Perseverance in the classroom: findings from a randomised educational intervention in primary schools in England*

Francesca Foliano

Sherria Hoskins

Heather Rolfe

Abstract

We evaluate a cluster randomised trial of a teacher-delivered programme aimed at increasing perseverance and academic achievement in primary school pupils across 100 schools in England. Year 6 teachers in treatment schools receive training in growth mindset theory and deliver an eight-week programme of structured classroom sessions. The intervention has no impact on Key Stage 2 test scores in reading, mathematics, or grammar, punctuation and spelling and this null result holds across all subgroups, including pupils eligible for Free School Meals. The intervention does produce a large shift in pupils' beliefs about the malleability of their intelligence (0.417 standard deviations), confirming that the programme was received as intended, and a positive effect on intrinsic motivation towards learning (0.127 SD).

JEL: I28, D04, C91, C93.

Key words: perseverance, mindset, attainment, beliefs.

*Correspondence: f.foliano@soton.ac.uk. We are grateful for discussions with participants at the International Workshop on Applied Economics of Education in Catanzaro, at the first J-PAL IDEE workshop at the Paris School of Economics, and at a seminar at the Center for Educational Leadership and Excellence, Stockholm School of Economics. This trial was registered on the International Standard Randomised Control Trial Number registry at `http://www.isrctn.com/ISRCTN14702744`. Ethics approval: University of Portsmouth, Science Faculty Ethics Committee (SFEC), 04/02/2016, Ref: SFEC 2016-007. The project has been funded by the Education Endowment Foundation, but the views expressed are those of the authors and not necessarily the Foundation. Visit `https://educationendowmentfoundation.org.uk`.

1 Introduction

Non-cognitive skills predict meaningful long-term outcomes such as schooling, employment, job performance, health and criminal behaviour (Heckman *et al.*, 2006; Almlund *et al.*, 2011; Borghans *et al.*, 2016). Unlike raw intelligence, non-cognitive skills are malleable and can be shaped until later years through investments and targeted interventions (Kautz *et al.*, 2014). The literature on the technology of skill formation emphasises that human development is a dynamic process where investments boost skills and skills are self-productive and exhibit dynamic complementarities (Cunha *et al.*, 2010). As a consequence early life investments in cognitive and non-cognitive skills maximise their returns, beginning with attainment in school.

Character traits such as self-perception, perseverance, motivation and self-control, have received particular attention in psychology and economics for their important role in fostering human capital accumulation and predicting later life outcomes (Heckman *et al.*, 2006; Almlund *et al.*, 2011; Kautz *et al.*, 2014). Perseverance in the form of drive for long term goals, often defined as grit, has been shown to be positively associated with educational attainment (Duckworth *et al.*, 2007, 2019; Jachimowicz *et al.*, 2018). A growing literature in educational psychology hypothesises that it is possible to increase perseverance in young people, and as a consequence improve their attainment and long-term outcomes, by teaching them that their brain has the potential to grow through effort exerted on challenging tasks (Dweck, 2006; Yeager and Dweck, 2012; Paunesku *et al.*, 2015; Yeager *et al.*, 2016). The belief that ability can grow through effort and dedication, called "growth mindset", is associated with greater incentives to exert effort over time, as pupils believe their work will pay off (Claro *et al.*, 2016).

Whether schools can foster perseverance at early ages, and thereby improve the achievement of young pupils, is a key question both for policy and for the economics of human capital: if schools can raise non-cognitive skills cheaply and at scale, the aggregate returns in terms of educational attainment and later labour market outcomes may be substantial (Heckman *et al.*, 2006; Kautz *et al.*, 2014). We evaluate a large-scale cluster randomised trial of a teacher-delivered programme aimed at increasing perseverance in primary school pupils across 100 schools in England. Year 6 teachers in treatment schools receive one day of training by educational psychologists introducing them to growth mindset theory and how to

embed it in everyday classroom practice, including how to give effort-focused feedback and frame mistakes as learning opportunities. Teachers then deliver an eight-week programme of structured weekly sessions (up to 2.5 hours per week, within normal class time). The aim of these sessions is to change the way pupils think about their intelligence: to build high expectations, motivation and resilience, and to encourage learning plans and goals that emphasise the development of skill and the expansion of knowledge. The underlying mechanism is that pupils who learn their intelligence can grow with effort have stronger incentives to persevere on challenging tasks and, as a consequence, improve their academic outcomes.

The primary outcomes are KS2 test scores in reading, mathematics, and grammar, punctuation and spelling (GPS): externally marked national assessments taken at the end of Year 6. To understand the mechanisms through which the intervention may affect attainment, we also examine its impact on pupils' beliefs about the malleability of their intelligence and on four dimensions of academic motivation: intrinsic value, self-efficacy, test anxiety, and self-regulation. Information on these non-cognitive outcomes is collected through questionnaires administered to all pupils before and after the intervention.

We supplement this quantitative evaluation of the intervention with a qualitative study, to i) better understand how the intervention was implemented and delivered in treatment schools, and whether this differed from the intended delivery model; ii) shed light on teachers' and pupils' responsiveness to the intervention, and the perceived impact on participants; iii) gather information on the activity of schools within the control group over the course of the trial period.

The intervention has no impact on test scores in reading, mathematics, or grammar, punctuation and spelling, as measured by externally marked national assessments at the end of Year 6. This null result applies across all pupils, including those from a disadvantaged background and those who self-reported low beliefs in the malleability of their ability at baseline. The intervention does produce a large shift in pupils' beliefs about the malleability of their intelligence (0.417 SD, 95% CI: 0.27 to 0.57), consistent with the programme having been received as intended. There is also a borderline significant positive effect on intrinsic motivation towards learning (0.127 SD).

An extensive literature in psychology and economics investigates the impact of school-

based interventions rooted in the growth mindset approach or more generally aimed at raising grit and perseverance among young people. Bettinger *et al.* (2018) show that a brief intervention targeting beliefs about the malleability of intelligence improves performance in a maths test three weeks after the intervention in Norway, with effects concentrated among pupils who initially reported a fixed mindset. With a similar brief intervention, Paunesku *et al.* (2015) find that treated pupils at risk of dropping out in 13 US schools perform better in end-of-semester grade point averages. Yeager *et al.* (2016) test a growth mindset intervention among 9th graders and find positive effects on core-course grade point averages for low-achieving pupils. In the largest study to date, Yeager *et al.* (2019) conduct a national experiment across 65 US public high schools involving over 12,000 students and find that a brief online intervention raises end-of-year GPA among lower-achieving students, with no significant effect on high achievers. A systematic review by Sisk *et al.* (2018) and a subsequent meta-analysis by Burnette *et al.* (2022) confirm that the average effect of growth mindset interventions on academic achievement is small and concentrated among students from disadvantaged backgrounds.

In all of these studies, subjects are secondary school pupils or college students and the intervention is brief, typically delivered online or via short classroom sessions. Huillery *et al.* (2025) test the impact of a more pervasive intervention conducted in 97 French disadvantaged middle schools to increase perceived return to effort among adolescents by developing an internal locus of control and the belief that intelligence is not a fixed trait. Treated pupils attend three sessions a year for four years led by external facilitators; they attain higher test scores, improve their behaviour and report a change in their mindset and higher educational and professional aspirations. Santos *et al.* (2022) conduct a nationwide field experiment with middle-school students comparing a self-directed and a teacher-delivered grit programme; the teacher-delivered arm produces positive effects on socio-emotional skills including grit, while impacts on grade point averages are more limited. The programme evaluated by Alan *et al.* (2019) differs from the previous studies in that subjects are younger children, enrolled in primary state schools in Istanbul, and the intervention is delivered by teachers over a set of sessions that cover several weeks of one academic year. Treated pupils in this trial are more likely to set challenging goals, exert effort and obtain higher grades.

