ImpactEd

Science Stars

Summary impact report

2021/2022

www.impacted.org.uk

Part 1: Introduction

Mission and Introduction 03

Executive Summary 07

Headline Findings 08

Evaluation Design 09

Part 2: Academic Outcomes

Results Summary 10

Part 3: Non-Cognitive Outcomes

Outcome Measures and Design 14

Results summary 15

Part 4: Qualitative Data Analysis

Results Summary 21

Part 5: Summarised Findings

Conclusion 24

Recommendations 25

Part 6: Appendix

Appendix A: Ernest Bevin College 27

Appendix B: Hartsdown Academy 28

Introduction from the Managing Director

I am delighted to provide you with an evaluation report outlining some of the key findings from the evaluation of the Science Stars programme over the last academic year

Our view is that data is only as useful as the purposes for which you use it. But done well, meaningful impact evaluation can help drive meaningful decision making to do more of what is working best, and less of what is not. Hopefully, this report helps provide some summary insights to support that process.

We also want to take this opportunity to say thank you for partnering with us. Please do continue to provide us with your feedback so that we can best develop our way of working with you.

Owen Carter
Co-Founder and Managing Director

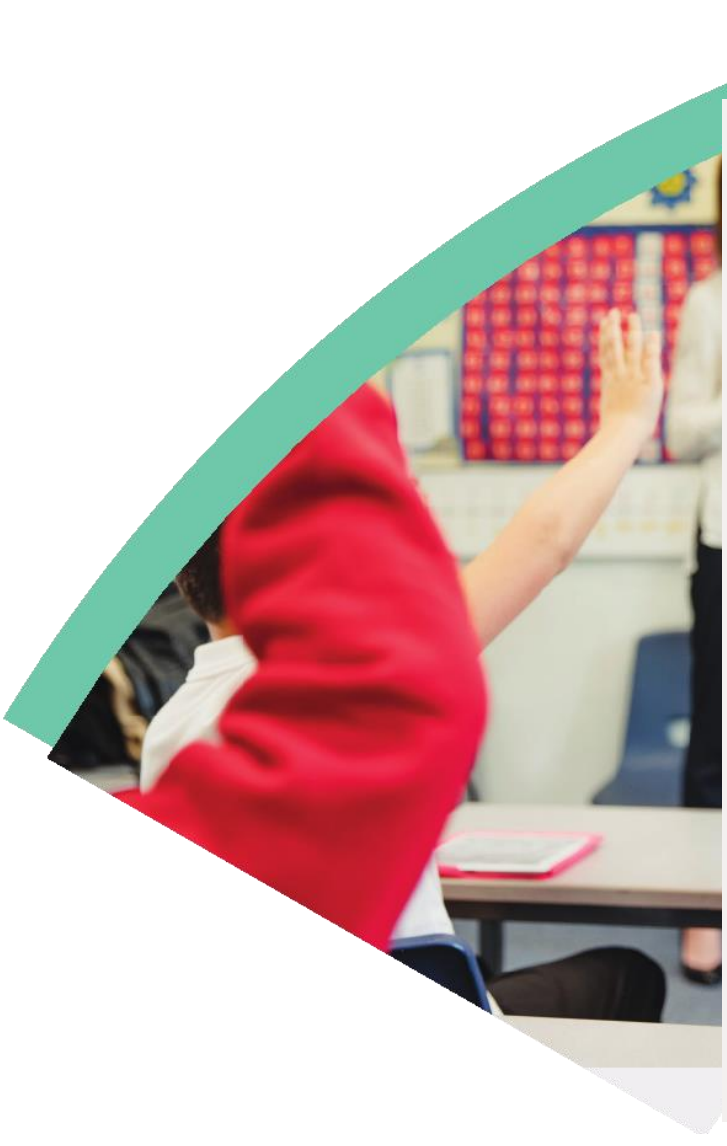
Owen Carter

Our Mission

To improve pupil outcomes and life chances by addressing the evaluation deficit in education

About us

ImpactEd is a non-profit organisation that exists to help evaluate, understand and improve impact in education. We support schools and those that work with them to reliably understand the impact of the programmes they run. We do this through partnership to build capacity for research and evaluation, and our digital platform which makes monitoring and evaluation easy



The Challenge

£4 billion

ANNUAL INTERVENTION SPEND BY ENGLISH SCHOOLS (INC STAFF TIME)

?

FROM 31 SCHOOLS INTERVIEWED, **JUST ONE** WAS CONFIDENT IN THEIR IMPACT EVALUATION

70%

IMPACT EVIDENCE RATED 'POOR' OR 'AVERAGE' BY PAUL HAMLYN FOUNDATION

Why do we exist?

'What works' in education is not an easy question to answer. Schools and education organisations invest substantial time, money and energy in different initiatives and interventions to improve outcomes for pupils. These might range from trialling new teaching and learning approaches, to curriculum redesigns, to mentoring, academic or behaviour interventions.

Through a range of evaluations, we know that some of these changes will be incredibly effective. However, we also know that all too often such programmes can work against their intended aims. And it is often extremely challenging for schools and intervention providers to reliably evaluate the effect that their programmes have on pupil outcomes.

We established ImpactEd to help schools and education organisations understand what is and isn't working in their context, giving them access to robust research methodologies to assess impact, and making evaluation quicker, easier and more effective.

How we work

Platform: Partners use our unique digital platform to make monitoring and evaluation easy to run, accessing reliable research methods for assessing impact on both academic achievement and a range of broader skills.

Partnership: We provide a tailored support and training programme that helps partners identify what it is they are trying to improve, how they are trying to do it, and ways in which they might measure this. Our training and ongoing consultation builds staff capacity for research and evaluation.

Impact: The platform generates live impact reports making it easy to understand what is working, where. We work with partners to discuss findings, informing evidence-based decisions about what is making the biggest difference to pupils, understanding what hasn't, and sharing successes.

About the Organisations

St George's is an independent medical university, affiliated with the University of London. With a strong historical commitment to widening participation activities, St George's is now increasingly working across the whole student lifecycle to support students from under-represented backgrounds. This year, St George's has run the Science Stars programme for the fourth year that focused specifically at school-based activities to raise attainment, the Science Stars programme.

ImpactEd is a not-for-profit organisation that exists to improve pupil outcomes by addressing this evaluation deficit in education. ImpactEd works in partnership across the education sector to support high-quality monitoring and evaluation that informs decisions about what will work most effectively to support students. Their work in access and widening participation has included evaluation projects with University College London, Goldsmiths University and London South Bank University among others.

Programme Overview

Science Stars is a sustained tutoring intervention designed to support Year 11 students to prepare for GCSEs and ultimately increase their attainment in science. The programme is delivered remotely by Student Ambassadors – current students at St George's, University of London – following a pre-designed curriculum developed by a former science teacher.

The programme aims to improve educational outcomes in GCSE Science for target students in Year 11. The key aims and objectives of the programme for participating students are as follows:

- Increase student understanding of the expectations of their GCSE Science examinations on a range of topics
- Provide practice opportunities for students to develop the skills to support them to successfully answer examination questions
- For students to improve their ability to think explicitly about their own learning such as increased self-efficacy, metacognition and reduced test anxiety
- For students to extend their revision repertoire

Executive Summary

This report provides an overview of the evaluation process and findings from the Science Stars tutoring intervention to assess changes in participants' science attainment and non-cognitive outcomes. Attainment and survey data was used to compare the academic and non-cognitive outcomes on Science Stars participants and a matched control group. **Findings are not statistically significant and suggest that the programme has not made a positive impact on the outcomes of Science Stars participants**. Science Stars participants made very similar academic progress to control group pupils, and their **levels of metacognition, self-efficacy and test anxiety decreased over time**, substantially more than the control group pupils. Despite this, **tutors were very positive about their tutoring experience** and thought that pupils had progressed in terms of confidence and exam technique. This is the fourth annual evaluation of the Science Stars programme conducted by ImpactEd. The evaluation process has remined consistent; however, **the programme design has changed significantly**, and it is important to see the findings in this report in terms of this wider context.

The **first two Science Stars programme evaluations had very positive results**. Between 2018 and 2020, the programme was delivered in-person; student tutors travelled to schools to teach sessions after school. In both years, **participants achieved between 0.5 and 1 grade higher than the matched control group pupils**, and results were **statistically significant**, meaning that we can be 95% confident that this difference was due to the intervention rather than by chance. Qualitative data from focus groups with tutors showed that **the programme was highly valued by tutors, teachers and school leadership**. **Parental engagement** through in-person events and check-in phone calls was seen as a **particular strength**.

Disruption caused by the COVID-19 pandemic in the year 2020-21 led to the decision to move the **tutoring programme online**. Tutors delivered sessions remotely to students throughout the period of online learning. Whilst the programme's ability to adapt to the crisis was clearly a strength**, results were less positive than previous years**. Participants' **grades were slightly higher** than the control group pupils' grades, but **findings were not statistically significant**. **Non-cognitive measures of metacognition, self-efficacy and test anxiety decreased**. However, the findings were rightly seen as **positive in the context of the pandemic** and the challenges posed by online learning. (It is important to note that the robustness of the evaluation was compromised because of having to use teacher assessed final grades rather than GCSEs).

