statistical analysis suggests that schools using +1 and Po2 show higher levels of mathematics attainment than those who may have used other resources. Furthermore, there is an important finding that assessment attendance is higher, and disapplication from the curriculum lower, in schools using +1 and Po2. This indicates that use of this one-to-one intervention improves access to national tests for children and represents an opportunity for school leaders to maximise the cost effectiveness of existing non-teaching staff or volunteers.

1. Introduction

It is universally accepted that good literacy and numeracy skills are essential for functioning adults in a developed society [1–3]. The 2012–2018 OECD surveys of adult skills, which included 33 industrialised countries, showed that even in high performing countries such as Japan as many as 10% of the adult population performed at the lowest levels of literacy or numeracy [4], though few adults in developed countries would be described as illiterate. Conversely, even in those countries ranked highest in the surveys there are significant numbers of adults whose skills could be described as functionally innumerate [5]. International data collected between 1994 and 2012 show that, over this period, numeracy proficiency declined everywhere apart from Italy [6]. Remedial adult catch-up programmes have demonstrated some benefit for basic skills, though the evidence of gain is more compelling for literacy than it is for numeracy [7].

Geary argues that the importance of poor mathematical skills is not widely recognised [8], and that the costs associated with this are frequently underestimated and potentially higher than those associated with poor literacy skills. For the individual, an above basic numeracy competency can carry a lifetime earnings premium of up to 10% [9]. For society as a whole, the estimated costs of low levels of numeracy are staggering: a 2014 report suggests that low levels of numeracy may cost the UK economy as much as £20 billion every year [10]. It is therefore unsurprising that national governments might seek effective numeracy interventions in order to mitigate this. A review of 10 years of adult literacy and numeracy intervention programmes in England concluded that the cost effectiveness of such provision was unclear, and that while there were measurable

personal and social gains for adults improving their literacy and numeracy skills, with consequent financial benefits for the individuals, there was insufficient evidence to measure the economic impact of the provision [11]. The situation in England is in line with that of other developed economies, but weak evidence has not diminished the calls for renewed national approaches to address shortcomings in adults' literacy and numeracy skills [12]. However, if it is difficult to make up lost educational ground once people have begun to play a part in a nation's economy it is even more important to ensure that little ground is lost in the first place—in other words, to address the gaps in learning earlier and look at the development of skills amongst younger children. Once an educational achievement gap opens between socioeconomically disadvantaged children and their more affluent peers it becomes increasingly difficult for those left behind to catch up: the so-called "Matthew effect" [13]. This can be compounded if children become anxious about their mathematical ability, generating a perpetual cycle of further underachievement leading to increased anxiety [14]. There is therefore an educational imperative to close achievement and attainment gaps as early as possible, and in particular before children move to secondary school.

In 1999 the UK government introduced a national numeracy strategy for England. The direction of travel was not entirely new, following from previous governments' attempts to influence both the curriculum and the practice in primary schools [15]. General levels of numeracy were raised through the strategy, but there is evidence that it had little impact on some of the weakest children [16]. The national numeracy strategy was ultimately absorbed into a more general framework for primary schools, but the emphasis on mathematical skills was retained and built into national accountability measures for schools, including publicly available league tables and inspection reports. In common with international comparison countries [17,18] England's most disadvantaged children are those whose mathematics attainment lags most behind their peers. To address the impacts of disadvantage, since 2011 English schools have received additional funding targeted at children from poor socio-economic backgrounds, with an expectation that this is used to raise their attainment in tangible ways that are evident in the accountability measures. While school leaders welcome this "pupil premium" funding, they experience additional pressures in the requirement to justify how it has been spent and to demonstrate progress on the part of their disadvantaged children [19]. Similar government policies around the world suggest that simply adding more money into the system, however specifically targeted, does not guarantee improvements for those children who have fallen furthest behind [20,21]. Despite this evidence, English school leaders are given autonomy over the pupil premium spend but must ensure that the funding drives improvements [22], and the need for school leaders to identify cost-effective interventions is therefore crucial.