Evidence on the effectiveness of teacher-delivered programmes aimed at raising perseverance and grit comes almost entirely from non-English-speaking settings, and no prior study

has tested such a programme against externally graded national test scores at national scale. We fill this gap with a cluster RCT across 100 primary schools in five English regions, using KS2 national assessments — blind-marked, high-stakes tests — as the primary outcome. The closest study is Alan *et al.* (2019), who find large effects on achievement in Istanbul primary schools; our null result on attainment, alongside a treatment effect on mindset beliefs of 0.417 standard deviations (approximately 50% larger than theirs), points to a context-specific explanation. England's national curriculum already centres the personal development of pupils (Department for Education, 2013) and Ofsted holds schools accountable for pupils' broader development alongside academic achievement (Ofsted, 2015). Qualitative evidence from our study suggests that some teachers in both treatment and control schools had prior awareness of the ideas underlying the intervention, though none had received formal training on them. Taken together, these findings suggest that the effectiveness of perseverance-oriented programmes on attainment depends critically on the baseline context, and that in school systems already oriented towards fostering non-cognitive skills, the marginal returns to such interventions may be limited.

The rest of the paper is structured as follows: section 2 describes the intervention and the evaluation design of the trial; section 3 presents the data and baseline descriptive statistics; section 4 presents the results; section 5 reports the main findings of the qualitative analysis; section 6 discusses the results and section 7 concludes.

2 Intervention and Evaluation

2.1 Content of intervention

The aim of the intervention studied in this trial is to increase the achievement of primary school pupils in numeracy and literacy by helping them to develop the belief that their intelligence is malleable and can grow with effort and dedication. The conjecture at the core of the intervention is that by changing teachers' behaviour and language as a result of a specific training, along with the delivery of precise sessions in the classroom, pupils change their own mindset and expectations about their own ability. This leads to changes in pupils' learning behaviour as they become more resilient and engage in challenge and opportunity. As a consequence, pupils' academic attainment improves (Blackwell *et al.*, 2007).

The intervention had two components. First, Year 6 teachers received one day of training that was prepared and delivered by a team of educational psychologists. The training introduced teachers to the growth mindset literature and evidence and provided suggestions on how to embed the approach in their classrooms and lessons. In particular teachers were advised to give feedback by praising pupils for their effort, approach and own outcome rather than for their ability or final result compared to others and to repeat the message that making mistakes is an opportunity to learn rather than a negative experience.

Second, to supplement these suggested changes to everyday practice and to effectively introduce pupils to the idea of incremental intelligence through structured teaching and learning activities, teachers were given the materials to run an eight-week programme (up to 2.5 hours a week) of weekly lessons with their Year 6 pupils. These sessions took place in normal class time. The aim of these sessions was to change the way that pupils think about their intelligence, in particular to build high expectations and resilience and to encourage specific plans and goals, emphasise the choice of the right learning strategies, the development of skill, and the expansion of knowledge. Teachers were given a teaching manual that included comprehensive lesson plans for the eight sessions and a USB stick with additional material to support their interaction with children and their parents. In addition, teachers were granted free access to online videos. These videos were referenced in the lesson plans and teachers were encouraged to use the videos to explain or emphasise particular concepts such as resilience and learning from mistakes. The videos focused on three famous people, Darwin, Einstein, and Wilma Rudolph, describing their lives and how they overcame adversity and succeeded, in line with the growth mindset message. Other materials provided by the project team included posters on growth mindset and pupil quizzes. During the training day teachers were carefully guided through the content of the materials and given practical tips on how to make these sessions more effective.[1]

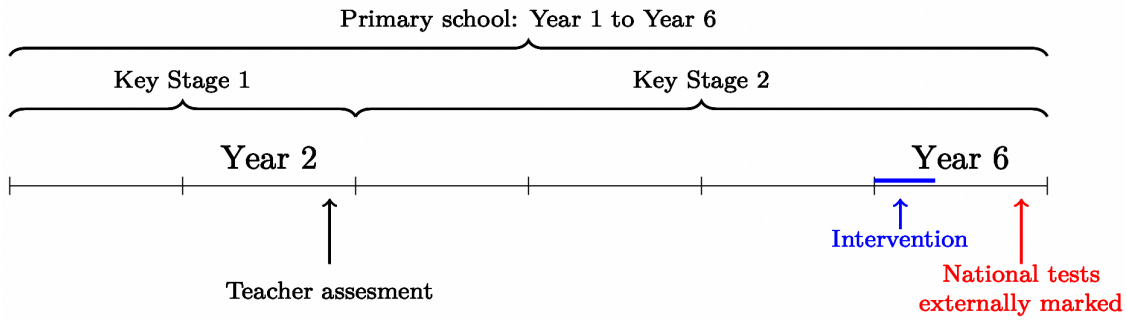
2.2 Evaluation Design

Primary education in England starts at age 5 in Year 1 and lasts for 6 years. During the primary education period there are two stages: Key Stage 1 (KS1), that ends in Year 2 when pupils are assessed by their teachers; Key Stage 2 (KS2) that runs between Year 3

[1] Appendix A includes a sample of slides from the training day.

and Year 6, ending with externally marked national tests.[2] The majority of pupils, around 93%, attend state schools. The intervention we evaluate in this study is aimed at teachers and pupils in primary state schools across England. It was implemented at the beginning of Year 6 to test its effect six months later on performance in externally marked national tests. Figure 1 summarises the institutional setting of primary education in England and the timing of the intervention.

Figure 1: Institutional settings and timing of the intervention



The design of this study is a two-arm cluster randomised trial, with primary schools participating in the trial randomly assigned to either the intervention group or the control group. School-level randomisation was chosen over class or pupil level to minimise the chance of contamination of control by treatment. The trial had a wait-list design to keep the control schools engaged and to minimise the risk that they would seek a similar formal training in the same academic year. All schools paid £500 and received the full training. Schools in the treatment group received their training in September/October 2016, while wait-list control schools received the training two academic terms later, after the date of the national Standard Assessment Tests (SATs).

Schools were allocated to the treatment or the control group using a block randomisation technique, controlling for prior attainment at the school level and geographical area. Blocks were defined according to terciles of school-level KS1 performance within each of five locations: Midlands; North East; North West; South East; and South West.[3] The purpose of

[2]Teachers' evaluations are also available for Year 6.

[3]Schools were allocated 50:50 to the intervention group or to the control group after their recruitment and consent forms had been received. Each school was assigned a random number between 0 and 1 from a uniform distribution. Schools were then sorted by block and random number. The first school in the list was

blocking was to improve the balance between the treatment and control groups in terms of key outcome-related characteristics. Figure 2 shows the geographical distribution of treatment and control schools after the randomisation. The trial was delivered as it would be at scale and 101 schools with about 5,000 Year 6 pupils in South East, South West, Midlands, North East and North West were recruited to the trial.

Figure 2: Treatment and control schools

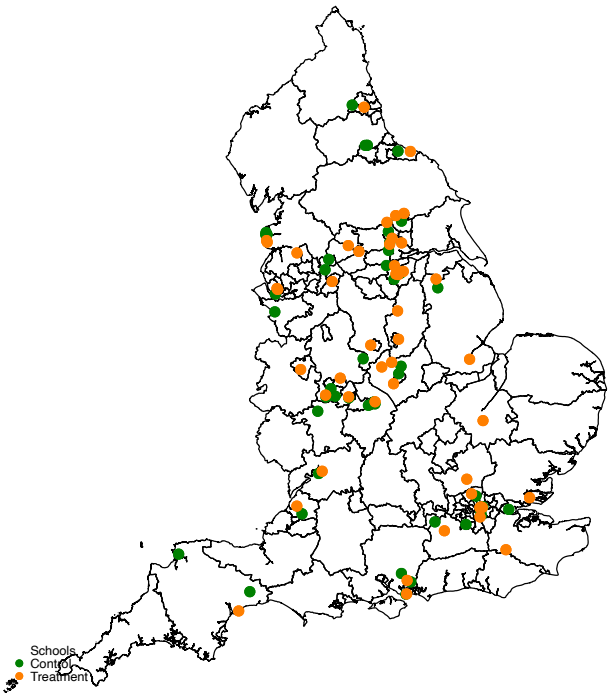


Table 1 reports the timeline of the trial. Recruitment involved adverts online, on social media (e.g. Facebook groups for Year 6 teachers) and emails from the National College of Teachers. A snowball approach was also used whereby interested schools and other school contacts were asked to pass on the advert to other schools in their own networks, and some schools volunteered via this route. 179 schools initially registered their interest. All volunteer schools were contacted to check that they met the inclusion criteria for the project and to outline the key commitments for the school. Schools were eligible to participate in the study

randomly assigned to be in the treatment or control group. Each subsequent school in the list was assigned to be in the opposite group of the previous schools, thus assuring an equal distribution allocation. The Stata code for this randomisation is reported in Foliano *et al.* (2019).

if: i) they had not used a systematic mindset programme with their Year 6 cohort before; ii) they were able to attend one of the training dates provided by a team of educational psychologists.