This report details the findings from the academic year 2021-22. It shows the impact of an online tutoring programme when it is delivered in-school to pupils following a normal school timetable and sitting normal GCSE exams. Whilst tutors continue to highlight the benefits of the remote delivery model, (most notably ease of access), there are **evident challenges in terms of engagement, behaviour and attendance**. Recent research has shown the varying impact the move to remote learning had on different groups (i.e Pupil Premium and SEND), and it is now widely accepted that **disadvantage is the most associated with less effective learning overall** (Howard, Kahn and Lockyer, 2021). The online model is therefore less likely to help pupils who require the most support and who have the most to gain from a programme like Science Stars. In light of these findings, it is recommended that the **programme's design is carefully considered** and that the programme **returns to its original design or a hybrid model** (consisting of a mix of online and in-person sessions) is adopted.

Headline Findings

The programme had very little impact on academic outcomes:

- Science Stars participants made very similar academic progress compared to control group pupils.
- However, more Science Stars participants achieved their target grade (64%) than control group pupils (61%).

The Science Stars programme did not have a positive impact on two out of the three measured non-cognitive outcomes:

- Metacognition levels decreased for Science Stars participants (-6%) but increased for control group pupils (+3%) across both schools.
- Science Stars participants saw a reduction in their test-anxiety levels, but results were very different across the two schools.
- Self-efficacy decreased very slightly for Science Stars participants (-4%) and increased very slightly for control group pupils (+3%), but results were very different across the two schools.

Tutors had a positive experience overall but faced challenges in terms of pupil engagement:

- Tutors felt well supported on the programme and rewarded by the experience of taking part.
- Tutors saw pupils progress in terms of confidence and exam technique.
- While online sessions are more convenient for tutors, they make tackling engagement and behaviour issues hard.

Evaluation Design

The evaluation had a combined focus. As well as looking at impact on science attainment, this evaluation also paid attention to non-cognitive outcomes with predictive validity i.e. which have been shown to be with associated improvements in long-term outcomes such as well-being, academic achievement, and employment destinations. Alongside academic achievement, there is evidence that these skills - such as metacognition and self-efficacy – can be particularly important in closing disadvantage gaps (Gutmann & Schoon, 2013).

In terms of specific outcomes, we used the following data:

- Science attainment data using school administered mock examinations at two time points
- Final GCSE results
- Pre/post assessment using validated questionnaire measures to measure student attitudes and perceptions for meta-cognition, self-efficacy and test anxiety
- Qualitative data on the experience of student tutors and graduate assistants, collected through focus groups

The evaluation used a control group design to better isolate the impact of the Science Stars programme beyond simply comparing pre-programme and post-programme data.

As pupil selection was conducted by the school and through a voluntary sign-up process, a randomised control group design was not possible. As such, we used a matched control group consisting of students from the same school and year group as the Science Stars participants, to control for prior attainment.

Some important caveats for this evaluation design are worth noting:

- As the control group was not randomised, there may be unobservable characteristics affecting performance beyond prior attainment.
- Particularly when looking at the schools separately, the overall sample size for both participants and the control group is small. As such, results may not be immediately generalisable to other school contexts.
- As programme is operating in very different contexts across the two schools, direct comparisons should not be made between them.

Notwithstanding these points, the design approach allows us to make relatively robust inferences within these constraints by collecting a range of datapoints to triangulate findings and assess if there was a common pattern across indicators.

Part 2: Academic Outcomes

Data for the academic outcomes analysis was provided by Hartsdown Academy and Ernest Bevin College. Baseline data from mock examinations taken in the autumn term (2021) – relatively early in the programme – is compared with interim data from mock examinations taken in the spring term (2022) and final GCSEs (2022).

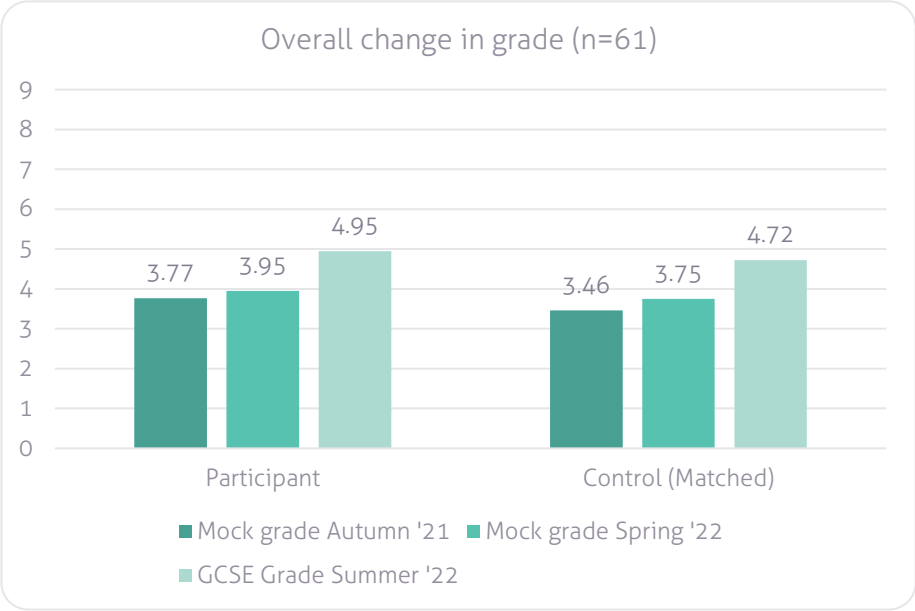
The primary outcome considered in the academic data analysis is relative progress between the three assessment points for Science Stars participants compared to the (matched) control group. The overall sample size includes 61 pupils; 23 participants and 38 (matched) control group pupils.

The **difference-in-difference method** was used to identify whether changes between the baseline and final points for participating and control group pupils were statistically significant. This analysis was undertaken for academic attainment and non-cognitive skills data. Descriptive analysis was undertaken to understand overall trends. We used parametric methods on the data to see if findings were statistically significant – that is, to see if we could rule out the possibility that any observed difference between the participating and control group was down to chance.

Results Summary

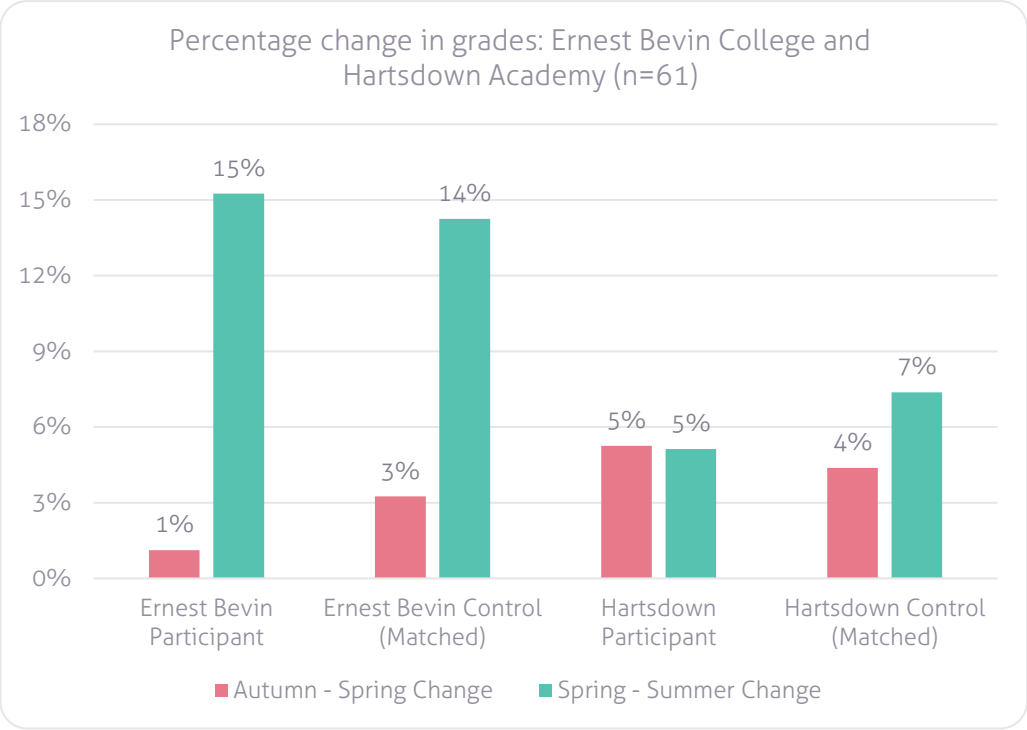
Key finding 1: Science Stars participants made very similar academic progress compared to control group pupils.

When comparing the progress made by Science Stars participants and control group pupils between mock examination taken in the autumn term of 2021, and GCSE examinations taken in the summer term of 2022, we see very little difference between the two groups. Science Stars participants' grades increased by an average of +15%, whilst control group pupils' grades increased by an average of +16%. This change was not statistically significant with a p-value of 0.9.

