Cambridge Assessment

The link between subject choices and achievement at GCSE and performance in PISA 2015: Executive summary

Research Report

Matthew Carroll & Tom Benton
December 2018

Author contact details:
Matthew Carroll & Tom Benton
Assessment Research and Development,
Research Division
Cambridge Assessment
The Triangle Building
Shaftesbury Road
Cambridge
CB2 8EA
UK

Carroll.m@cambridgeassessment.org.uk
Benton.t@cambridgeassessment.org.uk

http://www.cambridgeassessment.org.uk

As a department of Cambridge University, Cambridge Assessment is respected and trusted worldwide, managing three world-class examination boards, and maintaining the highest standards in educational assessment and learning. We are a not-for-profit organisation.

How to cite this publication:

Carroll, M. & Benton, T. (2018). *The link between subject choices and achievement at GCSE and performance in PISA 2015: Executive summary*. Cambridge Assessment Research Report. Cambridge, UK: Cambridge Assessment.

The purpose of this research was to compare individual pupil achievement on two high profile sets of assessments in England: the Programme for International Student Assessment (PISA), and the General Certificate of Secondary Education (GCSE). It should be noted to begin with that these two assessments are explicitly designed to serve different purposes. GCSEs are intended to test individual pupils' knowledge of and ability to apply specific content as defined in the qualification specification, with results having an impact upon their future educational and job opportunities. In contrast, the focus of PISA is upon the performances of whole countries rather than individuals, with the assessments designed to measure the extent to which pupils can apply particular areas of knowledge to solve problems within real world situations.

However, whilst acknowledging the different aims of the two assessments, understanding the links between performances on them is important as it affects the way we interpret findings from the PISA studies. Specifically, knowing how the PISA tests relate to what is actually taught and assessed as part of the GCSE curriculum may inform the extent to which we believe they provide a reasonable evaluation of our school system, as well as the extent to which it is justifiable to make decisions about the administration of GCSEs based upon PISA results. In addition, given that the PISA tests are sometimes criticised for failing to measure key skills such as creativity (Villalba, 2012), analysis of the relationship between PISA and individual GCSE subjects provides an empirical basis for us to form opinions about precisely what is being measured.

The data set used in this analysis was supplied by the Department for Education (DfE). It contained all available data from the PISA 2015 study for participants in England matched to key variables from the National Pupil Database (NPD) - in particular, students' GCSE grades in a number of individual GCSE subjects. This data set was anonymised before being provided to us so that no individual pupil could be personally identified. The full data set contained information on 5,194 pupils from a total of 206 schools. This data set was used to carry out an analysis of the strength of the associations between PISA, Key Stage 2 (KS2) and GCSE achievements. The PISA ability estimates (plausible values) within the data set were recalculated. In particular, the recalculation was done to ensure that the relatively short lengths of the PISA assessments would not in themselves result in estimated correlations being artificially low (see main report for technical details).

To begin with, the analysis explored the degree to which grades in individual GCSE subjects were correlated with PISA ability estimates. A sample of the results, restricted to GCSE subjects taken by at least 1,000 pupils within the matched data, is shown in Table 1. Although the results indicate moderately strong, positive correlations between GCSE grades and PISA abilities, these correlations are not particularly high when benchmarked against correlations between GCSE grades and KS2 scores. For example, the correlation between PISA maths ability and GCSE maths grade (0.78) is only slightly higher than the correlation between KS2 maths test marks and GCSE maths grade (0.75[1]). This is true despite the fact that KS2 tests and GCSEs are taken five years apart, as opposed to the gap of roughly 6 months between the PISA assessments and students taking their GCSEs. Data collated

[1] This, and more, correlation coefficients relating GCSE grades and key stage 2 performances are given in the main report.

from GCSE resits further confirmed that the relatively low correlation between PISA and GCSE cannot be explained simply by the time interval between the two tests. Rather, we interpret the results to indicate that the PISA tests measure something rather different to GCSEs. This fact need not solely relate to the content of the PISA tests but may also relate to other differences, such as the fact that PISA tests are low-stakes for the pupil (thus they may not apply full effort, and will not have prepared as thoroughly), or the fact that the PISA tests are fully computer-based. Although differences in content between GCSEs and PISA have been noted before (see, for example, Jerrim & Shure, 2016), this research confirms the impact of these differences empirically. As such, the skills[2] that are explicitly measured by PISA cannot be assumed to act as a proxy for all the skills that are not.

Table 1. Pearson correlations of PISA domain abilities with GCSE grades and key stage 2 scores. For further details on calculations and for results from the full range of subjects examined, see the main report. GCSE subjects are ordered by their correlation with PISA science from highest to lowest. The colouration from red to blue is used to highlight the highest and lowest correlations in the table.