Table 1: Timeline of the programme

Date	Activity
Jan 2016	Schools approached
Jan-May 2016	Schools recruited and agree to participate
May/July 2016	Pre-trial Mindset and MSLQ questionnaires
Jun-16	Schools randomly allocated to treatment/control group
Sep-16	Intervention group attended 1-day training event
Sept-Dec 2016	Intervention delivered in schools by intervention group
Sept 2016-Feb 2017	Fidelity survey to treatment schools
Dec 2016-Feb 2017	Fieldwork visits to 6 treatment schools
Jan-Feb 2017	End of project survey to treatment schools
Jun-17	Control group survey
May-17	Pupils sit Key stage 2 Exams
March/July 2017	Post-trial Mindset and MSLQ questionnaires
Jun-17	Control group attended 1-day training event
Autumn term 2017	Intervention delivered by control schools

Schools that were interested in taking part and that met the eligibility criteria were emailed a "Welcome pack" which included a description of the study in detail, a permission form for the headteacher to complete, an information letter to parents giving them the option to withdraw their child from the study, a form to request the children's unique pupil numbers (UPNs) and the "School Questionnaire" to be administered to pupils to collect baseline information on non-cognitive measures such as beliefs, motivations and learning attitudes. Prior to being randomised schools needed to provide the following: i) head teacher's consent to their school taking part in the trial; ii) confirmation that the information letters to parents were sent out; iii) UPNs for all Year 6 pupils; iv) completed pre-test non-cognitive measures.

Schools entered the randomisation phase when they returned the head teacher consent

form and were informed of their allocation only after returning all remaining material. All pupils attending Year 6 in selected schools were eligible for the treatment.[4]

Year 6 teachers in the treatment schools attended the training in September 2016. Teachers were asked to deliver the eight sessions of the programme to their pupils between September 2016 and January 2017. During these months, treatment schools were administered fidelity surveys. Fieldwork visits and end of project surveys were carried out between December 2016 and February 2017. Treated pupils sat their SATs in May 2017 and post-trial questionnaires on beliefs, motivations and attitudes towards learning were administered to all pupils between March and July 2017. Finally, schools in the control group responded to a survey in June 2017 and attended the training in the same year in September.

2.3 Outcomes of interest

The main cognitive outcomes considered in the evaluation of this intervention are the externally-marked tests that pupils sit at the end of KS2. These outcomes are better than teachers' assessments for two main reasons. First, they are blind assessments produced by external markers, therefore, they cannot be affected by teachers' knowledge of the treatment status of the pupils. Second, they are finely scored, providing high granularity that allows identification of potentially small effects of the treatment. We include scores from three tests: Reading, Maths and Grammar, Punctuation and Spelling (GPS).

The secondary set of outcomes includes measures of beliefs about the malleability of intelligence, and of motivations and attitudes towards learning. Questionnaires administered by teachers to pupils included three items designed to measure pupils' mindsets, based on Dweck (2006). The aim of these items was to measure whether pupils held a fixed mindset — the belief that intelligence is a fixed trait — or a growth mindset — the belief that intelligence can grow with effort. The mindset measure created with these questions has been used and validated in several experimental and economic studies (see for instance Yeager *et al.* (2016); Dweck (2006); Alan *et al.* (2019)).

[4]Six schools dropped out from the study after the randomisation and did not provide any pre-treatment information and therefore, were not included in the main analysis.

Questionnaires also included 31 items from the standard Motivated Strategies for Learning Questionnaire (MSLQ) (Pintrich and De Groot, 1990; Pintrich, 1991; Pintrich *et al.*, 1993). The MSLQ was developed originally to study college student learning. It draws on a social-cognitive theoretical framework that assumes motivations and learning strategies are not fixed traits of the learner, but vary by context and academic task (Duncan and McKeachie, 2005). Four sub-scales were selected: self-efficacy; intrinsic value; test anxiety and self-regulation — a widely used set of measures of motivations and learning strategies. The intrinsic value scale measures to what extent pupils' participation in an academic task is an aim in itself, rather than participation being a means to an aim; the self-efficacy scale measures how pupils judge their ability to accomplish a task as well as their confidence in their skills to perform that task; the test anxiety scale measures students' concern about taking an exam; finally, the self-regulation scale measures pupils' ability to plan, monitor and regulate their cognitive strategies to succeed in an academic task. The original questions were adapted by the educational psychologists that designed the content of the training to make them relevant to the age (i.e. primary school children) and location (i.e. England) of participants in our programme. The full set of questions is reported in Appendix B.

3 Data and baseline information

Initially 168 schools were approached; of these, 67 declined to participate in the study. 101 schools were included in the randomisation: 50 were allocated to the intervention group and 51 were allocated to the control group. After the randomisation, 6 schools (2 in treatment, 4 in control group) withdrew from the study and were not included in the final analysis as they did not provide baseline information.

The precise number of pupils initially randomised into the trial is 5,018. For these pupils the unique pupil numbers (UPNs) were obtained from the 101 schools participating in the study. 11 pupils were withdrawn from the trial by their parents and 132 UPNs could not be matched to administrative records of pupils in those schools. Of the remaining 4,875 pupils randomised in the trial, 291 are lost because 6 schools withdrew from the study after the randomisation (85 in 2 treatment schools and 206 in 4 control schools). The final number of pupils that could be included in the primary analysis is 4,584. The total rate of attrition for primary outcomes is 11.58% for reading, 11.24% for maths and 11.36% for GPS. Attrition is

a bigger problem for beliefs and attitudes towards learning: approximately 31% of originally randomised pupils are missing post-trial questionnaire data for the mindset measure and all four MSLQ subscales. The with-covariate analysis samples (N $\approx$ 2,900) are smaller still, owing to listwise deletion on pre-treatment questionnaire measures used as controls; this reflects the covariate specification rather than additional non-response. Appendix Table C2 compares baseline characteristics of questionnaire respondents and non-respondents: attrition is not differential by treatment arm (p = 0.973), and respondents and non-respondents are similar on prior attainment and all baseline non-cognitive measures, though girls and White pupils are somewhat more likely to respond.

Data used in the analysis come from two sources: administrative records on attainment and demographic characteristics from the National Pupil Database (NPD); information on pupils' beliefs and attitudes towards learning and school collected before and after the treatment with a questionnaire administered by the teachers in the classroom.[5]

The cognitive outcomes of interest are standardised test scores in reading, mathematics, and grammar, punctuation and spelling, as measured by the KS2 national assessment tests in Reading, Grammar, Punctuation and Spelling (GPS) and Maths. KS2 tests are taken by all pupils at the end of Year 6 (age 10-11) and are externally graded. KS1 teacher assessments in Reading, Writing and Maths are jointly used to define a comprehensive set of dummies for prior attainment. [6]

In the regression analysis we use standardised KS2 scaled point scores that, in their raw form, range from 80 to 120, as cognitive outcomes and dummies for categories of average

[5]Schools administered the pre-trial questionnaires to pupils when they were in Year 5 (March–July 2016). The post-trial questionnaires were administered by schools after the KS2 national assessments (May–July 2017).

[6]In KS1 teachers assess pupils against level of achievement (W, 1, 2C, 2B, 2A, 3, 4). To control for prior attainment we convert these levels into points by using the conversion tables provided by the Department for Education (Department for Education, 2015) and we then calculate the average KS1 attainment based on non-missing KS1 assessments. We then define the following categories: 11-14.99 points (did not achieve grade level), 15-16.99 points (achieved grade level), more than 17 points (achieved above grade level), or all missing. These categories are then used to create a set of dummy variables to control for prior attainment in the empirical analysis. KS1 assessments for the pupils in the study were held in the academic year 2012/13, when pupils in the trial were in Year 2 (age 6-7).

scores in KS1 as controls for prior attainment. KS2 test scores are considered robust measures of achievement. Dummies for categories of average KS1 attainment are a comprehensive measure of prior attainment as they contain information from three different assessments and they have the advantage of reducing the impact of missing KS1 assessments.