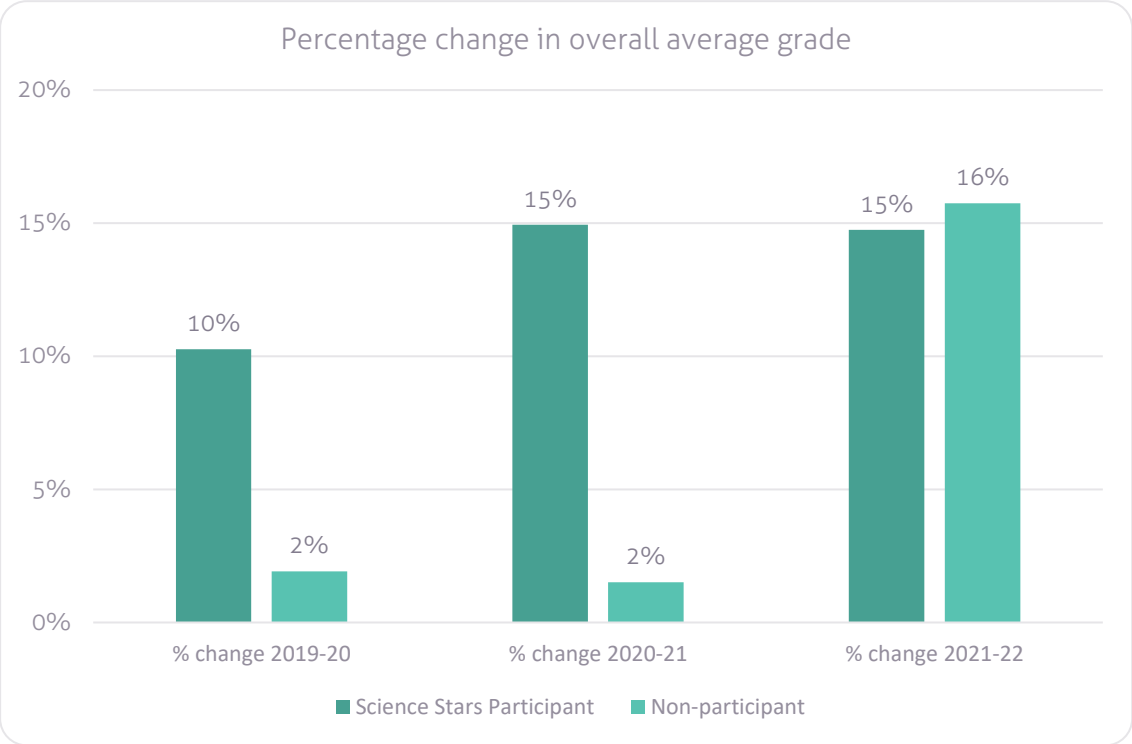


The same trend is seen when we look at each school separately. The participant group at Ernest Bevin College saw their grades increase by +16%, whilst control group pupils' grades increased by +18%. Similarly, at Hartsdown Academy, Science Stars participants saw their grades increase by an average of +10%, whilst control group pupils' grades increased by an average of +12%. Neither change was statistically significant. (See appendix A and B for a full breakdown of the results).

The graph below shows the percentage change in progress between the autumn and spring terms, and the spring and summer terms. We can see that pupils at Ernest Bevin College made the most progress between the spring and summer terms. However, the rate of progress remained relatively stable for students at Hartsdown Academy. In both cases, differences between the participants and control group pupils were not statistically significant.



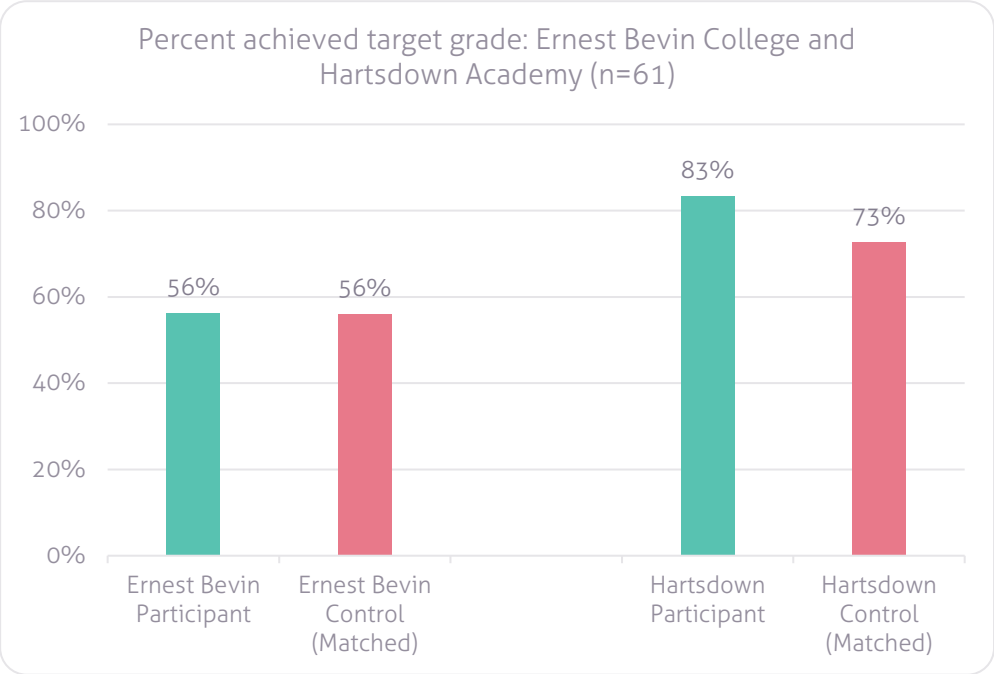
The graph below shows the percentage change in grades of the participant and control group pupils for the past three years. In 2019-20, participants made +8% more progress than control group pupils and in 2020-21 they made +13% more progress than control group pupils. However, in 2021-22, participants made -1% less progress than control group pupils. This suggests that the programme made less of an impact on the grades of participants this year compared to previous years.



Key finding 2: More Science Stars participants achieved their target grade (64%) than control group pupils (61%).

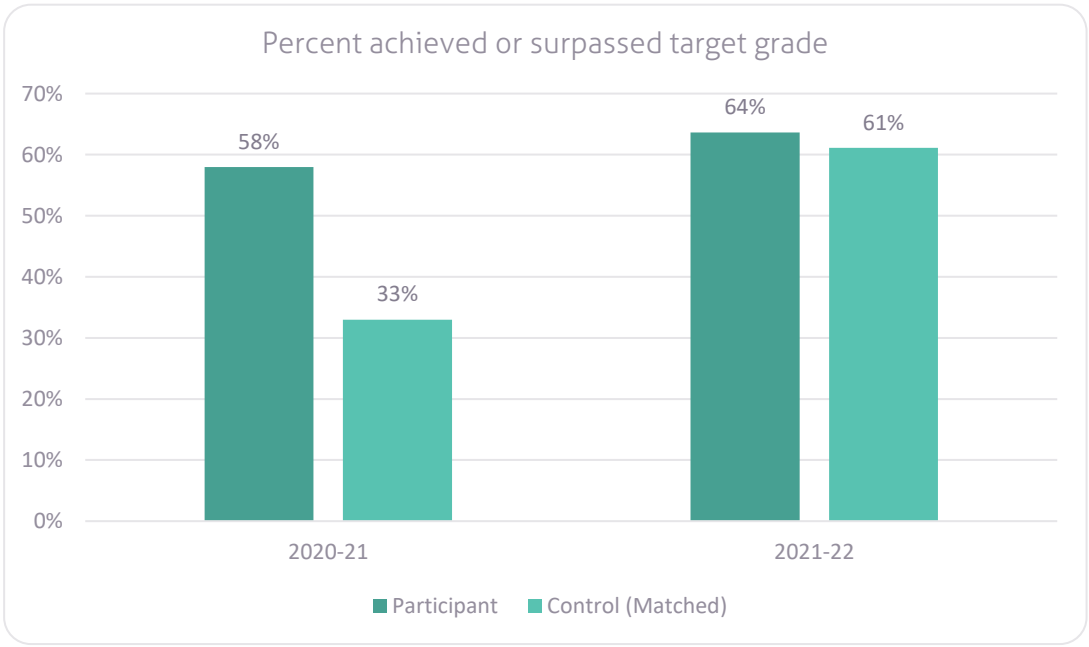
In both schools, participating and control group pupils were provided with a guided target grade. 3% more Science Stars participants achieved their target grade than control group pupils.

The graph below illustrates the difference between the two schools. At Ernest Bevin College, the same proportion of Science Stars participants and control group pupils achieved their target grade. In contrast, at Hartsdown Academy, +10% more Science Stars participants achieved their target grade than control group pupils.



Target grades use end of KS2 assessments to predict what pupils 'should' or 'are likely to' achieve based on the performance data of pupils with similar starting points in previous years. Whilst they underestimate the variation in pupil trajectories, they are useful when thinking about the progress made of large groups or cohorts. It can be said here that more Science Stars participants at Hartsdown Academy made the expected rate of progress than their control group peers.

When we look at how many achieved or surpassed their target grade compared to previous years, we can see that the difference between control group and participant pupils this year is far less than in previous years. The graph below shows that in 2020-21, +25% more participant pupils achieved or surpassed their target grade, a far bigger difference than in 2021-22. This reflects the previous findings which showed that the programme had less of an impact on pupils' grades this year than in previous years.