The injection of additional money into the English educational system spawned a proliferation of off-the-shelf products and strategies which promise learning gains for disadvantaged children as a result of spending the pupil premium funding. To aid school leaders the UK government established the Education Endowment Foundation (EEF) to generate and gather evidence in relation to effective approaches to improve outcomes for disadvantaged children. The EEF publishes an online Toolkit which summarises findings from meta-studies and purpose-designed research studies (often based on randomised control trial investigations). One-to-one tuition by a teacher or other adult is identified by the EEF as having a significant impact, though at a relatively high cost [23]. The evidence for this conclusion is based on several meta-studies, four of which focus entirely on literacy, and seven single studies, only two of which relate to numeracy: the development of mathematical skills is under-represented in this evidence base. A UK government review of literacy and numeracy catch-up strategies demonstrated similar bias, with literacy strategies outnumbering numeracy in the ratio 5 to 1 [24]. There is a pressing need for further studies of mathematics catch-up strategies in England.

2. Methods

Since the 1988 Education Reform Act, schools in England have had to publish data showing the performance of their pupils, as well as data relating to finances, workforce and pupil absences [25]. For primary schools the last dataset released prior to the COVID-19 pandemic (2019) listed more than 300 fields of data for each school, most relating to children's performance in reading, writing and mathematics. These are based on summative teacher assessments and externally set tests at the ages of 7 and again at 11, when they transfer to secondary school. With 15,000 primary schools in England this provides a rich source of data in the public domain which can be investigated by researchers interested in school and pupil performance. For the purposes of these national dataset children are grouped according to disadvantage, which in this case relates to the family income. The definition of disadvantage in English schools refers to families whose income has fallen below a nationally set threshold which enables the children to claim free school meals at any time in the previous 6 years. This entitles them to pupil premium funding. While this is not very sophisticated, it does mean that there is a common understanding of the meaning of disadvantage within the English school system. Similarly, three national prior attainment ability groups are defined by children's performance in assessments at age 7, averaged across reading, writing and mathematics.

This project made use of the publicly available school-level data to investigate the mathematics attainment of children in English primary schools that had used a particular one-to-one catch-up programme. The chosen mathematics intervention comprises two resources which are often purchased together and used sequentially: Plus 1 and Power of 2 (+1 and Po2), produced by 123Learning [26,27]. These consist of workbooks which are designed to be used by parents, teaching assistants or volunteers, and each book contains instruction for the adult, labelled a coach, to avoid the need for expensive training sessions. These particular resources were identified because they are designed to be used with any age group, including adults, though for this research we limit our investigation to the attainment of 11-year-olds. They represent an established resource that has been on the market for several years and have been sold internationally as well as in the UK. Importantly for this study, at the time of data collection over 1000 English schools were using +1 and Po2 allowing for robust statistical comparisons. The intention for +1 and Po2 to be used in one-to-one coaching sessions by adults other than trained teachers—potentially parents or volunteers—represents a relatively low-cost individual tuition programme that may well be attractive to school leaders. The chosen method assumes that schools have used the +1 and Po2 resource to help children to catch up with their peers by the time of their assessment at age 11. This does not imply or assume that the resources were used during children's final year in primary school, though this is likely.

This research used a quasi-experimental design to compare the mathematics performance of children in schools using +1 and Po2 with those not using this intervention. A list of 5000 schools which had purchased +1 and Po2 was compared with those included in the UK government's primary school performance tables for 11-year-olds in 2018; a total of 1071 schools matched. This allowed for independent 2-tailed *t*-tests to be carried out, comparing mathematics attainment of the sample of 1071 schools with the remaining 13,688 English schools not using this particular intervention. Although the UK government publishes progress data as well as attainment data, progress is based only on prior attainment (i.e., with no contextual information) and has been shown to be open to inherent bias. School progress measures "fall far short" of producing fair and robust measures of school performance [28] and are therefore not considered in this study which is restricted to attainment measures only.

The school-level approach permitted high-level comparisons to be made between two groups of schools without the need for any invasive or disruptive research within the schools. No schools were asked to change their behaviour, their approach to teaching, or their use of pupil premium funding. Although the performance data for primary schools is in the public domain with free open access, the study was conducted according to the prin-

ciples of the British Educational Research Association's ethical guidelines for educational research [29] and was approved by the ethics committee of Nottingham Trent University.