The three mindset items are scored from 1 to 6, with 1 indicating "Strongly disagree" and 6 indicating "Strongly agree". All three items are framed in the fixed mindset direction (for example, "You can't really change your basic intelligence"), so raw scores are reversed before constructing the composite, giving a scale where higher values indicate stronger growth mindset beliefs. All items on the adapted version of the MSLQ are scored from 1 to 7, with 1 indicating "Strongly disagree" and 7 indicating "Strongly agree". The measure of beliefs and the four subscales (intrinsic value, self-efficacy, test anxiety and self-regulation) are then calculated as a mean score of the relevant items.[7]

We use baseline measures to assess balance between the treatment and control groups. Table 2 reports the main school characteristics for the schools that participated in the randomisation and their pupils by treatment status. The most relevant difference at school level is the proportion of academy schools: 29.2% in the intervention group and 19.1% in the control group. Schools in the intervention group also have more pupils on roll than schools in the control group (a mean of 362.0 in treatment schools and of 340.6 in control schools). There is also a small difference in the school proportion of EAL pupils (17.5% in the intervention group and 15.6% in the control group). There are no other relevant differences between intervention and control schools in school-level characteristics.

Table C1 in Appendix C reports the comparison between the schools in the trial and the other schools in England. This comparison is important because it defines the external

[7]One item in the intrinsic value subscale is excluded from the composite measure used in the analysis. Item 14, "Even when I do poorly on a test I try to learn from my mistakes", does not appear in either the intrinsic goal orientation or task value subscales of the original MSLQ (Pintrich and De Groot, 1990; Pintrich *et al.*, 1993) from which the adapted scale is drawn. The remaining eight items each correspond directly to an item in one of those two original subscales. In addition, the content of item 14 directly mirrors a core message of the intervention, namely that making mistakes is an opportunity to learn, which introduces a risk of content contamination when estimating treatment effects on intrinsic motivation. The intrinsic value subscale is therefore constructed from eight items. Results based on the full nine-item scale are reported in Appendix D.

validity of our trial by showing whether a specific type of school self-selects into the programme, potentially biasing the results. Schools in the trial have more Free School Meal (FSM) pupils and more pupils with English as an Additional Language (EAL) and Special Educational Needs support. Pupils from more disadvantaged backgrounds are more likely to benefit from this type of intervention (Sisk *et al.*, 2018), therefore, we expect a higher average treatment effect than what we would obtain in a representative sample of English schools.

Table 2: Baseline comparison of intervention and control group: school characteristics

	Intervention	Control	difference	(p-value)
Religious school	0.229	0.255	-0.026	0.769
Academy	0.292	0.191	0.100	0.259
Community school	0.438	0.532	-0.094	0.363
Foundation school	0.146	0.064	0.082	0.197
Voluntary aided	0.125	0.191	-0.066	0.380
Ofsted: Outstanding	0.146	0.170	-0.024	0.748
Ofsted: Good	0.750	0.723	0.027	0.771
Ofsted: Satisfactory	0.083	0.085	-0.002	0.976
Ofsted: Inadequate	0.021	0.021	0.000	0.988
Number of pupils	361.958	340.574	21.384	0.474
% of Free School Meal	14.867	16.598	-1.731	0.462
% SEN with support	13.061	12.593	0.467	0.716
% SEN with statement	1.223	1.506	-0.283	0.218
% English Additional Language	17.513	15.577	1.936	0.675
N	*48*	*47*		

Table 3 shows that there are also a few small differences between treatment and control groups in pupil-level characteristics. For example, there is a marginally higher proportion of white pupils (73.1%) and a lower proportion of black pupils (4.0%) in the intervention compared to the control group (respectively 68.7% and 7.2%). Although the difference in the proportion of black pupils is small in terms of percentage points, it is notable in percentage terms given the low baseline proportion. The other individual characteristics and pupil pre-intervention outcomes in reading and maths, and the mindset and anxiety measures are,

on average, balanced between the two groups. There are small differences in the measures of intrinsic value, self-efficacy and self-regulation; standard errors clustered at the school level yield p-values above 0.05 for all three, and each is included as a control in the corresponding outcome regression.

Table 3: Baseline comparison of intervention and control group: individual characteristics

	Intervention	Control	difference	(p-value)
Female	0.493	0.480	0.013	0.388
Ever FSM	0.346	0.367	-0.021	0.650
Black	0.040	0.072	-0.033	0.135
White	0.731	0.687	0.044	0.600
Reading points	16.392	16.255	0.136	0.563
Writing points	15.005	15.123	-0.119	0.615
Maths points	16.208	16.157	0.050	0.809
N	*2,209*	*2,122*		
Mindset measure	3.825	3.856	-0.031	0.785
Intrinsic value	5.555	5.646	-0.091	0.168
Self-efficacy	5.206	5.294	-0.088	0.163
Anxiety	3.801	3.723	0.078	0.445
Self-regulation	4.740	4.838	-0.098	0.092
N	*2,034*	*1,759*		

In summary, there is a good balance between the treated and the control group in school and pupil characteristics and the small differences found seem fairly consistent with the amount of imbalance that would occur due to chance.

4 Results

We estimate the following model to assess the effect of the intervention on KS2 academic outcomes and on mindset beliefs and attitudes towards learning and school:

$$y_{i,j} = \beta_0 + \beta_1 T_j + X'_{i,j}\gamma + \epsilon_{i,j} \quad (1)$$

where T_j is an indicator that equals 1 if school j is in the treatment group and 0 otherwise, and X_{ij} is a set of observable characteristics for student i in school j that could be predictive of the outcomes of interest in our analysis.

When the outcome of interest is a test score, X includes achievement at KS1, an indicator for gender and for FSM status, a set of dummies for ethnicity and a set of dummies for the randomisation blocks. For beliefs and attitudes X includes pre-treatment measures in the same domain, an indicator for gender and FSM status, and a set of dummies for ethnicity and a set of dummies for the randomisation blocks. Standard errors are clustered at the school level. The estimated β_1 is the average treatment effect. We present results without and with covariates for the full sample; the no-covariate specification includes only the treatment indicator and block dummies. Heterogeneity analyses by socioeconomic status, gender and baseline mindset are reported in Section 4.2.

4.1 Main results

Table 4 reports the estimated average treatment effect for KS2 Maths, Reading and GPS for the full sample, without and with covariates.

Table 4: Treatment effect on attainment at Key Stage 2 (age 10–11)

	Without covariates			*With covariates*		
	Maths	Reading	GPS	Maths	Reading	GPS
T	-0.045	-0.011	0.008	-0.034	-0.012	0.010
	(0.055)	(0.051)	(0.058)	(0.048)	(0.039)	(0.045)
N	4,454	4,437	4,448	4,454	4,437	4,448
R^2	0.041	0.052	0.046	0.447	0.440	0.515

Note: KS2 measures of achievement are standardised to have mean zero and standard deviation one. Without covariates: treatment indicator and randomisation block dummies only. With covariates: adds KS1 attainment, gender, FSM status and ethnicity dummies. Standard errors in parentheses are clustered at the school level. *** p < 0.001, ** p < 0.01, * p < 0.05.

Across all three outcomes and both specifications, the point estimates are tiny and negative or near-zero, with confidence intervals that are symmetric around zero and comfortably within what would conventionally be considered negligible territory. The 95% confidence intervals (with covariates) are $(-0.13; 0.06)$ for Maths, $(-0.09; 0.06)$ for Reading, and $(-0.08; 0.10)$ for GPS; the treatment had no educationally meaningful effect on attainment. Appendix F confirms that this null holds for the subsample of pupils with post-trial questionnaire data.

Table 5 reports the estimated average treatment effect for the mindset measure and for the measures of intrinsic value, self-efficacy, test anxiety and self-regulation.