Variable	N	Science	Maths	Reading	Collaborative Problem Solving
GCSE Core Science	3,037	0.748	0.714	0.692	0.550
GCSE Additional Science	2,779	0.734	0.706	0.660	0.525
GCSE Maths	4,778	0.728	0.777	0.672	0.593
GCSE Geography	2,232	0.714	0.698	0.687	0.583
GCSE Physics	1,563	0.699	0.732	0.604	0.492
GCSE History	2,373	0.696	0.675	0.696	0.561
GCSE Biological Science	1,580	0.681	0.696	0.624	0.494
GCSE Chemistry	1,566	0.659	0.699	0.607	0.479
GCSE English	4,735	0.628	0.625	0.680	0.534
GCSE English Literature	4,287	0.613	0.592	0.637	0.534
GCSE Religious Studies	2,447	0.575	0.564	0.595	0.481
GCSE Physical Education	1,102	0.571	0.579	0.538	0.454
GCSE French	1,387	0.535	0.551	0.538	0.376
GCSE Information Technology	1,167	0.524	0.541	0.532	0.448
GCSE Art and Design	1,334	0.500	0.475	0.507	0.328
KS2 English: marks in reading test	4,564	0.648	0.601	0.638	0.572
KS2 Maths: total test marks	4,575	0.645	0.740	0.553	0.512
KS2 English: marks in writing test	4,564	0.519	0.524	0.549	0.441

Although the results comparing GCSE and PISA maths are interesting, those comparing GCSE English and PISA reading are even more striking. Naively, one might assume that these two tests measure similar skills. However, our correlation analysis shows that performance on the PISA reading test is at least as closely aligned with achievement in

[2] The term "skills" is used throughout this report as shorthand for "the propensity for students to correctly answer questions of the type asked in PISA/GCSE". It is not intended to imply a dichotomy between "skills" and "knowledge and understanding".

GCSE science as it is with GCSE English. This, initially surprising, finding becomes considerably easier to understand once a few of the items used to assess PISA reading have been explored. For example, if we look through a sample of released PISA items[3] we find that the several of the reading tasks actually ask students to read and interpret tables or figures, a skill that is assessed in GCSE science tests rather than in GCSE English. The substantive importance of this finding comes in that GCSE English forms a fundamental part of the way that school performance is judged in England, meaning that schools devote substantial time to teaching skills such as essay writing that are fundamental to success in this subject. However, it is clear that when it comes to judging the performance of our education system as a whole using PISA, this particular skill is not measured. Thus the performance of our country's education system is judged whilst ignoring some of the key skills[4] that schools are trying to teach.

PISA collaborative problem solving displayed by far the lowest correlations with achievement in GCSEs. This is interesting for two reasons. Firstly, it is the domain where England displays its strongest performance relative to the OECD average (Jerrim & Shure, 2017). As such, we may ask how this is being achieved given that the skills being assessed show little similarity with those tested (and, thereby, presumably taught) within any mainstream GCSE subject. One possibility is that these skills are not widely taught in other countries either, so that the performances of countries may be more affected by activities out of school than is the case for the other PISA domains.

Secondly, the relatively low correlation between GCSE performance and PISA collaborative problem solving raises the question of whether all of our assessment at GCSE fails to recognise certain pupils' skills. For example, our analysis suggests that roughly a tenth of students averaging at grade C at GCSE (i.e. of middling ability in GCSEs, see Gill, 2017) are amongst the top fifth of students when it comes to collaborative problem solving (see main report). Even amongst pupils with an average GCSE grade of D, one in twenty will be amongst the top fifth of performers in collaborative problem solving. Whether we see this as a genuine deficiency in the current GCSE system will depend upon our view of the importance of this skill in its own right and the validity of the OECD's approach to assessing it. For a further discussion of merits of the OECD's collaborative problem solving assessment see Shaw and Child (2017).

The findings presented above are robust to different methods of calculating correlations, and to restricting analysis to pupils that appear to have made a reasonable level of effort in the tests. For further details on these checks see Appendix 1 of the main report.

This research also explored the association between GCSE subject choices and pupil abilities as measured by PISA. The strongest of such relationships was found with whether students studied three separate science GCSEs or whether they studied the combined science GCSEs (core and additional science). Before accounting for any pupil

[3] Available from https://www.oecd.org/pisa/pisaproducts/Take%20the%20test%20e%20book.pdf. The "Lake Chad" item gives an example of a PISA reading item that requires looking at scientific information.