3. Results

The result for the key headline indicators for mathematics are shown in Table 1. The first 6 lines of this table show the mean scaled scores for pupil groups in schools that have been using +1 and Po2, compared with the control group of all other schools. The scaled scores are based on externally set and marked tests which the children sit during the summer term of their final year in primary school. Raw scores from these tests are adjusted using a non-linear 80–120 scaling such that a score of 100 equates to the "expected standard" and a score of 110 or more is defined as a high score.

Table 1. Comparative performance measures for +1 and Po2 schools and all other schools.

Row	Performance Measure	Used +1 and PO2?	N	Mean	Significance [1]	
1	Average scaled score in mathematics	No	13,672	104.26	0.000	**
		Yes	1068	104.62		
2	Average scaled score in mathematics for pupils with low prior attainment	No	2842	94.2	0.235	
		Yes	219	93.8		
3	Average scaled score in mathematics for pupils with medium prior attainment	No	12,975	102.7	0.065	
		Yes	1037	102.9		
4	Average scaled score in mathematics for pupils with high prior attainment	No	10,409	109.9	0.019	*
		Yes	865	110.1		
5	Average scaled score in mathematics for disadvantaged pupils	No	8984	101.9	0.929	
		Yes	692	101.9		
6	Average scaled score in mathematics for non-disadvantaged pupils	No	8984	105.1	0.022	*
		Yes	692	105.4		
7	Percentage of pupils reaching the expected standard in mathematics	No	13,686	76.4%	0.002	**
		Yes	1071	77.8%		
8	Percentage of pupils with low prior attainment reaching the expected standard in mathematics	No	2851	21.9%	0.217	
		Yes	220	20.2%		
9	Percentage of pupils with medium prior attainment reaching the expected standard in mathematics	No	12,988	73.9%	0.048	*
		Yes	1040	75.0%		
10	Percentage of pupils with high prior attainment reaching the expected standard in mathematics	No	10,420	98.0%	0.952	
		Yes	867	98.0%		
11	Percentage of disadvantaged pupils reaching the expected standard in mathematics	No	8995	65.1%	0.899	
		Yes	695	65.2%		
12	Percentage of non-disadvantaged pupils reaching the expected standard in mathematics	No	8995	80.6%	0.226	
		Yes	695	81.3%		
13	Percentage of pupils achieving a high score in mathematics	No	13,686	23.1%	0.000	**
		Yes	1071	24.8%		
14	Percentage of pupils with low prior attainment achieving a high score in mathematics	No	2851	1.1%	0.038	*
		Yes	220	0.7%		
15	Percentage of pupils with medium prior attainment achieving a high score in mathematics	No	12,988	10.5%	0.172	
		Yes	1040	11.0%		
16	Percentage of pupils with high prior attainment achieving a high score in mathematics	No	10,420	53.8%	0.017	*
		Yes	867	55.7%		
17	Percentage of disadvantaged pupils achieving a high score in mathematics	No	8995	13.8%	0.872	
		Yes	695	13.7%		
18	Percentage of non-disadvantaged pupils achieving a high score in mathematics	No	8995	26.5%	0.024	*
		Yes	695	27.7%		

Table 1. *Cont.*

Row	Performance Measure	Used +1 and PO2?	N	Mean	Significance [1]	
19	Percentage of pupils reaching the expected standard in mathematics teacher assessment	No	13,688	80.1%	0.000	**
		Yes	1071	81.5%		
20	Percentage of pupils absent from or not able to access the test in mathematics	No	13,686	0.38%	0.082	
		Yes	1071	0.32%		
21	Percentage of pupils absent or disapplied in mathematics teacher assessment	No	13,688	0.16%	0.042	*
		Yes	1071	0.12%		

[1] Key: * = 95% confidence limit. ** = 99% confidence limit.

Lines 7–12 show the percentage of children in various groups that have reached the government-defined expected level in mathematics at the age of 11, averaged across the sample schools which use +1 and Po2, and also the mean for the national comparator schools. Similarly lines 13–18 shows the percentages of children achieving a high score on the tests.