Table 5: Treatment effect on beliefs and attitudes towards learning and school (age 10–11)

	Mindset	Intrinsic	Self-efficacy	Anxiety	Self-regulation
			Without covariates		
$T = 1$	0.353***	0.083	-0.056	-0.029	0.066
	(0.071)	(0.075)	(0.059)	(0.051)	(0.058)
N	3,460	3,440	3,440	3,440	3,427
R^2	0.057	0.029	0.021	0.012	0.025
			With covariates		
$T = 1$	0.417***	0.127†	-0.025	-0.016	0.090
	(0.074)	(0.067)	(0.058)	(0.053)	(0.055)
N	2,902	2,917	2,916	2,916	2,899
R^2	0.139	0.265	0.275	0.235	0.244

Note: The mindset measure and the MSLQ sub-scales are standardised to have mean zero and standard deviation one. Without covariates: treatment indicator and randomisation block dummies only. With covariates: adds pre-treatment measures in the same domain, gender, FSM status and ethnicity dummies. The larger sample in the without-covariate specification reflects the absence of listwise deletion on pre-treatment measures. Standard errors in parentheses are clustered at the school level. † p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001.

The results show that the intervention produced a large shift in beliefs about the malleability of intelligence, consistent with the programme having been received as intended. Under the with-covariate specification, the average treatment effect is 0.417 of a standard deviation with confidence interval (0.27; 0.57); the estimate without covariates is 0.353 SD, confirming that the result is not driven by covariate adjustment. The average treatment

effect on the intrinsic-value measure (with covariates) is positive and statistically significant at the 10% level: 0.127 with confidence interval (−0.01; 0.26). Without covariates, the point estimate is smaller (0.083) and no longer significant at the 10% level, suggesting that controlling for the pre-treatment measure increases precision. The average treatment effect is not statistically different from zero for the remaining three sub-scales under either specification. The effect for self-efficacy is −0.025 with confidence interval (−0.14; 0.09); for test anxiety is −0.016 with confidence interval (−0.12; 0.09); and for self-regulation is 0.090 with confidence interval (−0.02; 0.20). With an additional set of exploratory regressions we examine which items of the intrinsic value sub-scale show the largest treatment effects. Table D1 in Appendix D shows that the clearest pattern is for the statement "I prefer class work that is challenging so I can learn new things", for which treated pupils consistently reported greater agreement. This is in line with the content of the intervention to which they were exposed, and may point to a genuine shift in learning orientation; the full composite is borderline significant at the 10% level ($p \approx 0.06$). Appendix E confirms adequate unidimensional structure for each scale via confirmatory factor analysis and replicates the with-covariate results from Table 5 replacing item means with Bartlett-weighted factor scores; estimates are near-identical, confirming that the results are not sensitive to the choice of aggregation method.

4.2 Heterogeneity analyses

Tables 6 and 7 present treatment effects on KS2 attainment for subgroups defined by socioeconomic status and gender (Table 6) and by baseline mindset beliefs (Table 7).[8] In all subsamples, the treatment effect on attainment is close to zero and precisely estimated, mirroring the full-sample null result.

Table 8 presents subgroup results for the secondary outcomes. The treatment effect on mindset beliefs is large and statistically significant for both FSM pupils (0.477 SD) and girls (0.464 SD), with confidence intervals that do not overlap zero. Point estimates for intrinsic value are similarly positive for FSM pupils (0.161) and girls (0.142). The average treatment effect is not statistically different from zero for self-efficacy or test anxiety in either subgroup, nor for self-regulation among FSM pupils. Among girls, the point estimate for self-regulation

[8]Pupils with fixed mindset are those in the lowest tercile of the distribution of beliefs about the malleability of intelligence; pupils with growth mindset are instead those in the highest tercile.

Table 6: Treatment effect on attainment at Key Stage 2 by subgroup

	FSM			*Girls*		
	Maths	Reading	GPS	Maths	Reading	GPS
T	-0.028	0.006	-0.017	-0.008	-0.023	0.007
	(0.061)	(0.051)	(0.055)	(0.049)	(0.044)	(0.042)
N	1,579	1,574	1,576	2,190	2,181	2,186
R^2	0.448	0.436	0.536	0.428	0.448	0.515

Note: KS2 measures of achievement are standardised to have mean zero and standard deviation one. All regressions include achievement at KS1, an indicator for gender (FSM columns) or FSM status (girls columns), a set of dummies for ethnicity and a set of dummies for the randomisation blocks. Standard errors in parentheses are clustered at the school level. *** p < 0.001, ** p < 0.01, * p < 0.05.

Table 7: Treatment effect on attainment at Key Stage 2 by baseline mindset

	Fixed			*Growth*		
	Maths	Reading	GPS	Maths	Reading	GPS
T	-0.002	0.035	-0.015	-0.131	-0.045	-0.033
	(0.073)	(0.062)	(0.065)	(0.067)	(0.060)	(0.066)
Observations	1,243	1,237	1,241	1,033	1,035	1,033
R-squared	0.392	0.418	0.490	0.472	0.457	0.502

Note: KS2 measures of achievement are standardised to have mean zero and standard deviation one. All regressions include achievement at KS1, an indicator for gender and for FSM status, a set of dummies for ethnicity and a set of dummies for the randomisation blocks. Standard errors in parentheses are clustered at the school level. *** p < 0.001, ** p < 0.01, * p < 0.05.

is positive (0.136 SD, SE = 0.069, p = 0.051) and significant at the 10% level. Girls exhibit higher baseline self-regulation than boys (Duckworth and Seligman, 2006; Zimmerman and Martinez-Pons, 1990), which may make them better placed to act on the intervention's emphasis on learning strategies , a candidate explanation for the significant effect observed here.

Table 8: Treatment effect on beliefs and attitudes towards learning and school by subgroup

	Mindset	Intrinsic	Self-efficacy	Anxiety	Self-regulation
			FSM		
$T = 1$	0.477***	0.161†	-0.077	0.043	0.086
	(0.108)	(0.096)	(0.085)	(0.073)	(0.085)
N	953	955	955	955	948
R^2	0.140	0.240	0.219	0.153	0.221
			Girls		
$T = 1$	0.464***	0.142†	0.005	-0.008	0.136†
	(0.081)	(0.075)	(0.067)	(0.069)	(0.069)
N	1,465	1,473	1,473	1,473	1,466
R^2	0.155	0.235	0.271	0.204	0.219

Note: The mindset measure and the MSLQ sub-scales are standardised to have mean zero and standard deviation one. All regressions include pre-treatment measures in the same domain, an indicator for gender (FSM panel) or FSM status (girls panel), a set of dummies for ethnicity and a set of dummies for the randomisation blocks. Standard errors in parentheses are clustered at the school level. † p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001.

5 Qualitative analysis

The empirical evaluation of the intervention was complemented by a qualitative analysis with the aim to understand how the intervention programme was implemented and delivered in treatment schools, and to what extent this differed (if at all) from the intended delivery model. It also aimed to shed light on teachers' and pupils' responsiveness to the intervention, and the perceived impact the programme had on those involved. Finally, the qualitative analysis sought to gather information on the activity of schools within the control group over the course of the trial period, and to assess the extent to which control schools were already engaging with similar approaches.

The research methods used included the following:

- fieldwork visits to six treatment schools conducted from December 2016 to February 2017, which included:

  i) semi-structured interviews with nine Year 6 teachers, three deputy/head teachers,

and four teachers from other Year groups;

ii) three focus groups with eight to ten Year 6 pupils receiving the intervention; and

iii) three lesson observations of the final Changing Mindsets session;

iv) end of project survey, administered to treatment schools from January to February 2017 (24 schools participated);

- fidelity survey, administered to treatment schools over the course of the intervention from September 2016 to February 2017; and
- control group survey, administered in June 2017.

While all the different activities of the qualitative study involve small samples of schools, they provide interesting insights that could help interpret the null effect of the intervention. In treated schools most teachers said they had some familiarity with growth mindset theory and its pedagogical implications, although few had yet tried to embed them in their teaching.

Teachers welcomed the training day, though some preferred more practice with activities rather than the theoretical basis for growth mindset. Some attendees noted that they would have liked the training to have spent more time going through and discussing the intervention materials, such as the lesson plans, and to have had further opportunities to practise delivering some of the suggested activities and reflect on how they may best be applied in the classroom.