[4] The term "skills" is used throughout this report as shorthand for "the propensity for students to correctly answer questions of the type asked in PISA/GCSE". It is not intended to imply a dichotomy between "skills" and "knowledge and understanding".

characteristics, the mean PISA score of those studying separate sciences was around 70 PISA points ahead of those studying combined science. After accounting for differences in prior attainment at key stage 2, the difference was around 35 points - equivalent to just under a year's progress using a rule of thumb suggested by OECD (see, for example, OECD, 2010, page 55). Even after accounting for differences in achievement at GCSE, in every domain, those studying separate sciences were around 20 PISA points ahead of those studying combined science – a difference equivalent to roughly 6 months progress. If we interpreted these differences in a purely causal manner they would suggest that, without even needing to improve GCSE results, if the two-thirds of pupils studying combined science moved to studying separate sciences, the performance of the UK in PISA could rise to being close to that of Korea (ranked between 4th and 9th across different domains). However, whilst these differences may possibly indicate that studying separate sciences helps students to develop the skills measured by PISA, it is also possible they may be explained by unmeasured aptitudes of the pupils choosing to study separate sciences in the first place.

The differences between those studying separate and combined sciences are illustrated for the case of PISA maths in Figure 1. The figure shows how the distribution of PISA abilities in maths varied according to pupils' grades in GCSE maths and whether they studied combined or separate sciences at GCSE. As can be seen, for any given grade in GCSE maths, pupils that studied separate science GCSEs tended to have higher scores in PISA maths than those who studied combined sciences. There is some indication that this difference was more prominent amongst those with lower grades at GCSE. Specifically the differences in medians between those taking separate and combined sciences was around 40 PISA points amongst those with grade D, around 25 points amongst those with grade A, and 10 points amongst those with grade A*. This may suggest that the separate science effect is partially caused by the impact of students being entered and prepared for different tier exams as the difference is smaller amongst students with grades B and above who were definitely entered for higher tier GCSE papers than amongst lower grades[5]. Having said this, it is clear that some gap persists across all grades (and hence both tiers). Also, given the restricted sample sizes amongst those with different grades, these comments should be treated with some caution. Similar findings could be shown across all other PISA domains (science, reading and collaborative problem solving). More detailed analyses using mixed effects modelling confirmed that these differences were statistically significant (see the main report for details).

[5] For maths GCSE exams in England students could either enter higher tier examinations where grades A*-E were available or foundation tier exams where grades C-G were available.