Part 3: Non-Cognitive Outcomes

Outcome Measures and Design

As well as looking at impact on science attainment, this evaluation also examined non-cognitive outcomes with predictive validity i.e. which have been shown to be with associated improvements in long-term outcomes such as well-being, academic achievement, and employment destinations. Alongside academic achievement, there is evidence that these skills - such as metacognition and self-efficacy – can be particularly important in closing disadvantage gaps (Gutmann & Schoon, 2013).

These non-cognitive outcomes were measured using psychometrically validated questionnaires, administered to pupils pre and post Science Stars. The evaluation followed a pre-post-test design. Pupils were assessed at the beginning (baseline collection) and end (final collection) of the programme. Collecting data at these two time points allows us to analyse the level of change over the course of the programme for each specific outcome.

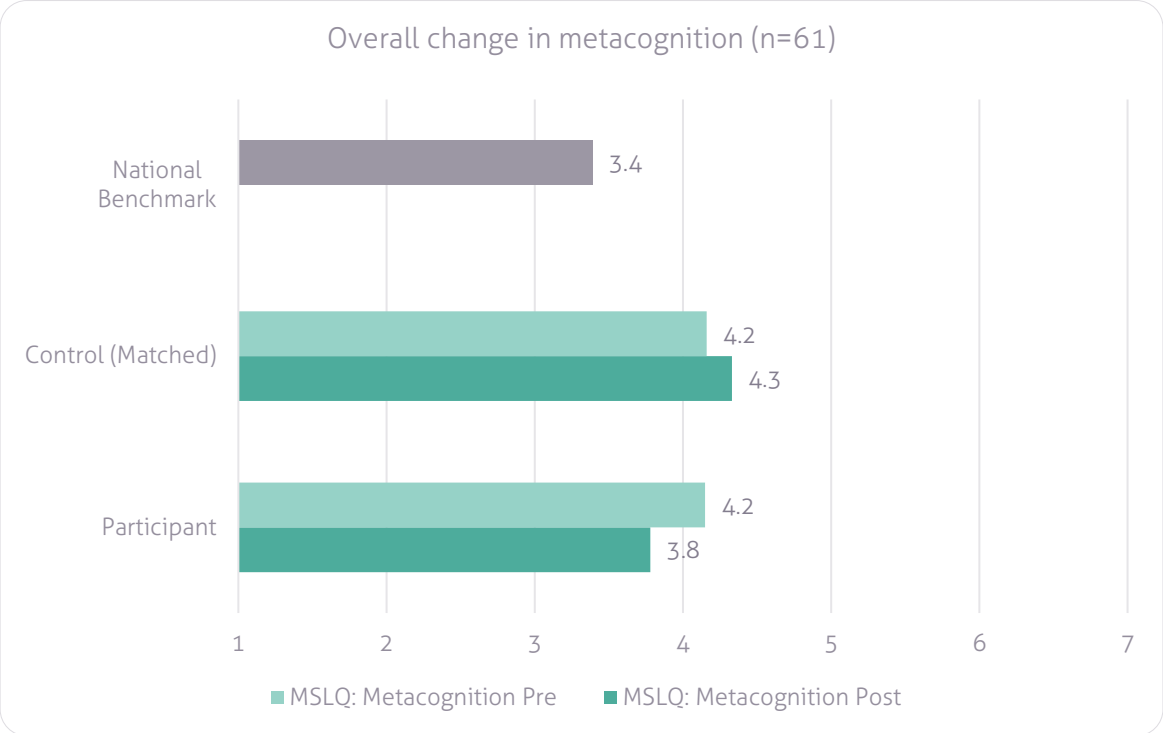
Our core outcome measures for this evaluation were:

Outcome	Measurement details
Metacognition	Metacognition means 'thinking about thinking': pupils' ability to think explicitly about their own learning. It is strongly associated with academic progress and improves other skills required for learning, such as critical thinking (Flavell, 1979; Higgins et al., 2016). We measured metacognition using the Cognitive Strategies Use and Self-Regulation subscales of the Motivated Strategies for Learning Questionnaire.
Self-efficacy	Self-efficacy is a measure of pupils' belief in their ability to achieve a specific task in the future. Self-efficacy is correlated with higher academic achievement and persistence, and also contributes to pupil wellbeing (Gutman & Schoon 2013, DeWitz et. al. 2009). We measured self-efficacy using the Self-efficacy subscale of the Motivated Strategies for Learning Questionnaire.
Test Anxiety	Test anxiety is concerned with pupils' emotional responses to tests (Pintrich and De Groot, 1990). Greater levels of test anxiety can result in worse performance in exams but in some situations may be linked to increased motivation.

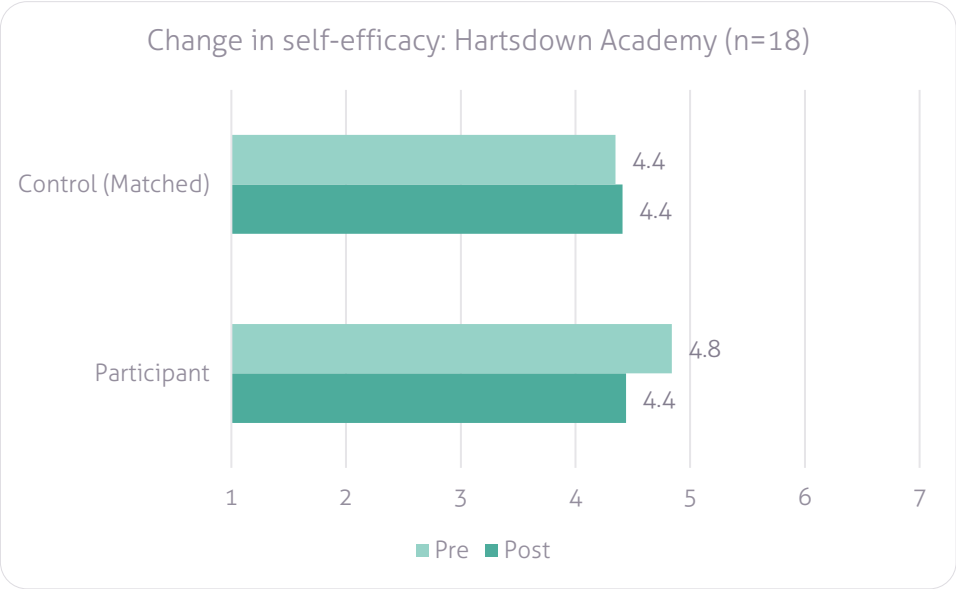
Results Summary

Key finding 3: Metacognition levels decreased for Science Stars participants (-6%) but increased for control group pupils (+3%) across both schools.

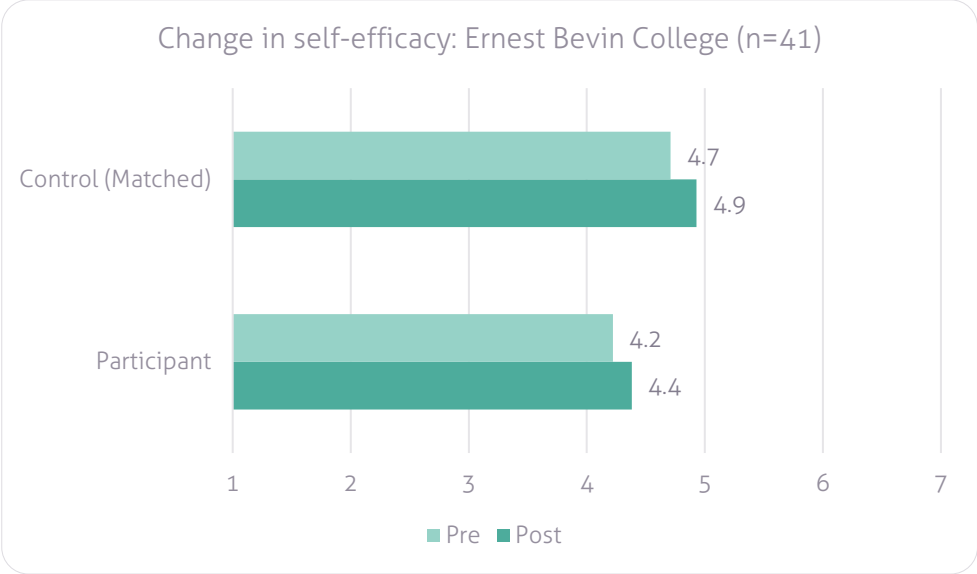
Metacognition saw the largest decrease between baseline and final points out of all the non-cognitive measures. The non-cognitive skills surveys administered at the start and end of the programme found that participating pupils' metacognition levels decreased by -6%, whilst control group pupils increased by +3%. This change was not statistically significant with a p-value of 0.07. However, upon comparing against the national benchmark for metacognition, Science Stars participants have higher than average scores suggesting that they have better than average skills when it comes to thinking about their own learning. This is illustrated in the graph below.



When we look at the schools separately, we see the same trend. Science Stars participants at Hartsdown Academy saw a relatively large decrease in their metacognition scores (-11%), while the control group pupils increased only very slightly (<1%). This change is not statistically significant with a p-value of 0.23.

