sparx

A digitally-delivered, double-blind randomised controlled trial investigating the effect of homework length on year 7 students' attainment in mathematics

Dr Sidra Nawaz

Dr Stephen Welbourne

Abstract

Background

The relationship between homework length and attainment is a hotly debated topic in educational research (Cooper, Robinson, & Patall, 2006) but one that, to our knowledge, has thus far not been investigated in a double-blind, randomised controlled trial with randomisation at student level and with precisely controlled conditions in each trial arm.

Method

A four-arm, double-blind, randomised controlled trial was carried out involving a total of 368 Year 7 pupils from three secondary schools in Devon, which used a digital platform (Sparx) to deliver their maths homework. The trial took place over four weeks, during which the pupils were taught algebra as part of their schools' scheme of learning. In each class, pupils were matched by ability (KS2 maths scaled score) into groups of four and, within each group, randomly assigned to one of the four trial arms. In each trial arm, students received a set percentage by time of their normal algebra homework (0%, 50%, 100% and 150%). Students were assessed at the beginning and end of the half-term during which the trial took place. Attainment at each stage was measured using a one-parameter item response theory (IRT) analysis of the assessments.

Results

Complete test results at the pre- and post-test stage were successfully obtained for 354 out of 368 students (96%). A multivariate linear mixed model analysis was used to analyse the data, revealing that homework length was very significantly associated with the change in student attainment between the two test points after controlling for fixed effects of prior attainment and time and random effects of school, class and student ($p < 0.0001$).

Conclusions

This study provides very strong evidence that homework does, in some circumstances, significantly improve attainment. This is the first time that this has been tested using a double-blind RCT and illustrates the potential of using digital platforms to resolve long-standing educational controversies.

Introduction

The effectiveness of homework is a hotly debated topic and has been since the latter half of the last century (Cooper, 1989; Cooper et al., 2006; Trautwein, Köller, Schmitz, & Baumert, 2002; Vatterott, 2009), with questions regarding optimum homework length and frequency dominating the discourse (Cooper et al., 2006; Lam, 1996). In particular, the traditional view that more homework always equates to more learning has been increasingly challenged in recent years (Eren & Henderson, 2011; Rønning, 2011; Vatterott, 2009). A comprehensive and highly-cited review on the effectiveness of homework was published in 2006 (Cooper et al., 2006). This included 15 studies that had investigated the effect of doing homework on academic attainment and the relationship between time spent on homework and academic attainment in primary and secondary school pupils. While the results from these studies indicated a largely positive association between homework, or time spent on homework, and student attainment, the author highlighted a lack of rigorous randomisation and control of possible confounding variables in the study design, concluding that further research was necessary to establish the association between homework and attainment more definitively. More recently, large-scale research based on data from the international PISA tests also yielded ambiguous results (Dettmers, Trautwein, & Lüdtke, 2009). Simply put: it does seem to be the case that students receiving regular homework do better than those who don't, but it is not clear if this is due to the effect of the homework itself, or because schools that set regular homework tend to be more efficient educators.

Isolating the effect of time spent on homework on student achievement is difficult given the vast number of confounding variables at play. These include the student's prior ability, the subject on which homework is to be given, the student's learning experience prior to attempting homework, teacher skill, school facilities, the behaviour of other students in class, and the student's home environment. Further, it is difficult to control the exact time each student spends on their homework as, in most circumstances, students will be given a fixed number of questions and will take varying times to complete them. This could lead to counter-intuitive results where the faster and more able students appear to spend less time on homework, but achieve more (Trautwein, 2007).

Well-designed randomised controlled trials (RCTs) offer the ability to control for many of the confounding variables (Connolly, Keenan, & Urbanska, 2018; Torgerson, 2001). The optimum RCT design involves creating matched groups of individual students and randomly assigning students to the trial arms within these groups, such that each group is represented in each trial arm, allowing the trial arms to be as similar as possible. Hitherto, it has been practically impossible to randomly assign students within a class to different trial arms while preserving the double-blindness that is essential to avoid results being biased or contaminated. However, the use of digital platforms to deliver personalised homework has created an opportunity for a completely new type of study where students can be randomly assigned to trial arms without the teacher, the student, or the researcher being aware of the assigned arm for each student. Furthermore, sophisticated online learning platforms, such as the one used in this study, have the ability to predict how long each student will likely spend on each question, thus allowing homework assignments to be precisely calibrated and minimising the impact of the confounding variables described above.