In addition to the tests, a classroom-based summative teacher assessment is carried out during the last term of primary education. Line 19 of Table 1 shows the percentage of children achieving the expected standard in the teacher assessment; in the national performance tables this is not broken down into performance of different groups of children. The last two lines of the table show the percentages of children absent from the tests and the teacher assessment, respectively.

4. Discussion

4.1. Scaled Scores

We can say with some certainty that the mean aggregated scaled score for mathematics is higher for those schools using +1 and Po2 than it is for all other schools (>99% confidence). The magnitude of the difference between the means is relatively small at less than one point (0.35), and a superficial analysis of effect size using Cohen's *d* [30] might lead to a conclusion that the effect is unimportant (in this case *d* = 0.12). However, we would do well to learn from the medical community that "effect size is not omnipotent" [31] and note that a simplistic interpretation of effect size could mask a profound educational impact on individuals—comparable with estimates of clinical impact in the medical world [32].

By way of illustration, consider two schools of 100 pupils, where the average scaled score for the first school is 104.00. If the average scaled score for the second school is 104.35, and for the sake of argument the difference is due to the performance of five children, those children must each score 7 additional points to generate this effect size. This would take them above the threshold for high performance and would represent important learning gains over and above their peers if it were generated by a specific intervention. School leaders might well consider such an intervention worth investment, even though, statistically, the effect size is small. Those schools that purchased +1 and Po2 did so in modest quantities: an average of nine and eleven books, respectively, totalled over 4 years. This compares with a mean cohort size of 40 pupils, so it seems reasonable to assume that they are using the resources with less than 10% of the cohort. We cannot make assumptions about cause and effect on the basis of this measure alone, but if the statistically significant advantage seen by the +1 and Po2 schools were driven either fully or partially by the intervention then it would appear that the one-to-one support may be having an important impact on a small number of pupils.

The other mean scaled scores in Table 1 give a breakdown of the pupil cohorts by low, medium and high prior attainment (as defined by UK government), and also by disadvantage (where pupils in receipt of the pupil premium funding is used as a proxy indicator for disadvantage). These scores show smaller differences, with only pupils with a high prior attainment and those who are not disadvantaged showing a statistically significant benefit in the +1 and Po2 schools (to 95% confidence limit). Note that the

number of schools included in the samples are lower in these measures because data for groups with small numbers of children is suppressed at source to prevent individuals being identified. Given that the +1 and Po2 resources are targeted at children who have fallen behind in their numeracy development it would be unlikely that they have been used to secure improved scores for those children with high prior attainment, and it seems reasonable to infer that this is due to an alternative stimulus. Conversely, +1 and Po2 schools might be disappointed that the performance of low prior achieving children has not been enhanced by the intervention, although suppression of data means that almost 4/5 of +1 and Po2 schools have been excluded from the sample for this measure.

4.2. Expected Standards

The basic headline performance measure for mathematics in English primary schools is the percentage of children achieving the expected standard. From line 7 of Table 1 we can see that this is 1.4 percentage points higher for +1 and Po2 schools than it is for all other schools (99% confidence limit). A similar difference is confirmed by the figures in line 19 which show the percentages of children achieving expected standard in the teacher assessment. This is hugely important for school leaders: a difference of 1% at this level places a school 400 places higher in the national league tables. While this might raise questions about the validity of the league tables, the pragmatic reality is that such tables are used by politicians and public to make judgements about schools, and again a small but statistically significant effect could be of critical importance to a school.

Lines 8–12 in Table 1, which break down the expected progress measure by pupil groups, suffer from the suppression of large quantities of data due to small numbers, such that it is not possible to determine exactly which groups of children contributed to the expected standards figure. However, the percentage for pupils with medium prior attainment is higher for +1 and Po2 schools (95% confidence), and this middle group is perhaps where the biggest difference might be expected. The +1 and Po2 intervention is designed as a catch-up programme for children whose progress has slowed in mathematics, and a successful intervention here should enable those children who are only a little behind to cross the expected progress threshold. No differences are seen for the children with low and high prior attainment, and there is no difference when stratified by disadvantage.