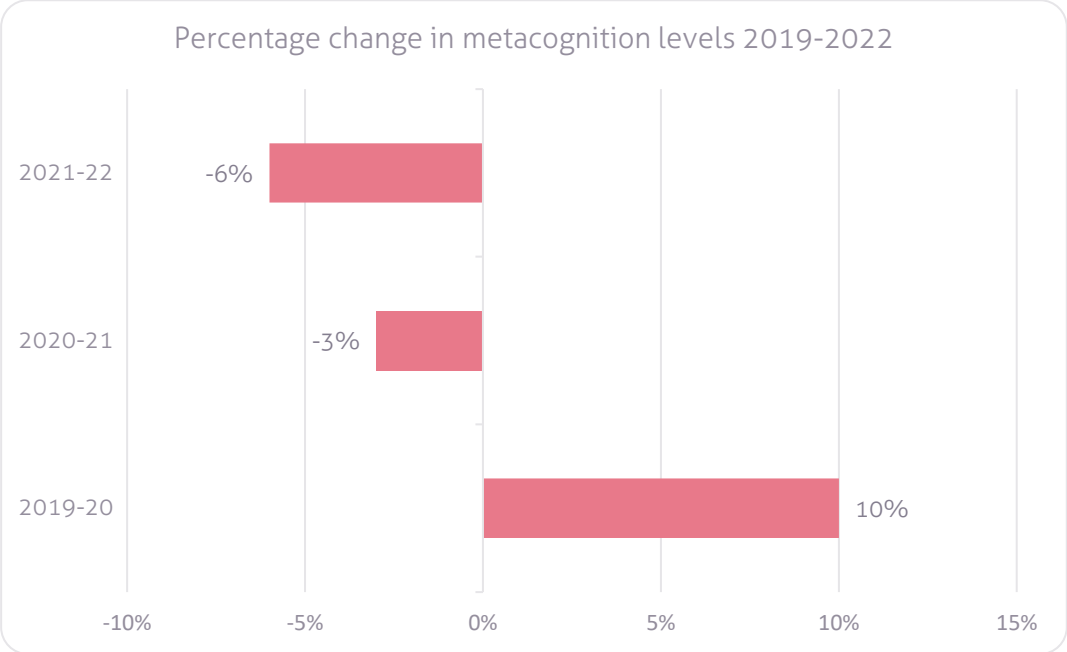


Science Stars participants at Ernest Bevin College saw a -4% decrease in their metacognition scores, while the control group pupils saw a +5% increase. This change is not statistically significant with a p-value of 0.19.



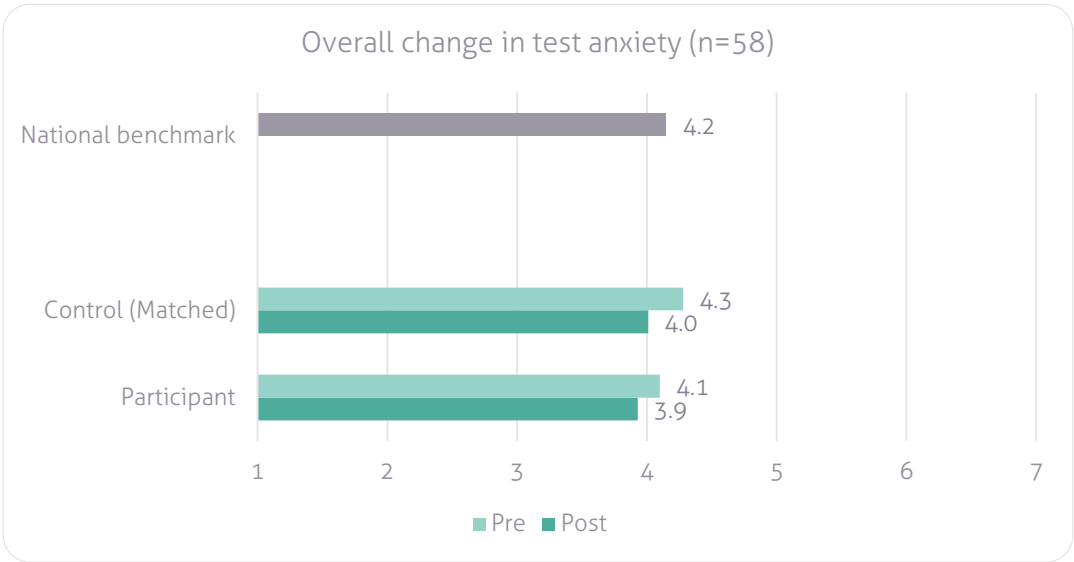
It should be noted that last years' Science Starts participants at Ernest Bevin College (2020-21) also saw a decrease of -3% in their metacognition levels. These changes are relatively small and could be caused by a greater awareness or understanding of revision techniques which makes pupils think more critically of their own learning.

The graph below shows that the % change in participants' metacognition levels has decreased over time. In 2019-20, the metacognition levels of participants increased by +10% over the course of the programme. This dropped to a change of -3% in 2020-21, and dropped again to -6% in 2021-22.



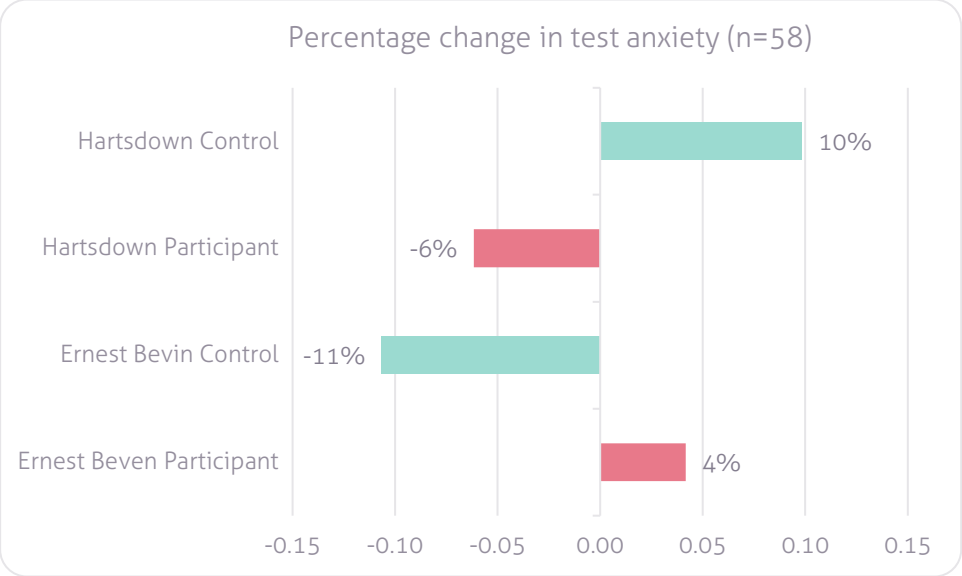
Key finding 4: Science Stars participants saw a reduction in their test-anxiety levels, but results were very different across the two schools.

The MSLQ Test Anxiety scale measures pupil anxiety using various statements such as 'I worry a great deal about tests'. The scale is out of 7, 1 being 'Not at all true of me', and 7 being 'Very true of me'. A reduction in Test-Anxiety levels is therefore positive and, overall, Science Stars participants saw a decrease of -3% in their Test Anxiety levels. This is slightly less than the control group who saw a decrease of -5%. This change was not statistically significant with a p-value of 0.57.

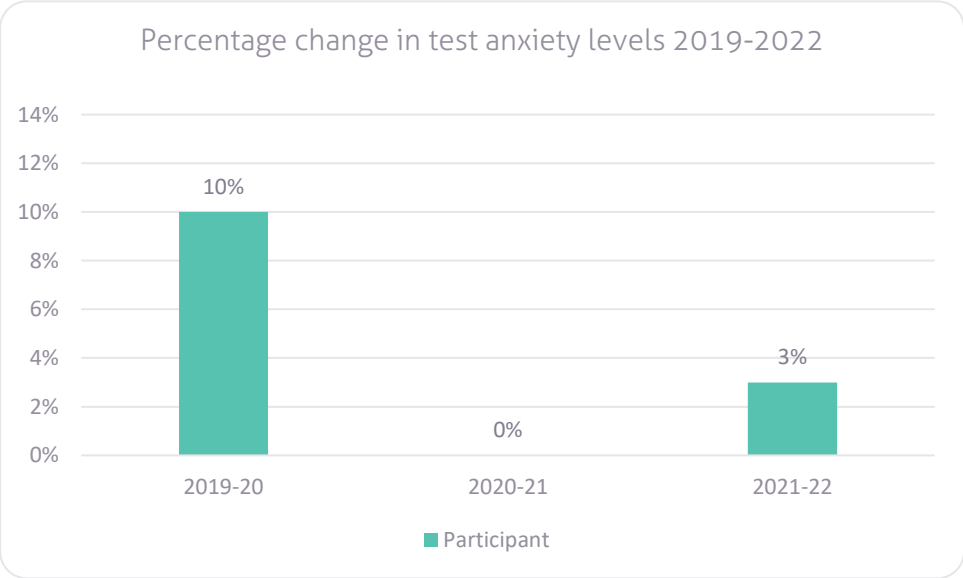


Interestingly, the control group participants Test Anxiety levels stayed very close to the national average in the baseline and final points. The participant group however, started with anxiety levels slightly lower than the national average, and ended the programme -5% below the national average.