In the schools that participated in the fieldwork, and in those that responded to the fidelity survey, teachers reported variations in the length of the sessions, their regularity, and the number of suggested activities Year 6 teachers were able to get through. However, the majority of respondents who were involved in delivering these sessions felt that pupils had understood the ideas and concepts they had been taught and engaged with the material of the sessions.

Finally, teachers reported that the most widespread change in practice was altering the nature of the feedback and praise that they provided to pupils. This included omitting any fixed mindset language that was person- and/or ability-centred and focusing instead on praising the process by which pupils complete a particular task, and their outcomes in relation to their progress in learning, as opposed to the end result compared to others.

6 Discussion

There are three possible reasons for the result of the evaluation: the programme was not delivered as intended, or was too short, so that pupils did not take on its messages and change their attitudes, behaviours and, consequently, their performance; the pupils were too young and older children are better at using the theory of incremental intelligence to improve their motivation in academic tasks and raise their performance, particularly as reflected in tests; treatment and control schools were already using some of the main concepts of the intervention in the classroom, or more generally all schools were already committed to fostering the individual development of their pupils to some extent.

We can rule out the possibility that the programme was not delivered as intended. The qualitative analysis indicates a good level of fidelity with limited adaptations to the programme. The programme was very well received, with high buy-in from schools and teachers, and was seen to address a recognised need to change pupils' view of their own intelligence and approach to learning. The correct implementation of the programme is also supported by the consistent shift in pupils' beliefs about the malleability of their intelligence.

A further explanation could be that the pupils were too young to translate the intervention's messages into changed learning behaviours and academic outcomes. It is also possible that, at age 11, pupils may be too young to self-direct their learning in a way which is possible for older pupils, such as those involved in the studies in US, Norway and France. There may be intermediate steps between adopting the incremental theory of intelligence and improving academic achievement that younger pupils are less able to utilise. These may include, for example, understanding strengths and weaknesses, when to seek help, how to express the help they need and then to listen and process the feedback. These abilities may require a degree of maturity and self-understanding which may not be possible for children of primary school age.

Finally, schools which had already delivered a systematic programme based on the growth mindset theory were excluded from the trial. However, the qualitative analysis highlighted that some teachers in both treatment and control schools were already familiar with the main concepts delivered by the programme. In addition, the national curriculum in England highlights the importance of supporting the personal development of pupils in the classroom.

These two features of the schools in the study may have affected the extent to which the intervention could add value.

Placed in broader perspective, these explanations converge on a single mechanism: the returns to perseverance-oriented interventions on attainment depend on the institutional baseline. Our null result alongside the large effects found by Alan *et al.* (2019) in Istanbul public schools suggests that programmes of this kind face diminishing returns as the institutional environment already moves in the same direction. Yeager *et al.* (2019) reach a similar conclusion in US high schools, finding effects concentrated in schools that provide a receptive environment for growth-oriented messages; England's national curriculum and Ofsted accountability framework, which require schools to foster pupils' personal development (Department for Education, 2013; Ofsted, 2015), may already provide such messages at baseline, limiting the marginal contribution of an additional intervention to academic attainment.

7 Conclusions

We study the effect of an intervention aimed at raising perseverance among primary school pupils in England, a teacher-delivered, eight-week programme evaluated against nationally administered, externally marked test scores. The intervention has no impact on KS2 scores in reading, mathematics, or grammar, punctuation and spelling, and this null result holds across all subgroups including pupils eligible for FSM. The intervention does produce a large shift in pupils' beliefs about the malleability of their intelligence (0.417 SD, approximately 50% larger than the shift reported by Alan *et al.* (2019) in Istanbul primary schools), confirming that the programme was received as intended. The only other positive findings are a borderline significant effect on intrinsic motivation towards learning (0.127 SD) and a significant effect on self-regulation among girls (0.136 SD). These findings point to a context-specific explanation: teachers in both treatment and control schools had some prior familiarity with the ideas underlying the intervention, and England's national curriculum is more broadly designed to foster the personal development of pupils. Taken together, the evidence suggests that the effectiveness of perseverance-oriented programmes on academic attainment depends critically on the institutional baseline, and that in school systems already oriented towards non-cognitive development, the marginal returns to such interventions may be limited.

References

Alan, S., Boneva, T. and Ertac, S. (2019). Ever failed, try again, succeed better: Results from a randomized educational intervention on grit. *The Quarterly Journal of Economics*, **134** (3), 1121–1162.

Almlund, M., Duckworth, A. L., Heckman, J. and Kautz, T. (2011). Personality psychology and economics. In *Handbook of the Economics of Education*, vol. 4, Elsevier, pp. 1–181.

Bettinger, E., Ludvigsen, S., Rege, M., Solli, I. F. and Yeager, D. (2018). Increasing perseverance in math: Evidence from a field experiment in norway. *Journal of Economic Behavior & Organization*, **146**, 1–15.

Blackwell, L. S., Trzesniewski, K. H. and Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child development*, **78** (1), 246–263.

Borghans, L., Golsteyn, B. H., Heckman, J. J. and Humphries, J. E. (2016). What grades and achievement tests measure. *Proceedings of the National Academy of Sciences*, **113** (47), 13354–13359.

Burnette, J. L., Billingsley, J., Banks, G. C., Knouse, L. E., Hoyt, C. L., Pollack, J. M. and Simon, S. (2022). A systematic review and meta-analysis of growth mindset interventions: For whom, how, and why might such interventions work? *Psychological Bulletin*.

Claro, S., Paunesku, D. and Dweck, C. S. (2016). Growth mindset tempers the effects of poverty on academic achievement. *Proceedings of the National Academy of Sciences*, **113** (31), 8664–8668.

Cunha, F., Heckman, J. J. and Schennach, S. M. (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica*, **78** (3), 883–931.

Department for Education (2013). National curriculum in england: primary curriculum. *DFE-00178-2013*.

Department for Education (2015). *KS1 and KS2 test and examination point scores used in the 2015 school performance tables*. Tech. rep., Department for Education.

Duckworth, A. L., Peterson, C., Matthews, M. D. and Kelly, D. R. (2007). Grit: perseverance and passion for long-term goals. *Journal of personality and social psychology*, **92** (6), 1087.

—, Quirk, A., Gallop, R., Hoyle, R. H., Kelly, D. R. and Matthews, M. D. (2019). Cognitive and noncognitive predictors of success. *Proceedings of the National Academy of Sciences*, **116** (47), 23499–23504.

— and Seligman, M. E. P. (2006). Self-discipline gives girls the edge: Gender in self-discipline, grades, and achievement test scores. *Journal of Educational Psychology*, **98** (1), 198–208.

Duncan, T. G. and McKeachie, W. J. (2005). The making of the motivated strategies for learning questionnaire. *Educational psychologist*, **40** (2), 117–128.

Dweck, C. (2006). *Mindset: The New Psychology of Success*. New York: Random House.

Foliano, F., Rolfe, H., Buzzeo, J., Runge, J. and Wilkinson, D. (2019). Changing mindsets: Effectiveness trial. *National Institute of Economic and Social Research*.

Heckman, J. J., Stixrud, J. and Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor economics*, **24** (3), 411–482.

Huillery, E., Bouguen, A., Charpentier, A., Algan, Y. and Chevallier, C. (2025). The role of mindset in education: A large-scale field experiment in disadvantaged schools. *The Economic Journal*, **135**.

Jachimowicz, J. M., Wihler, A., Bailey, E. R. and Galinsky, A. D. (2018). Why grit requires perseverance and passion to positively predict performance. *Proceedings of the National Academy of Sciences*, **115** (40), 9980–9985.

Kautz, T., Heckman, J. J., Diris, R., Ter Weel, B. and Borghans, L. (2014). Fostering and measuring skills: Improving cognitive and non-cognitive skills to promote lifetime success.

Ofsted (2015). *The common inspection framework: education, skills and early years*. Tech. rep., Ofsted.

Paunesku, D., Walton, G. M., Romero, C., Smith, E. N., Yeager, D. S. and Dweck, C. S. (2015). Mind-set interventions are a scalable treatment for academic underachievement. *Psychological science*, **26** (6), 784–793.