4.3. High Scores

Turning our attention to the percentage of children achieving high scores in mathematics, we can see from line 13 that +1 and Po2 schools outperform others by 1.7 percentage points (99% confidence). This is perhaps surprising: the +1 and Po2 programme would not be expected to raise the attainment above the high score threshold. It is more likely that those schools investing in +1 and Po2 are also using additional strategies to support more able children and are seeing the benefits of this in their results.

The breakdown by prior attainment and disadvantage (Table 1, lines 14–18) poses some intriguing questions. This shows that the +1 and Po2 schools enable a greater proportion of more able children to achieve a high score than others (95% confidence, line 16). It is unlikely that this is due to a catch-up intervention, and supports the suggestion above that these schools are not just investing in children who have fallen behind, but have also identified ways to stretch their more able pupils. Line 14 of Table 1 shows that +1 and Po2 schools see fewer low ability pupils achieving a high score than comparison schools. Though the sample size is smaller due to data suppression, the difference is nonetheless statistically significant (95% confidence). This might suggest that they are supporting basic numeracy, but using a method that puts an artificial ceiling on the higher levels of achievement. We note, however, that the numbers here are very small and the difference will amount to just one child in a cohort.

Considering the difference between disadvantaged and more affluent pupils, the results show that the percentage of non-disadvantaged pupils achieving high scores is higher in +1 and Po2 schools than in others (line 18, 95% confidence). Again, this is

unlikely to be related to +1 and Po2, and is more likely to be caused by a strategy for higher achieving children, albeit an approach that appears to have a disproportionate effect for more affluent families.

4.4. Absent and Disapplied

The last two lines of Table 1 give some information about those pupils who were absent from or unable to access the tests, and those who were absent for or disapplied from the National Curriculum teacher assessment. In this case the term "unable to access" means either that the pupils were not entered because their work indicated they would not have been able to complete the tests, or that they attempted the test but did not achieve enough raw marks to meet the threshold for scaling. While there is no evidence of a difference in test performance between +1 and Po2 schools and others, it can be seen that the percentage of pupils absent or disapplied in the teacher assessments is lower in the +1 and Po2 schools (line 21, 95% confidence). The term disapplication refers to withdrawal from the content of the National Curriculum for mathematics, and is usually applied to children with particular needs for whom the National Curriculum content is deemed by the school to be inappropriate or inaccessible.

This difference indicates that in schools using +1 and Po2 more children are able to be involved in the national teacher assessments. Anecdotally, teachers say that using +1 and Po2 gives children more confidence in their mathematics ability, and this may lie behind the measurable difference. As with previous data, although the number of children involved is small, the difference between being disapplied and excluded from the National Curriculum, or included in the National Curriculum teaching and assessment, is a profound one that can have long lasting implications throughout the period of secondary education.

5. Conclusions

This exploratory school-level analysis suggests that in this particular instance, the use of a one-to-one mathematics catch-up intervention has the potential to lift some pupils above the threshold for expected progress. The data for pupils with low prior attainment is inconclusive, and there may be a danger that schools have focused their efforts on pupils close to the performance threshold in order to ensure that the school-level headline figures are enhanced. Such "performance-maximising behaviour" [33] was certainly found to be a problem in the early years of secondary school performance tables in England, where dysfunctional effects were caused by schools concentrating on the targets defined by the neoliberal system at the expense of other important objectives [34]. At that time school leaders confessed to targeting resources towards borderline pupils who could potentially boost a pivotal indicator [35]. The introduction of progress measures into the accountability framework was described by headteachers as a "step forward" [36] and potentially mitigates against dysfunctional behaviour, although progress measures are not well understood by the general public—to the extent that the UK government had to publish explanatory notes for parents [37]—so attainment measures are always likely to carry more weight and may take priority for school leaders.

This study is limited by using only school-level data, and further study is required to examine at pupil level how effective the +1 and Po2 publications are in helping children to catch up with their peers and national expectations. Ideally this would be in the form of a large-scale study conducted as a randomised control trial (RCT) with the potential to control for selection bias and alternative effects, and provide strong counterfactuals to infer causation [38,39]. RCTs in education remain controversial among the academic community [40] but in practice they are commonly used and can provide considerable insight into the benefit of a particular intervention despite being somewhat intrusive in nature [41]. In addition to a quantitative study it would also be beneficial to carry out a qualitative investigation into how the products are actually used in schools.