In this study, we used the Sparx Maths Homework platform to conduct a four-arm, double-blind, randomised controlled trial to investigate the effect of algebra homework length on attainment. The Sparx platform enables precise manipulation of questions included in students' homework, allowing the proportion asked on a particular topic to be adjusted without increasing or reducing the overall length of homework. During the trial, the students were introduced to algebra for the first time in class. We tested the students immediately before and after the trial period and analysed the results using a linear mixed modelling approach to test the relationship between homework length and student algebra ability after controlling for confounding variables.

Method

Participants

The trial ran for four weeks during the second half of the autumn term of the 2018/19 academic year. All Year 7 students from three secondary schools in Devon participated in the trial with the following exceptions: those who opted out of the trial, those whose KS2 maths scaled score was not available, and those randomly excluded due to their class size not being a multiple of four. There were 20 participating classes in total across the three schools with a range of 4 to 28 students per class recruited in the trial. The total number of students recruited was 368. In every class, students were grouped into quartets based on ability measured by KS2 maths scaled score, such that within each quartet the ability of the four students was as similar as possible. Then, each member of the quartet was randomly assigned to receive one of the four specified percentages (measured by predicted time) of their normal algebra homework throughout the trial. These percentages were 0%, 50%, 100% and 150%, which equates to 0, 15, 30 and 45 minutes per week respectively. Students excluded from the trial were given homework as normal.

Assessment

Students were required to sit a test both at the beginning and end of the half-term during which the trial took place. The test was composed of two papers: one to evaluate a student's mathematical knowledge and skills (with an allowed time of 20 minutes), and one to evaluate a student's problem-solving ability (with an allowed time of 22 minutes). Questions in the second paper were taken from the pre-2008 national Key Stage 3 SATs exams. These papers were prepared separately for four levels of student ability, but critically there was a high level of question overlap between adjacent levels. This was necessary to allow for a global IRT analysis across all paper levels. As all classes participating in the trial were streamed by ability, paper levels were assigned per class. Besides algebra, these tests also included questions on two other topics taught during this half-term (number and ratio). All students followed the same scheme of learning with algebra being the second of the three topics taught.

Computing student abilities

Student abilities were evaluated using a one-parameter item response theory (IRT) analysis. This allowed the scale of abilities to be preserved despite students sitting different levels of the test, facilitated by the question overlap between papers for adjacent levels. IRT analyses can operate in two different ways: either both student ability and question difficulty can be estimated in parallel, or, if question difficulty has already been calibrated, these fixed difficulties can be used to estimate student ability. In this study we analysed the post-test results using the first method (estimating student abilities and question difficulties) and then analysed the pre-test results using fixed question difficulties estimated from the post-tests.

While student ability estimates obtained at the post-test stage were from algebra questions only, they were compared with ability estimates for all questions (algebra and non-algebra) at the pre-test stage, as the increased number of items provided a more accurate baseline measurement.

Multivariate analysis using linear mixed modelling

We performed multivariate mixed analysis to detect the relationship between homework length and attainment while accounting for factors known to influence these outcomes, such as prior student ability as measured by KS2 maths scaled score. Other factors, such as being in a specific class or school, can also be said to influence attainment, but in a random way. These factors are therefore best modelled as random rather than fixed effects, because we cannot hypothesise whether a given class or school will have a positive or negative influence on attainment.

We began this analysis by defining a baseline model, and then added in effects in the order of least to most experimentally manipulated. For example, a student's KS2 scaled maths score (fixed effect), stream (fixed effect), class (random effect) and school (random effect), were added one-by-one and evaluated as predictors before homework length (fixed effect) was added to the model and evaluated. This approach allowed us to identify important confounding variables (fixed or random effects) and control for them when testing the effect of homework length on student algebra ability, with time of test as a predictive factor. Homework length was introduced into the model as an interaction with time of test, since it could not have had an effect on students' pre-test attainment.

The Akaike Information Criterion (AIC) was used to test model improvement after each effect was added in turn. The AIC was chosen as a measure to test model improvement rather than the R2 value as it penalises increased model complexity. We took a drop in AIC with a p value of 0.01 or less as a significant result. Each effect was added in turn and an ANOVA carried out to test model performance with and without the effect.