When the two schools are looked at separately however, a slightly different picture emerges. Ernest Bevin College saw an increase of +4% in the anxiety levels of Science Stars participants, and a decrease of -11% in the control group levels. In contrast, Hartsdown Academy saw a decrease of -6% in the Science Stars participants anxiety levels, and an increase of +10% in the control group levels. Neither was statistically significant with p-values of 0.4 and 0.78 respectively. The difference between the two schools is illustrated in the graph below.



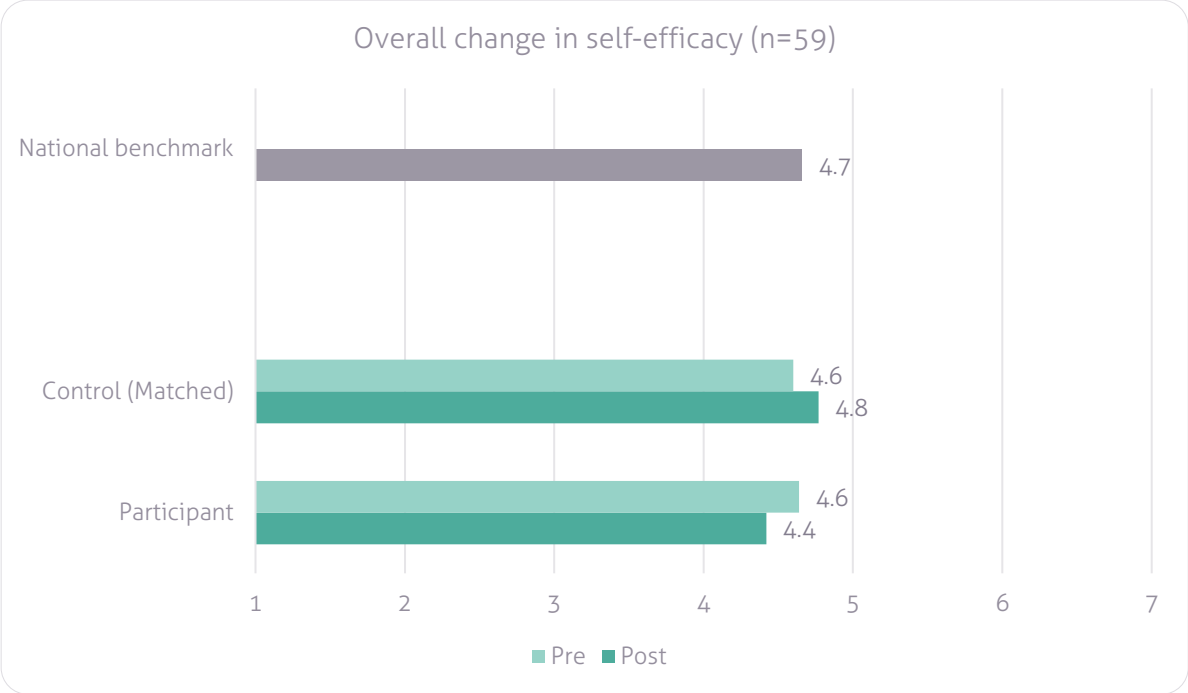
The graph below shows how the impact of the programme on the anxiety of participants has changed over time. In 2019-20, participants felt on average 10% less anxious than at the start of the programme. In 2020-21, there was no change in anxiety levels, and in 2021-22, participants felt 3% less anxious than at the start of the programme. In line with previous findings, this suggests that the programme in its current form, is not having as big an impact on the non-cognitive outcomes of participants as it was previously.



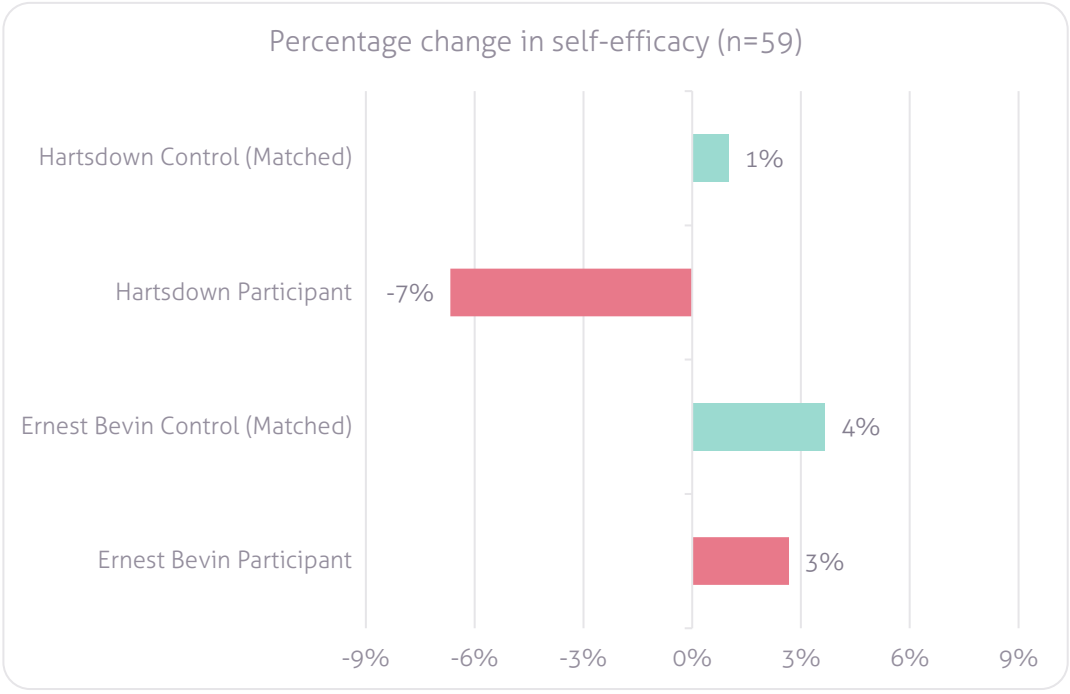
However, the difference this year between the two schools suggests that anxiety levels are linked to external circumstances. 2020-21 was the year of the pandemic, so we would expect that cohort to be particularly anxious and for the programme to make less of an impact in that respect. With the return to normal schooling this year, we can see that participants are slightly less anxious than during the pandemic, but more anxious than they were pre-pandemic.

Key Finding 5: Self-efficacy decreased very slightly for Science Stars participants (-4%) and increased very slightly for control group pupils (+3%), but results were very different across the two schools.

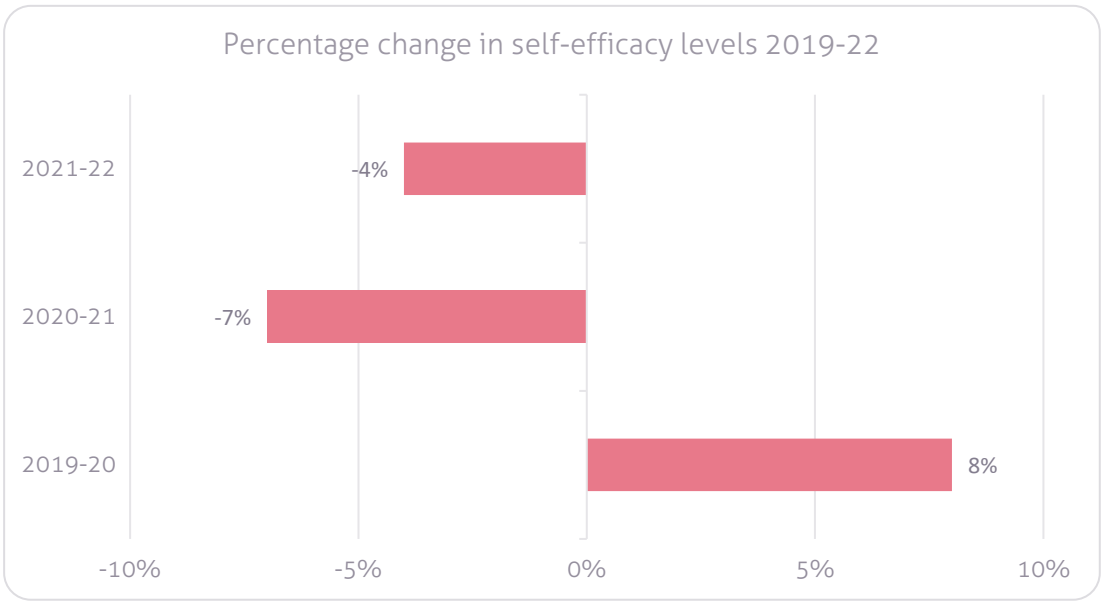
Overall, self-efficacy levels remained relatively stable for Science Stars participants and control group pupils between the baseline and final surveys. Science Stars participants levels decreased by -4%, dropping from an average of 4.6 to 4.4. Control group pupils' self-efficacy levels increased by +3%, rising from an average of 4.6 to 4.8. However, as the graph below illustrates, both groups' scores are very close to the national average and appear to have remained relatively stable over the course of the year. This change was not statistically significant with a p-value of 0.32.