Pintrich, P. (1991). A manual for the use of the motivated strategies for learning questionnaire (mslq).

Pintrich, P. R. and De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of educational psychology*, **82** (1), 33.

—, Smith, D. A., Garcia, T. and McKeachie, W. J. (1993). Reliability and predictive validity of the motivated strategies for learning questionnaire (mslq). *Educational and psychological measurement*, **53** (3), 801–813.

Santos, I., Petroska-Beska, V., Carneiro, P., Eskreis-Winkler, L., Munoz Boudet, A. M., Berniell, I., Krekel, C., Arias, O. and Duckworth, A. (2022). *Can Grit Be Taught? Lessons from a Nationwide Field Experiment with Middle-School Students*. IZA Discussion Paper 15588, Institute of Labor Economics (IZA).

Sisk, V. F., Burgoyne, A. P., Sun, J., Butler, J. L. and Macnamara, B. N. (2018). To what extent and under which circumstances are growth mind-sets important to academic achievement? two meta-analyses. *Psychological science*, **29** (4), 549–571.

Yeager, D. S. and Dweck, C. S. (2012). Mindsets that promote resilience: When students believe that personal characteristics can be developed. *Educational psychologist*, **47** (4), 302–314.

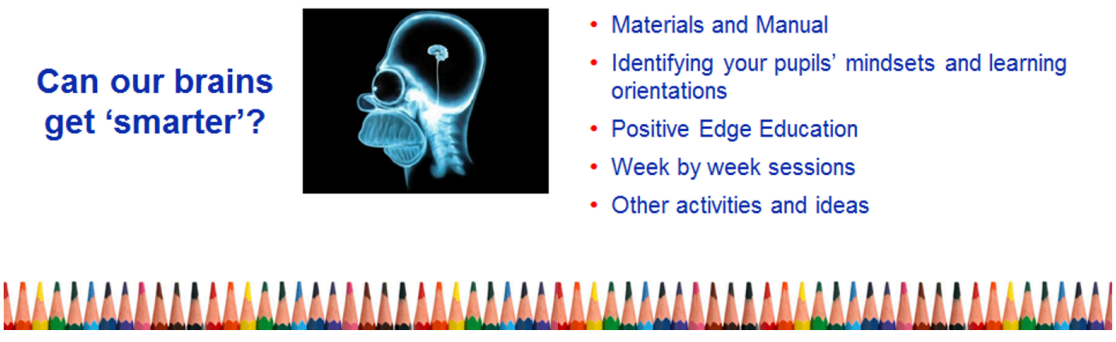
—, Hanselman, P., Walton, G. M., Murray, J. S., Crosnoe, R., Muller, C., Tipton, E., Schneider, B., Hulleman, C. S., Hinojosa, C. P. *et al.* (2019). A national experiment reveals where a growth mindset improves achievement. *Nature*, **573** (7774), 364–369.

—, Romero, C., Paunesku, D., Hulleman, C. S., Schneider, B., Hinojosa, C., Lee, H. Y., O'Brien, J., Flint, K., Roberts, A. *et al.* (2016). Using design thinking to improve psychological interventions: The case of the growth mindset during the transition to high school. *Journal of educational psychology*, **108** (3), 374.

Zimmerman, B. J. and Martinez-Pons, M. (1990). Student differences in self-regulated learning: Relating grade, sex, and giftedness to self-efficacy and strategy use. *Journal of Educational Psychology*, **82** (1), 51–59.

Appendix A Content of the intervention

Sample of slides from the training day



8 Week Pupil Programme

- Materials and Manual
- Identifying your pupils' mindsets and learning orientations
- Positive Edge Education
- Week by week sessions
- Other activities and ideas

Materials and Manual

- **Manual**
  - Overview of academic literature for you and your teachers
  - Mindset questions and the Your Learning Questionnaire (pupil learning orientations with norm data) and scoring instructions for both
  - Lesson plans and materials for each of the weeks described
- **USB stick**
  - Electronic versions of all materials
  - Mindsets questions and Your Learning Questionnaire
  - Powerpoint for your teacher briefing
  - Powerpoint for a parents briefing
- **Positive Edge Education**

Weekly Lesson Plans

1. **Introducing Key Concepts; Exploring Pupils' Mindsets**
2. **Dealing with Mistakes and Emotions**
3. **Understanding the Brain and Encouraging Challenge**
4. **Appreciating the Importance of Effort and Persistence in Improvement and Progression**
5. **Reflecting on the Language we Hear**
6. **Understanding the Impact of Stereotypes (Maths Challenges)**
7. **Unravelling Talent and Luck Myths: The Journeys of Inspirational People**
8. **Consolidating Key Concepts**

Appendix B Mindset measure questions, MSLQ questions and subscales

Scoring of the 'Mindset measure' - Section 1
All items of the 'Mindset measure' are scored from 1–6, with 1 indicating 'Strongly disagree' and 6 indicating 'Strongly agree'.

- You have a certain amount of intelligence, and you really can't do much to change it.
- Your intelligence is something about you that you can't change very much.
- You can learn new things, but you can't really change your basic intelligence.

Scoring of the 'School Questionnaire' - Section 2
All items on the adapted version of the MSLQ are scored from 1–7, with 1 indicating 'Strongly disagree' and 7 indicating 'Strongly agree'. (NB: three items on the self-regulation scale are reverse-coded before constructing the sub-scale score.) Sub-scales (intrinsic value, self-efficacy, test anxiety and self-regulation) are then calculated as a mean score.

Intrinsic value

- 1) I prefer class work that is challenging so I can learn new things.
- 4) It is important for me to learn what is being taught in class.
- 5) I like what I am learning in class.
- 7) I think I will be able to use what I learn in one class in other classes.
- 10) I often choose class and homework activities I will learn something from even if they require more work.
- 14) Even when I do poorly on a test I try to learn from my mistakes.[9]

[9]This item was added to the adapted questionnaire and does not appear in either the intrinsic goal orientation or task value subscales of the original MSLQ (Pintrich and De Groot, 1990; Pintrich *et al.*, 1993). It is excluded from the composite intrinsic value measure used in the analysis; see Section 3.

- 15) I think that what I am learning in class is useful for me to know.
- 17) I think what we are learning in class is interesting.
- 21) Understanding school subjects is important to me.

Self-efficacy

- 2) Compared with other pupils in class, I expect to do well.
- 6) I am certain I can understand the ideas taught in class.
- 8) I expect to do very well in class.
- 9) Compared with others in class, I think I am a good pupil.
- 11) I am sure I can do an excellent job on the problems and tasks given in class.
- 13) I think I will receive a good mark in class.
- 16) My learning skills are excellent compared with other pupils in class.
- 18) Compared with other pupils in class I think I know a great deal about the subject.
- 19) I know that I will be able to learn the information in class.

Test anxiety

- 3) I am so nervous during a test that I cannot remember facts I have learned.
- 12) I have an uneasy, upset feeling when I take a test.
- 20) I worry a great deal about tests.
- 22) When I take a test I think about how poorly I am doing.

Self-regulation

- 23) I ask myself questions to make sure I know the information I have been learning.
- 24) When work is hard I either give up or learn only the easy parts.*

- 25) I do extra work and practice exercises even when I do not have to.
- 26) Even when lessons are dull and uninteresting, I keep working until I finish.
- 27) Before I begin school work, I think about the things I will need to do to learn.
- 28) I often find that I have been reading but I do not know what it is all about.
- 29) I find that when the teacher is talking I think of other things and do not really listen to what is being said.
- 30) When I am reading, I stop once in a while and go over what I have read.
- 31) I work hard to get a good mark even when I don't like a subject.