Results

We computed the mean progress as indicated by improvement in student ability (post-test algebra ability – pre-test mathematical ability) made by students in each of the four trial arms. The results are shown in Figure 1. Mean progress was seen to jump considerably between the 0 and 50% points and then increase linearly with homework length. To confirm the statistical significance of this observation the data were subjected to a linear mixed model analysis, the results from which are summarised in Table 1. The baseline model consisted of KS2 maths scaled score and time of test as fixed effects, and student ID as a random effect. This model had an AIC of 2233.9. Additional random effects of school, stream and class were sequentially added to the model, but only class had a significant effect (AIC reduced to 2225.7, p=0.001) and was therefore retained in the model. Finally, homework length was modelled as an interaction with time of test and was found to be very significantly associated with change in student ability. The AIC of the model reduced to 2209, p = 0.000056. The amount of variance explained by the final model was 71.8% (47.4% by fixed effects and 24.4% by random effects). The R2 and RMSE values for the ability model were 0.814 and 0.743 respectively.

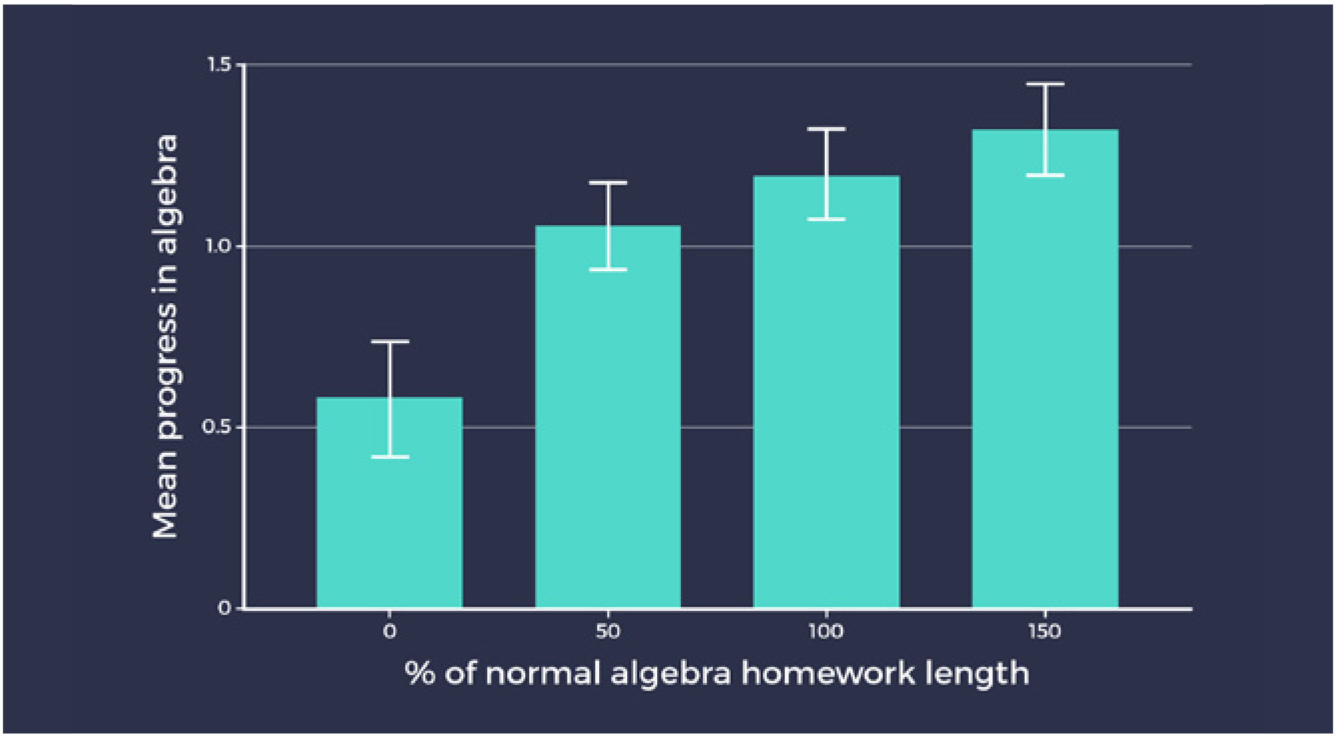

Figure 1. Mean improvement in student ability on algebra across the four trial arms. Each error bar represents the standard error of the mean. Note: abilities obtained from IRT analysis are on a logit scale.