As was the case for test anxiety levels, when we look at the two schools separately a slightly different picture emerges. Science Stars participants at Hartsdown Academy saw a relatively large drop in their self-efficacy levels (-7%), whilst the control group pupils at Hartsdown remained stable (+1%). In contrast, at Ernest Bevin College, the self-efficacy levels of both the Science Stars participants and control group pupils rose slightly, by +3% and +4% respectively. The difference between the schools is shown in the graph below. Neither change at Ernest Bevin College or Hartsdown Academy was statistically significant with p-values of 0.22 and 0.88 respectively.



The graph below illustrates how the impact of the programme on the self-efficacy levels of participants has changed over time. We can see that in 2019-20, when the programme was delivered in-person, it made a positive impact of +8% on the self-efficacy levels of participants. This contrasts with the self-efficacy levels of the 2020-21 and 2021-22 cohorts, whose self-efficacy levels decreased by -7% and -4% respectively. These findings suggest that when the programme is delivered remotely, it does not have a positive impact on self-efficacy.



Part 4: Qualitative Data Analysis

In addition to survey and attainment data collected from pupils and schools, Science Stars tutors and group assistants were each invited to participate in a post-programme focus group to understand the implementation factors of the programme and overall pupil experience. The focus groups inquired about participants' perceptions of programme delivery and design, their roles and pupils' engagement to understand the why and how behind the data captured in the previous sections.

The following section summarises the key themes that have emerged from the two focus groups conducted at the end of the programme.

Results Summary

Theme 1: Tutors felt well supported on the programme and rewarded by the experience of taking part.

All tutors involved in the focus group said that they found the training useful. The midway training was seen as particularly beneficial because it was participatory and allowed the tutors to raise issues they were facing. Tutors liked having scenarios to talk through because it enabled them to bring in their own personal experiences to the training and hear ideas from others about how to overcome problems.

All tutors in the focus group spoke highly of Marlene and her team, and the support they had offered them. The tutors said that Marlene was excellent at communicating, offering 1-1 support, contacting schools on their behalf and providing them with the resources they needed. The tutors felt prepared for the sessions and knew that Marlene would help to tackle problems if they arose.

Tutors found tutoring a rewarding experience. They were motivated by the desire to help disadvantaged pupils. They spoke about the joy of seeing pupils progress, and they particularly enjoyed the graduation ceremony at the end.

> When I first met my students, they would not want to ask questions, they struggled to engage, and they'd get confused really easily. However, we got to a point where they knew so much more than I ever expected. It's so rewarding to see that. That's why I wanted to be part of it again.
>
> *- Science Stars Tutor*

Theme 2: Tutors saw pupils progress in terms of confidence and exam technique.

The engagement of participants increased over the course of the programme. Tutors described how, despite concerns about engagement levels at the start of the programme, by the end, participants were confident, engaged and supporting each other; 'At the beginning, my group are very quiet. It took a very long time to for them to answer the exam questions. But towards the end, they were quick with it. They were asking questions or even helping each other when they got stuck'.

Tutors also noticed a marked difference in pupils' exam technique. They spoke about how participants learnt to deconstruct exam questions and use metacognitive techniques to decide how best to approach the answer.

> At the start, they would see an exam question and just think that it looked really complicated. Whereas at the end, they would break the question down by identifying key information, and what they needed to do to answer it. Not only could you see an improvement in what they scored in the tests, you could also see an improvement in their thought process.
>
> *- Science Stars Tutor*

Tutors highlighted issues to do with attendance and pupil turnover, which they said negatively impacted pupil progress. They liked it when they kept the same class for the full length of the programme, and they could guarantee good attendance because the sessions didn't clash with anything else in the school day.

Theme 3: While online sessions are more convenient for tutors, they make tackling engagement and behaviour issues hard.

> Because they weren't in a controlled classroom situation, it made it easier for them to pretend that they'd understood... I'd ask them something and they wouldn't know, and I didn't know if it was because they hadn't been paying attention or if it was because the way I was explaining it made it harder for them to understand.
>
> *- Science Stars Tutor*

Tutors liked the convenience of running the sessions remotely. On the whole, the technology worked fine, and teachers were quick to respond when there was a problem. Sessions were the most successful when teachers set up the session in school and stayed to help pupils who had problems accessing the resources.

The downside of the remote model was that it made it harder for tutors to form positive relationships with tutees. Pupils were often shy at the start and reluctant to use the microphone. Waiting for pupils to type answers in the chat-box slowed the pace of the session and made it harder to get through content.

School policy that pupils cannot not use cameras made it hard for tutors to know if pupils were actively engaged. One described how they 'ended up talking to a blank screen for the whole of the session ... You didn't know whether it was silence because they were confused or

silence because they weren't listening or messing around in the room that they were in... That problem was quite difficult to overcome'.

Tutors were, however, able to overcome some of the barriers posed by online learning. One described how after a few sessions, pupils 'realised that using the microphone would benefit them because it would make more time for actual content and learning', going on to describe how 'gradually (the tutees) began to use the microphone and became much better at answering questions ... They even began to ask me (the tutor) questions through the microphone'. As the tutors and tutees became more confident teaching and learning online, they found ways to communicate more effectively.

Part 5: Summarised Findings

Conclusions

The analysis of the mock examination and GCSE results of Science Stars participants and control group pupils suggests that **both groups made relatively similar progress, despite the intervention**. The differences-in-differences method showed that the **differences identified between the participants and control group grades were not statistically significant**. Despite this, **Science Stars participants made positive progress on average**, and 64% achieved their target grade (slightly higher than the control group pupils of whom 61% achieved their target grade).

The analysis of non-cognitive outcomes suggests that the **programme had a negative impact on the pupils' self-efficacy and metacognition** levels. All changes were **not statistically significant.** However, when looked at separately, **the impact of the programme was very different in each of the schools** (see appendix A and B for more details). This is something which should be explored in further evaluations, particularly if more schools become involved.

The qualitative data gathered from in-depth tutor and group assistant focus groups highlights the **merits of the programme in offering engaging science content to participants**. Tutors and group assistants **felt well supported** by Marlene and her team. Whilst they **faced some difficulties in terms of engagement**, overall, they found the **programme a rewarding experience** and felt that they had **made a positive difference** to their tutees' science GCSEs.

Previous evaluations have shown that the **programme had a significant positive impact in previous years**. The delivery model has changed significantly as a result of the pandemic, and it is worth exploring the **option of returning to the original model** now that the pandemic is over. Whilst there are **clearly benefits of remote delivery** in terms of **flexibility and accessibility** for tutors, it appears that it is at the **cost of academic and non-cognitive outcomes** for pupils. The online model means that **tutors do not know when pupils are not engaging and find it hard to assess their understanding.** Future evaluations should further **investigate the differential impact of online and face-to-face delivery models** in order to refine the programme delivery model to meet pupil needs.

Recommendations

Programme Recommendations:

Recommendation 1: Consider returning to an in-person tutoring model in order to increase pupil engagement, improve relationships between tutors and pupils and enhance tutor understanding of pupil progress. Alternatively, explore the possibility of a **hybrid model** and evaluate the impact of online and in-person tutoring separately (see evaluation recommendations below).

Recommendation 2: If the online tutoring model remains, train tutors in how to set clear expectations for pupils regarding online learning and ask schools to set the same expectations. Such expectations should be around communication, engagement and participation. The possibility of **using cameras** during the sessions should be explored with the senior leadership teams in the schools.

Recommendation 2: Ensure tutors are aware of schools' behaviour policies and know how to enforce them. Establish clear communication channels between tutors and teachers so that issues to do with behaviour and attendance can be delt with quickly during sessions.

Recommendation 3: Create more opportunities in tutoring sessions for formative assessment so that tutors gain a better understanding of pupils' gaps in knowledge and understanding. This may take the form of low-stakes quizzes, polls, entry and exit slips or short answer questions. Ideally, each session would begin and end with an assessment-for-learning task.

Recommendation 4: Provide more opportunities for pupils to collaborate, share ideas and teach each other. This may help to build confidence, improve self-efficacy and develop metacognitive skills.

Recommendation 5: Hold in-person induction events or tutoring sessions at the start of the course to allow tutors to build positive relationships with pupils.

Recommendation 6: Consider offering more opportunities for tutors to share ideas and build peer support networks. A similar session to the 'midway training' could be offered earlier on.

Recommendation 7: Consider engaging parents again through in-person events and 'check-in' phone calls. This would help to boost pupil engagement and hopefully therefore academic attainment.

Recommendations for future evaluations:

Recommendation 1: Explore the possibility of evaluating the impact of in-person and online tutoring sessions separately. For example, the programme could be run online in one school, and in-person in another. This would make it possible to compare the impact of both models in the evaluation.