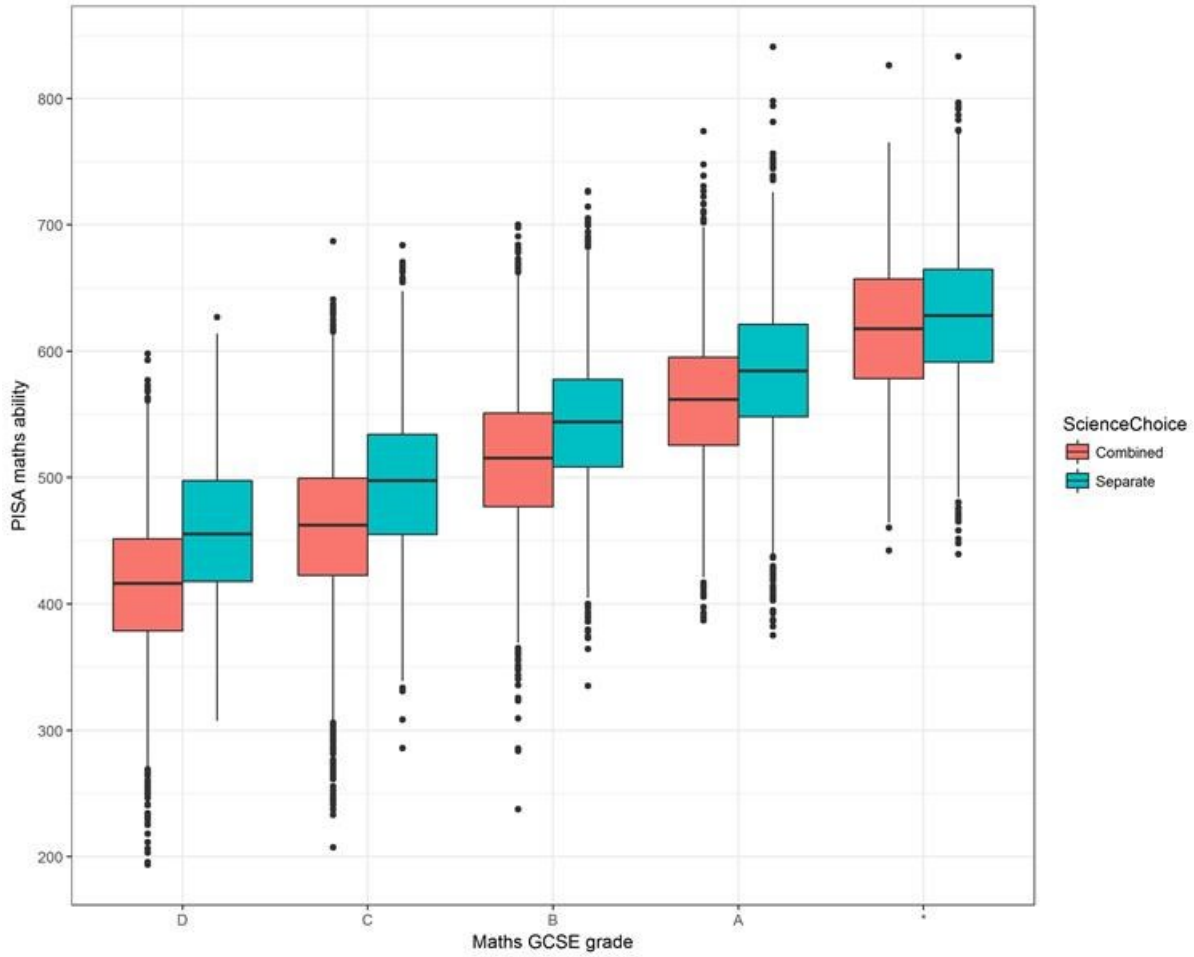

Figure 1: The relationship between Maths GCSE grade and PISA plausible values, split by whether students studied separate or combined sciences at GCSE.

The analysis also identified a small number of other statistically significant associations between GCSE subject choices and abilities within particular PISA domains. Within PISA reading, students who had chosen to study history, German or French displayed higher abilities (by around 15 points) than those, with equally good overall GCSE performance, who had not studied these subjects. Although alternative explanations are possible, this may indicate that performance in PISA reading is associated with the extent to which students are taking GCSE subjects that develop relevant skills (e.g. comprehension of texts). Similarly, after accounting for GCSE performance, the results suggested a positive association between taking GCSE drama and performance in collaborative problem solving (just under 20 points) - an understandable finding given that the collaborative problem solving tasks are themselves a form of role play. Conversely, a negative association was found between taking physical education GCSE and performance in PISA science (just under 10 points difference). This may potentially indicate that PISA favours students taking subjects that are fully classroom based. Having said this, there was no indication that pupils taking arts subjects such as Art & Design or music performed worse in PISA than those choosing less artistic options.

It is worth noting that even if all the results relating pupil subject choices to PISA scores were genuine causal effects (and we provide no guarantee that they are), this does not

necessarily imply those subjects associated with higher PISA scores are necessarily "better". For example, would it necessarily be right to force all pupils to study separate sciences up until the age of 16, regardless of their interest in the subject, *purely* to boost results in PISA? Likewise, would it be correct to discourage the uptake of the GCSE in physical education, regardless of pupils' interest or aptitudes, just because taking this subject is associated with lower scores in PISA science? Rather than leading to such conclusions, we hope that our work will encourage reflection on what exactly is (and isn't) measured by PISA and that this will help inform the way results from these studies are interpreted.

This research highlights the fact that PISA and GCSEs measure different skills, and has also shown, perhaps as might be expected, that an individual's performance in PISA is likely to depend upon their choices with regard to what they study. As such, the popularity of subjects at a national level could have some influence on England's performance in PISA. These findings warn against uncritical use of the PISA results without careful consideration of exactly what is being measured. They also give the opportunity to reflect upon the skills that are currently assessed both by PISA and by our own GCSEs and the extent to which these match wider societal goals.

References

Gill, T. (2017). *Candidates awarded A* and A grades at GCSE in 2015*. Statistics Report Series No. 111. Cambridge, UK: Cambridge Assessment. Available from http://www.cambridgeassessment.org.uk/Images/352252-candidates-awarded-a-and-a-grades-at-gcse-in-2015.pdf.

Jerrim, J., & Shure, N. (2016). *Achievement of 15-Year-Olds in England: PISA 2015 National Report*. London: Department for Education DFE-RR630. Available from https://www.gov.uk/government/publications/pisa-2015-national-report-for-england.

OECD (2010) *PISA 2009 Results: What students know and can do – student performance in reading, mathematics and science (Volume I). OECD: Paris.* Available online: http://dx.doi.org/10.1787/9789264091450-en.

Shaw, S. and Child, S. (2017). Utilising technology in the assessment of collaboration: A critique of PISA's collaborative problem-solving tasks. *Research Matters: A Cambridge Assessment publication*, 24, 17-22. Available from http://www.cambridgeassessment.org.uk/Images/426176-research-matters-24-autumn-2017.pdf.

Villalba, E. (2012). Searching for the holy grail of measuring creativity. *Creativity Research Journal*, *24*(1), 1-2.