Baseline/effect added *	AIC	P value	Retain in model?
Baseline **	2233.9	-	-
Stream (F)	2231.7	0.041	N
Class (R)	2225.7	0.00138	Y
School (R)	2227.4	0.557	N
Homework length*time	2208.5	0.000056	Y

* F and R in brackets after each variable denote whether it was treated as fixed or random respectively.

** The baseline model included time of test and KS2 maths scaled score as fixed effects and student ID as a random effect.

Table 1. Modelling student ability using a linear mixed model.

Discussion and conclusions

We carried out a four-arm, double-blind, randomised controlled trial to investigate the relationship between homework length and student attainment. Students were matched by ability and randomly assigned to one of the trial arms, such that there was no difference in student abilities between the trial arms. Students then received a set percentage of their normal algebra homework depending on the trial arm to which they were assigned. We found that algebra homework length accounted for a very significant proportion of the change in student attainment during the trial period.

While the results from this trial demonstrated a strong relationship between attainment and homework length, they did not identify an upper limit for homework length beyond which homework does not result in better outcomes. Nevertheless, it seems intuitively obvious that some upper limit must exist beyond the range used in this study.

Despite the high relevance of this topic to educators, and the large volume of research papers over decades that address this and similar questions, it has hitherto been impossible to come to any firm conclusion regarding the utility of homework. By adopting a completely new experimental paradigm that leverages some of the unique features of digital educational platforms, the current study unequivocally demonstrates that, at least in some circumstances, homework can be very effective at improving educational outcomes. What is particularly impressive is that these striking results were obtained in a relatively short timescale (4 weeks) with a manageable sample size (n = 368).

Of course, it is important to acknowledge that the relationship between homework length and student attainment was investigated in a very specific context: Year 7 students getting bespoke algebra homework assigned by and delivered through the digital Sparx platform. Whether these findings are applicable to any regularly set paper-based homework, or indeed homework distributed via other digital platforms, remains to be investigated.

References

Connolly, P., Keenan, C., & Urbanska, K. (2018). The trials of evidence-based practice in education: a systematic review of randomised controlled trials in education research 1980–2016. *Educational Research,* 60(3), 276–291. https://doi.org/10.1080/00131881.2018.1493353

Cooper, H. (1989). Synthesis of Research on Homework. *Educational Leadership*, 47(3), 85–91. Retrieved from https://eric.ed.gov/?id=EJ398958

Cooper, H., Robinson, J. C., & Patall, E. A. (2006). Does homework improve academic achievement? A synthesis of research, 1987-2003. *Review of Educational Research*, 76(1), 1–62. https://doi.org/10.3102/00346543076001001

Dettmers, S., Trautwein, U., & Lüdtke, O. (2009). The relationship between homework time and achievement is not universal: evidence from multilevel analyses in 40 countries. *School Effectiveness and School Improvement,* 20(4), 375–405. https://doi.org/10.1080/09243450902904601

Eren, O., & Henderson, D. J. (2011). Are we wasting our children's time by giving them more homework? *Economics of Education Review*, 30(5), 950–961. https://doi.org/10.1016/J.ECONEDUREV.2011.03.011

Lam, J. W. (1996). The employment activity of Chinese-American high school students and its relationship to academic achievement (Master's thesis, University of Texas at Arlington, 1996). In *Masters Abstracts International* (Vol. 34, pp. 317–337).

Rønning, M. (2011). Who benefits from homework assignments? *Economics of Education Review*, 30(1), 55–64. https://doi.org/10.1016/J.ECONEDUREV.2010.07.001

Torgerson, C. J. (2001). The need for randomised controlled trials in educational research. *British Journal of Educational Studies,* 49(3), 316–328. https://doi.org/10.1111/1467-8527.t01-1-00178

Trautwein, U. (2007). The homework–achievement relation reconsidered: Differentiating homework time, homework frequency, and homework effort. *Learning and Instruction,* 17(3), 372–388. https://doi.org/10.1016/J.LEARNINSTRUC.2007.02.009

Trautwein, U., Köller, O., Schmitz, B., & Baumert, J. (2002). Do Homework Assignments Enhance Achievement? A Multilevel Analysis in 7th-Grade Mathematics. *Contemporary Educational Psychology*, 27(1), 26–50. https://doi.org/10.1006/CEPS.2001.1084

Vatterott, C. (2009). *Rethinking homework: Best practices that support diverse needs. Principal-Arlington-* (Vol. 86).