Assessing the effect of reducing frequency of ‘data drops’: a retrospective quantitative analysis

GORDANO SCHOOL

George Reakes

PURPOSE OF RESEARCH

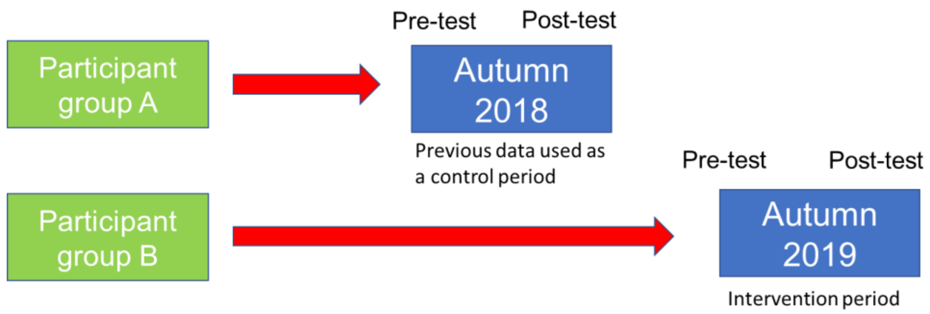
‘Data Drops’ (centralised data collection recording the attainment of whole cohorts) have taken place three times a year for each year group. In surveys, staff have reported that this creates a workload burden. We wanted to evaluate the impact on student progress if we reduce the ‘data drops’ to twice a year.

The supposed benefit of ‘data drops’, beyond information for parents, is that the data can be used by pastoral, curriculum and senior leaders to intervene, coach and track the progress of underachieving students to have an impact on their attainment. We wanted to reduce data drops to improve staff workload, but the risk was that student progress might be adversely affected. We used a measure of students’ progress in the core subjects to gather evidence about the impact of the data drop reduction.

THE RESEARCH DESIGN

This is a retrospective quantitative analysis comparing core subject estimated GCSE scores for this year’s Year 10 pupils, with last year’s Year 10. Last year’s Year 10 pupils had teachers who had not been exposed to the reduced workload intervention. As described in the Results section, Year 9 data for both participant groups was used as a covariate in order to improve the comparison.

Figure 1: Research design



LIMITATIONS

The dependent variable in this study is teacher-reported ‘projected grades’. These are likely to be based on a range of assessments carried out by teachers which may vary significantly between subjects and certainly between pre-tests and post-tests (for example the post-test data took into account data from a significant GCSE past-paper test at the end of Year 9, whereas the pre-test did not). However within subjects we expect that teachers would apply a consistent method for assessing attainment and recording projected grades, and so it is reasonable to compare the attainment of the Autumn 2018 cohort to the Autumn 2019 cohort.

METHODS

Participants and sample size

All Year 10 pupils in a large secondary school participated in this study, in the academic years 2018/19 and 2019/20. Data from the three ‘core subjects’ was analysed: 591 English Literature grades, 593 Mathematics grades, 171 Biology grades, 171 Chemistry grades, 399 Combined Science grades, and 172 Physics grades across the two years.

Procedures

Data input in 2018/19

All teachers record ‘projected grades’, ‘effort’ and ‘behavior’ as numerical scores in a centrally-stored spreadsheet for all pupils they teach three times per year. For projected grades these are on the GCSE 1-9 scale, and for effort and behavior these are on a scale of 1-4. For this study, the data for projected grades was gathered and analysed.

Data input in 2019/20

Exactly as in 2018/19 except that the number of times teachers record ‘projected grade’, ‘effort’ and ‘behavior’ is twice per year instead of three times.

Materials (and apparatus)

No additional materials were required. Teachers continued to use the same method to record and input data, but in the intervention period they did so less frequently (twice per year instead of three times per year).

CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE RESEARCH

Overall, any effect size differences that were detected were too small to be of concern (bearing in mind the nature of the analysis and its limitations) suggesting that no harm was being done by changing teacher practice. This said, a small non-significant effect size difference in Physics was found. Discussions with the Physics teachers suggested this may have been the results of naturally occurring differences in pupil ability between last year’s Year 10 and the current cohort.

Future research should consider moving to a more tightly controlled study in which a parallel comparable control is used. In addition, looking at school data in this detail raises questions about the use of 1-9 GCSE estimates without a clear and fully understood rubric.

RESULTS

Originally we planned to use ‘gain scores’ in the analysis. However, because of doubt in the direct comparability of pre-test to post- test data, an alternative method of analysis was selected. Although children are assessed on the same nine-point scale, the comparability of this nine-point scale to GCSE results may only be reliable at the post-test stage, where the scores were based on pupil performance in GCSE past-paper tests.

Preliminary assumption testing showed that the data did not meet the assumptions necessary to use parametric ANCOVA, so we used a non-parametric alternative (Quade’s F). Using pre-test scores as the covariate, separate analyses were conducted on the core subject data (Table 1). In addition, and for completeness, the effect was assessed across all subject areas in combination.

There was a significant positive, but small, effect size in English Literature. There was a significant negative, but negligible effect size in Combined Science. In Mathematics, Biology, Chemistry and Biology, there were no significant effects.

Table 1

Subject	Effect size r	CI (95%)	p-value	d	n
English Literature	0.051	0.02-0.08	<0.001	0.101	591
Mathematics	0.015	-0.01-0.04	0.156	0.030	593
Biology	-0.086	-0.026-0.08	0.326	-0.174	171
Chemistry	-0.081	-0.20-0.04	0.184	-0.163	171
Combined Science	-0.034	-0.06-0.01	0.014	-0.069	399
Physics	-0.148	-0.36-0.06	0.162	-0.299	172
All core subjects	0.015	0.01-0.02	<0.001	0.030	2097

This research was carried out with funding and support from the Department for Education and Education Development Trust.

Education Development Trust