Appendix C Supplementary descriptives

Table C1: Comparison between schools in the trial and the rest of primary schools in England

	Other schools in England	Trial	difference	(p-value)
Religious school	0.329	0.242	0.087	0.071
Academy	0.180	0.158	0.022	0.573
Community school	0.000	0.000	0.000	0.926
Foundation school	0.000	0.000	0.000	0.948
Voluntary aided	0.076	0.084	-0.008	0.759
Ofsted: Outstanding	0.198	0.158	0.040	0.330
Ofsted: Good	0.667	0.737	-0.070	0.150
Ofsted: Satisfactory	0.109	0.084	0.025	0.434
Ofsted: Inadequate	0.026	0.021	0.005	0.768
Number of pupils	381.685	351.379	30.306	0.381
% of Free School Meal	12.120	15.723	-3.603	0.001
% SEN with support	11.126	12.829	-1.703	0.021
% SEN with statement	2.163	1.363	0.801	0.301
% English Additional Language	11.648	16.555	-4.907	0.005
N	22,118	95		

Table C2: Baseline characteristics by post-trial questionnaire response status

	Respondents	Non-respondents	Difference	(p-value)
Treatment arm	0.509	0.512	−0.004	0.973
Female	0.501	0.439	0.061	0.000
FSM eligible	0.341	0.408	−0.067	0.073
White	0.756	0.560	0.196	0.046
Black	0.052	0.070	−0.019	0.538
KS1 reading points	16.385	16.119	0.266	0.223
KS1 writing points	15.125	14.851	0.274	0.257
KS1 maths points	16.246	15.967	0.279	0.179
Mindset (baseline)	3.865	3.749	0.116	0.256
Intrinsic value (baseline)	5.615	5.532	0.083	0.122
Self-efficacy (baseline)	5.250	5.235	0.015	0.817
Test anxiety (baseline)	3.743	3.842	−0.100	0.286
Self-regulation (baseline)	4.788	4.776	0.012	0.830
N	3,514	1,070		

Note: Respondents are pupils with at least one non-missing post-trial questionnaire response (mindset measure or any MSLQ subscale at wave 2). Non-respondents are pupils missing all wave-2 questionnaire data. Difference = respondents minus non-respondents. p-values from OLS with standard errors clustered at the school level. Baseline non-cognitive measures are wave-1 (pre-trial) scores; missing for pupils who did not complete the pre-trial questionnaire.

Appendix D Additional results

Table D1: Treatment effect on the items of the intrinsic value sub-scale

	mslq1	mslq4	mslq5	mslq7	mslq10	mslq15	mslq17	mslq21
					All			
$T = 1$	0.191*	-0.013	0.042	0.074	0.164	0.115	0.214	0.043
	(0.078)	(0.058)	(0.094)	(0.074)	(0.104)	(0.099)	(0.110)	(0.091)
N	2,887	2,847	2,832	2,842	2,868	2,854	2,860	2,845
R-squared	0.147	0.085	0.153	0.085	0.125	0.128	0.134	0.137
					FSM			
$T = 1$	0.303*	-0.014	0.036	0.082	0.374*	0.133	0.163	0.051
	(0.134)	(0.091)	(0.154)	(0.120)	(0.151)	(0.121)	(0.159)	(0.121)
N	945	924	918	928	937	930	938	928
R-squared	0.134	0.085	0.140	0.081	0.135	0.115	0.153	0.147
					Girls			
$T = 1$	0.322**	0.003	0.049	0.139	0.079	0.181	0.237*	0.020
	(0.102)	(0.061)	(0.104)	(0.078)	(0.133)	(0.114)	(0.116)	(0.094)
N	1,453	1,434	1,424	1,436	1,446	1,450	1,445	1,443
R-squared	0.152	0.094	0.134	0.086	0.100	0.115	0.120	0.127

Note: Column labels are abbreviated; full item wording is in Appendix B. All regressions include pre-treatment measures in the same item, an indicator for gender and for FSM status, a set of dummies for ethnicity and a set of dummies for the randomisation blocks. Standard errors in parentheses are clustered at the school level. All items are scored from 1–7, with 1 indicating 'Strongly disagree' and 7 indicating 'Strongly agree'. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

Appendix E Robustness: Bartlett factor scores

We first verify the unidimensional structure of each scale by estimating a single-factor confirmatory factor analysis (CFA) using Stata's `sem` command, with standard errors clustered at the school level, on the post-trial (wave 2) data. Self-regulation items 24, 28, and 29 are reverse-coded prior to analysis (Appendix B); the mindset scale is exactly identified ($df = 0$) and fit statistics are not reported for it. All standardised factor loadings are positive and significant at $p < 0.001$. Loadings range from 0.58 to 0.76 for mindset, 0.50 to 0.72 for intrinsic value, 0.60 to 0.75 for self-efficacy, and 0.62 to 0.81 for test anxiety. For self-regulation, eight of nine items load in the range 0.43–0.65; one item (mslq28, "I often find that I have been reading but I do not know what it is all about", reverse-coded) has a lower loading of 0.168 but is retained as it belongs to the published MSLQ subscale. SRMR values are 0.051, 0.040, 0.011, and 0.073 for intrinsic value, self-efficacy, test anxiety, and self-regulation respectively, all below the conventional 0.08 threshold, confirming adequate unidimensional structure.

Table E1 replicates the full-sample results from Table 5, replacing the simple-mean composites with Bartlett-weighted factor scores. Bartlett scores weight each item by its contribution to the latent factor from a single-factor analysis on the post-trial items, producing minimum-variance unbiased estimates of the underlying trait. The two aggregation methods are highly correlated ($r > 0.98$ for all subscales); estimates and standard errors are nearly identical across columns, confirming that the choice of aggregation method does not drive any of the results. The intrinsic value coefficient remains positive and borderline significant at the 10% level under both specifications. The slightly smaller N under Bartlett scores reflects listwise deletion: the factor-score prediction requires all scale items to be non-missing.

Table E1: Treatment effects on secondary outcomes: simple mean vs Bartlett factor score (full sample)

	Simple mean (Table 5)					*Bartlett factor score*				
	Mindset	Intrinsic	Self-eff.	Anxiety	Self-reg.	Mindset	Intrinsic	Self-eff.	Anxiety	Self-reg.
$T = 1$	0.417***	0.127†	-0.025	-0.016	0.090	0.417***	0.133†	-0.034	-0.008	0.059
	(0.074)	(0.067)	(0.058)	(0.053)	(0.055)	(0.074)	(0.069)	(0.061)	(0.053)	(0.054)
N	2,902	2,917	2,916	2,916	2,899	2,902	2,768	2,708	2,822	2,746
R^2	0.139	0.265	0.275	0.235	0.244	0.139	0.278	0.283	0.236	0.253

Note: The mindset measure is the simple mean in both specifications (3 items; Bartlett score near-identical). Intrinsic value uses the 8-item composite (mslq14 excluded). Self-regulation items 24, 28, 29 are reverse-coded before factor analysis. All regressions include the pre-treatment measure of the same domain, an indicator for gender and for FSM status, a set of dummies for ethnicity and a set of dummies for the randomisation blocks. Standard errors in parentheses are clustered at the school level. † p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001.

Appendix F Robustness: KS2 attainment in the non-cognitive subsample

Approximately 31 per cent of originally randomised pupils are missing post-trial questionnaire data, raising the question of whether the attainment null holds for the subsample that also produced the observed mindset shift. Table F1 replicates the with-covariate specification from Table 4 restricted to pupils with at least one non-missing post-trial questionnaire response (any item from the mindset measure or the MSLQ at wave 2). This yields approximately 4,200 observations — around 5 per cent fewer than the full attainment sample — and closely mirrors its composition. The estimated treatment effects are −0.024 (SE 0.048), 0.000 (SE 0.040), and 0.022 (SE 0.046) for Maths, Reading and GPS respectively, virtually identical to the full-sample estimates. The attainment null is not an artefact of selecting on pupils who did not engage with the post-trial questionnaire.

Table F1: Treatment effects on KS2 attainment, restricted to pupils with post-trial questionnaire data

	Maths	Reading	GPS
$T = 1$	−0.024	0.000	0.022
	(0.048)	(0.040)	(0.046)
N	4,213	4,202	4,208
R^2	0.446	0.439	0.513

Note: OLS. KS2 outcomes standardised to mean 0 SD 1. Sample: pupils with at least one non-missing post-trial questionnaire response (mindset measure or MSLQ wave 2). Controls: KS1 attainment categories, gender, FSM status, ethnicity dummies, randomisation block dummies. Standard errors in parentheses are clustered at the school level. *** p < 0.001, ** p < 0.01, * p < 0.05.