The intention of the publishers of +1 and Po2 is that the workbooks are used by children, supported on a one-to-one basis by adults other than teachers, and it is this

feature which might make the resource attractive to schools which cannot afford to free up teachers for one-to-one tuition. In principle there could be zero personnel costs for a school if it were to use volunteers or parents (or indeed, peers) as the coaches for +1 and Po2, making it extremely cost effective if it produces learning gains. We cannot tell how the schools in this study used the workbooks, but it would be fair to assume that they did not pay qualified teachers for one-to-one tuition as their resources would not stretch to this. Whatever the qualifications and expertise of the coaches, it would indeed be surprising if a one-to-one coaching model did not bear some fruit, whichever resource or programme is employed. Having demonstrated the potential for learning gains in this small study, a future study would benefit not just by comparing the progress of children using +1 and Po2 with all other children, but by comparing with those using another one-to-one mathematics support package. This would help to identify the contribution made specifically by +1 and Po2, compared with improvements due to one-to-one tuition generally.

The suggestion from this analysis that +1 and Po2 have made the prescribed National Curriculum teacher assessments in mathematics more accessible for pupils is important. Teachers may feel that if an intervention builds confidence it will be worth investment, even if it does not generate learning gains in and of itself. It has been shown that children's attainment in the primary classroom is influenced by their self-confidence, and in a virtuous circle improved attainment builds confidence further [42]. Test anxiety is a learned phenomenon experienced by children as young as 7 [43] and may relate to both tests and teacher assessments, especially if children perceive that there are high personal stakes, as they frequently do with National Curriculum assessments [44]. This is not a problem confined to the UK but applies wherever there are regular testing regimes—one American study describes test anxiety in school children as a national epidemic [45]. While a lack of confidence in mathematics is not the same as general test anxiety, an intervention that builds confidence in children such that the least able have the opportunity to demonstrate success in assessments must be of some benefit. This study focuses on the potential to accelerate learning for 11-year-olds in primary schools, though the +1 and Po2 resources are marketed for any age group, including adults. With some researchers reporting very high proportions of mathematics anxiety among adults in various developed countries [46] the use of a simple one-to-one coaching model to build confidence should be seen as the potential to transform lives and contribute to economic prosperity for adults as well as children: Gross et al. estimate a 19:1 return on investment in a similar programme [47].

This preliminary study would appear to be particularly prescient, as it used data from immediately before the coronavirus pandemic which slowed children's progress in mathematics and other subjects worldwide. Around the globe school leaders will be seeking cost effective catch-up programmes which can help to get their pupils back on track. The principle that one-to-one tuition can help children catch up is one that is accepted by policymakers, who may direct funding towards it, and the notion of using adults other than teachers as coaches will undoubtedly resonate with school leaders. Further research into one-to-one tuition for accelerating learning in mathematics in particular is essential, and qualitative studies might help to understand who the users of this resource are, to what extent they might be considered coaches, and the nature of the coaching in practice.

Institutional Review Board Statement: This study was conducted according to the principles of the British Educational Research Association's ethical guidelines for educational research [29] and was approved by the ethics committee of Nottingham Trent University (application number 2020/48, 25 February 2020).

Informed Consent Statement: Not applicable: data is at school level and is in the public domain.

Data Availability Statement: School-level data supporting this work can be found on the UK government website: https://www.compare-school-performance.service.gov.uk/download-data (accessed on 12 November 2021).

References

1. Ahmed, M. A Case Study on the Development of Adult Language, Literacy and Numeracy Skills. *EAI Endorsed Trans. E-Learn.* **2020**, *6*, 1–6. [CrossRef]
2. Windisch, H.C. *Adults with Low Literacy and Numeracy Skills: A Literature Review on Policy Intervention*; OECD Education Working Papers No. 123; OECD Publishing: Paris, France, 2015. [CrossRef]
3. OECD. *OECD Skills Outlook 2013: First Results from the Survey of Adult Skills*; OECD Publishing: Paris, France, 2013. [CrossRef]
4. OECD. *Skills Matter: Additional Results from the Survey of Adult Skills*; OECD Skills Studies; OECD Publishing: Paris, France, 2019. [CrossRef]
5. Evans, J.; Yasukawa, K.; Mallows, D.; Creese, B. Numeracy skills and the numerate environment: Affordances and demands. *Adults Learn. Math. Int. J.* **2017**, *12*, 17–26.
6. Paccagnella, M. *Literacy and Numeracy Proficiency in IALS, ALL and PIAAC*; OECD Working Paper No. 142; OECD Publishing: Paris, France, 2016.
7. Carpentieri, J.D. The evolution of evidence problems in adult literacy and numeracy: How did we get into the current impasse and how can we get out of it? *Int. J. Lifelong Educ.* **2019**, *38*, 644–656. [CrossRef]
8. Geary, D.C. Consequences, characteristics, and causes of mathematical learning disabilities and persistent low achievement in mathematics. *J. Dev. Behav. Pediatrics* **2011**, *32*, 250–263. [CrossRef] [PubMed]
9. Litster, J. *The Impact of Poor Numeracy Skills on Adults. Research Review*; NRDC and IoE: London, UK, 2013.
10. Martin, R.; Hodgson, H.; Maloney, A.; Rayner, I. *Cost of Outcomes Associated with Low Levels of Adult Numeracy in the UK*; Pro Bono Economics & National Numeracy: London, UK, 2014.
11. Vorhaus, J.; Litster, J.; Frearson, M.; Johnson, S. *Review of Research and Evaluation on Improving Adult Literacy and Numeracy Skills*; BIS Research Paper No. 61; BIS: London, UK, 2011.
12. Further Education Trust for Leadership. *Skills for Life: A New Strategy for English, Maths, ESOL and Digital*; FETL: London, UK, 2021.
13. Walberg, H.J.; Tsai, S.L. Matthew effects in education. *Am. Educ. Res. J.* **1983**, *20*, 359–373.
14. Field, A.P.; Evans, D.; Bloniewski, T.; Kovas, Y. Predicting mats anxiety from mathematical achievement across the transition from primary to secondary education. *R. Soc. Open Sci.* **2019**, *6*, 191459. [CrossRef]
15. Brown, M.; Bibby, T.; Johnson, D.C. Turning our attention from the what to the how: The National Numeracy Strategy. *Br. Educ. Res. J.* **2000**, *26*, 457–471. [CrossRef]
16. Gross, J. Supporting children with gaps in their mathematical understanding: The impact of the National Numeracy Strategy (NNS) on children who find mathematics difficult. *Educ. Child Psychol.* **2007**, *24*, 146–156.
17. Anders, Y.; Rossbach, H.; Weinert, S.; Ebert, S.; Kuger, S.; Lehrl, S.; von Maurice, J. Home and preschool learning environments and their relations to the development of early numeracy skills. *Early Child. Q.* **2012**, *27*, 231–244. [CrossRef]
18. Denton, K.; West, J. Children's reading and mathematical achievement in kindergarten and first grade. *Educ. Stat. Q.* **2002**, *4*, 19–26.
19. Copeland, J. A critical reflection on the reasoning behind, and effectiveness of, the application of the Pupil Premium Grant within primary schools. *Manag. Educ.* **2019**, *33*, 70–76. [CrossRef]
20. Ladd, H.; Fiske, E. Weighted student funding in the Netherlands: A model for the U.S.? *J. Policy Anal. Manag.* **2011**, *30*, 470–498. [CrossRef]
21. Clark, J.; Roy Das, S.; Menclova, A. Evaluating the effectiveness of school funding and targeting different measures of student disadvantage: Evidence from New Zealand. *Econ. Rec.* **2017**, *93*, 576–599. [CrossRef]
22. Morris, R.; Dobson, G. Spending the pupil premium: What influences leaders' decision-making? *Educ. Manag. Adm. Leadersh.* **2021**, *49*, 284–302. [CrossRef]
23. EEF. One to One Tuition. Teaching & Learning Toolkit 2018. Available online: https://educationendowmentfoundation.org.uk/evidence-summaries/teaching-learning-toolkit/one-to-one-tuition/ (accessed on 30 August 2021).
24. DfE. *Literacy and Numeracy Catch-Up Strategies*; OGL: London, UK, 2018.
25. Leckie, G.; Goldstein, H. The evolution of school league tables in England 1992–2016: 'Contextual value-added', 'expected progress' and 'progress 8'. *Br. Educ. Res. J.* **2017**, *43*, 193–212. [CrossRef]
26. Sharp, D.J. *Plus 1: The Introductory Coaching System for Maths Success*; 123Learning: Nottingham, UK, 2005.
27. Sharp, D.J. *Power of 2: The Introductory Coaching System for Maths Success*; 123Learning: Nottingham, UK, 2005.
28. Perry, T. English Value-Added Measures: Examining the Limitations of School Performance Measurement. *Br. Educ. Res. J.* **2016**, *42*, 1056–1080. [CrossRef]
29. BERA. *Ethical Guidelines for Educational Research*, 4th ed.; BERA: London, UK, 2018.

30. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed.; Lawrence Erlbaum Associates: New York, NY, USA, 1988.
31. Lee, D.K. Alternatives to P value: Confidence interval and effect size. *Korean J. Anesthesiol.* **2016**, *69*, 555–562. [CrossRef]
32. McGough, J.J.; Faraone, S.V. Estimating the size of treatment effects: Moving beyond p values. *Psychiatry* **2009**, *6*, 21–29. [PubMed]
33. West, A. High stakes testing, accountability, incentives and consequences in English schools. *Policy Politics* **2010**, *38*, 23–39.
34. Wiggins, A.; Tymms, P. Dysfunctional Effects of League Tables: A Comparison between English and Scottish Primary Schools. *Public Money Manag.* **2002**, *22*, 43–48. [CrossRef]
35. Wilson, D.; Croxson, B.; Atkinson, A. "What gets measured gets done" Headteachers' responses to the English Secondary School Performance Management System. *Policy Stud.* **2006**, *27*, 153–171. [CrossRef]
36. Tymms, P.; Dean, D.C. *Value-Added in the Primary School League Tables: A Report for the National Association of Head Teachers*; CEM Centre, University of Durham: Durham, UK, 2004.
37. DfE. *Information for Schools and Parents: Primary School Progress Measures*; OGL: London, UK, 2017.
38. Banerjee, P.A. A systematic review of factors linked to poor academic performance of disadvantaged students in science and maths in schools. *Cogent Educ.* **2016**, *3*, 1178441. [CrossRef]
39. Higgins, S. Room in the Toolbox? The place of randomised controlled trials in educational research. In *Mobilising Teacher Researchers: Challenging Educational Inequality*; Childs, A., Menter, I., Eds.; Routledge: London, UK, 2017.
40. Connolly, P.; Biggart, A.; Miller, S.; O'Hare, L.; Thurston, A. *Using Randomised Controlled Trials in Education*; Sage: London, UK, 2017.
41. Styles, B.; Torgerson, C. Randomised controlled trials (RCTs) in education research–methodological debates, questions, challenges. *Educ. Res.* **2018**, *60*, 255–264. [CrossRef]
42. Nunes, T.; Bryant, P.; Sylva, K.; Barros, R. *Development of Maths Capabilities and Confidence in Primary School*; DCSF: London, UK, 2009.
43. Whyte, J.; Anthony, G. Maths anxiety: The fear factor in the mathematics classroom. *N. Z. J. Teach. Work* **2012**, *9*, 6–15.
44. Howard, E. *A Review of the Literature Concerning Anxiety for Educational Assessments*; Ofqual: Coventry, UK, 2020.
45. Lamb, E.; Scott, D.A.; Bentley, J.K.; Sumner, C. Test anxiety in school-age children: An examination of a national epidemic. *Glob. J. Guid. Couns. Sch.* **2018**, *8*, 8–16.
46. Blazer, C. Strategies for Reducing Math Anxiety. *Inf. Capsul.* **2011**, *1102*, 1–8.
47. Gross, J.; Hudson, C.; Price, D. *The Long Term Costs of Numeracy Difficulties*; Every Child a Chance Trust: London, UK, 2009.