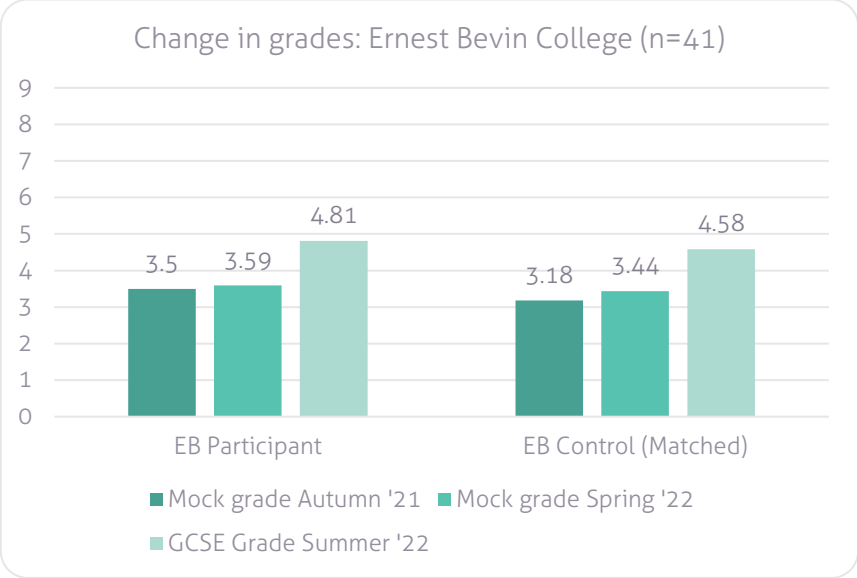
Recommendation 2: Collect qualitative data on pupils' experience of the programme. It would be helpful to understand how easy it is for pupils to engage with online sessions.

Recommendation 3: Collect additional demographic data on Pupil Premium eligibility, gender and ethnicity so that we can decipher the differential impact of the programme. This is particularly important in light of recent research showing the differential impact online learning has on disadvantaged groups (Howard, Kahn and Lockyer, 2021).

Appendix A: Ernest Bevin College

Academic outcomes

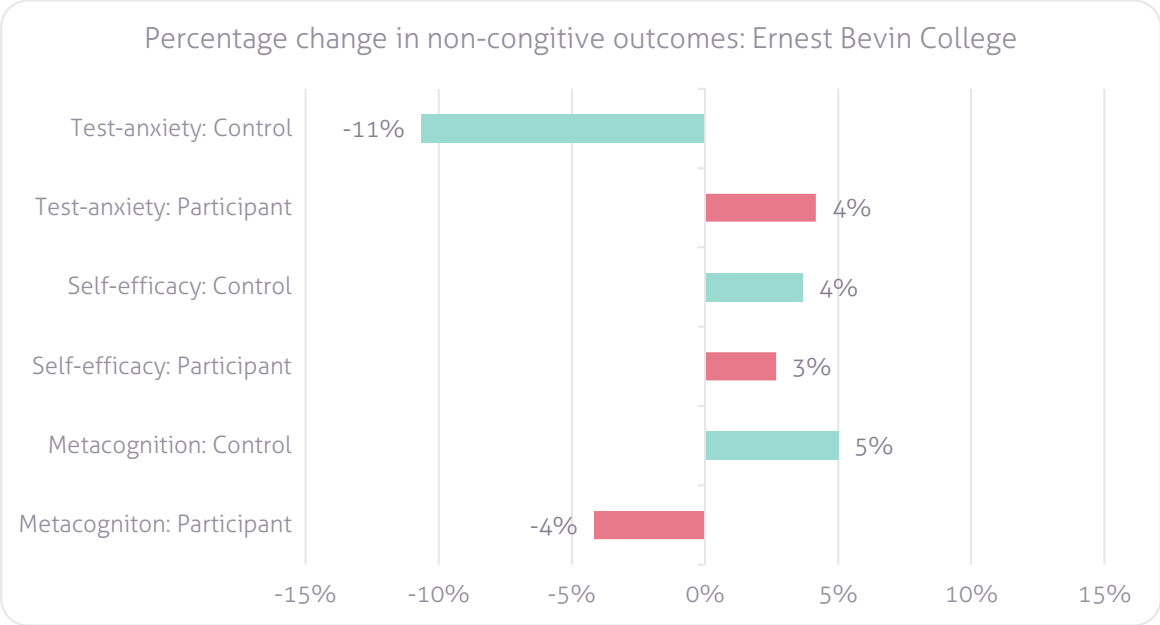
Science Stars participants' grades increased by +16% between the autumn 2021 mock exams and the summer 2022 GCSE exams, whilst control group pupils increased by +18%. Both groups made the greatest gains in terms of progress in the summer term.



Non-cognitive outcomes

The only 'positive' gains made in non-academic outcomes by Science Stars participants at Ernest Bevin College were in their self-efficacy levels, which increased by +3% (in line with the control group pupils who increased by +4%). The biggest difference between the participants and the control group pupils can be seen in their anxiety levels, where the control group pupils' anxiety decreased by -11%, whilst the participants increased by +4%. Metacognition levels for both control and participant pupils stayed relatively close to the national average and saw small changes of +5% and -4% respectively. None of these changes were statistically significant.

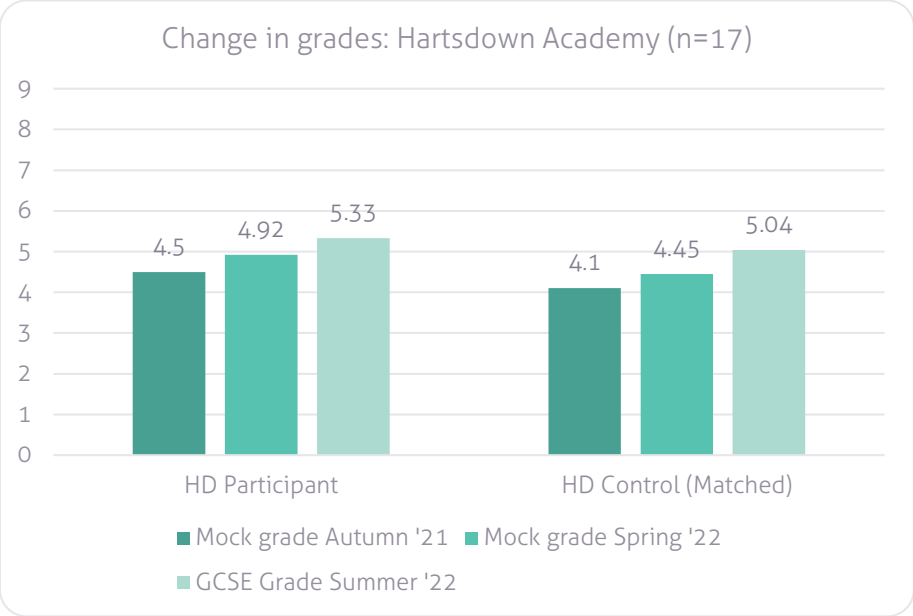
The differences between the percentage changes in the non-cognitive outcomes can be seen in the graph below.



Appendix B: Hartsdown Academy

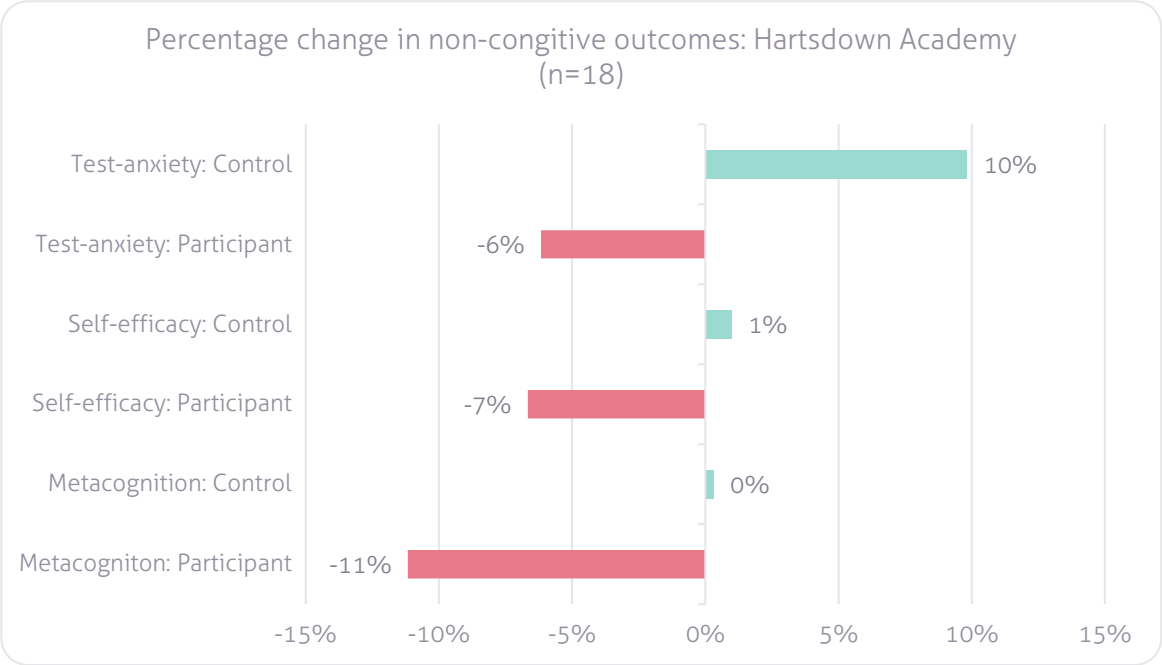
Academic outcomes

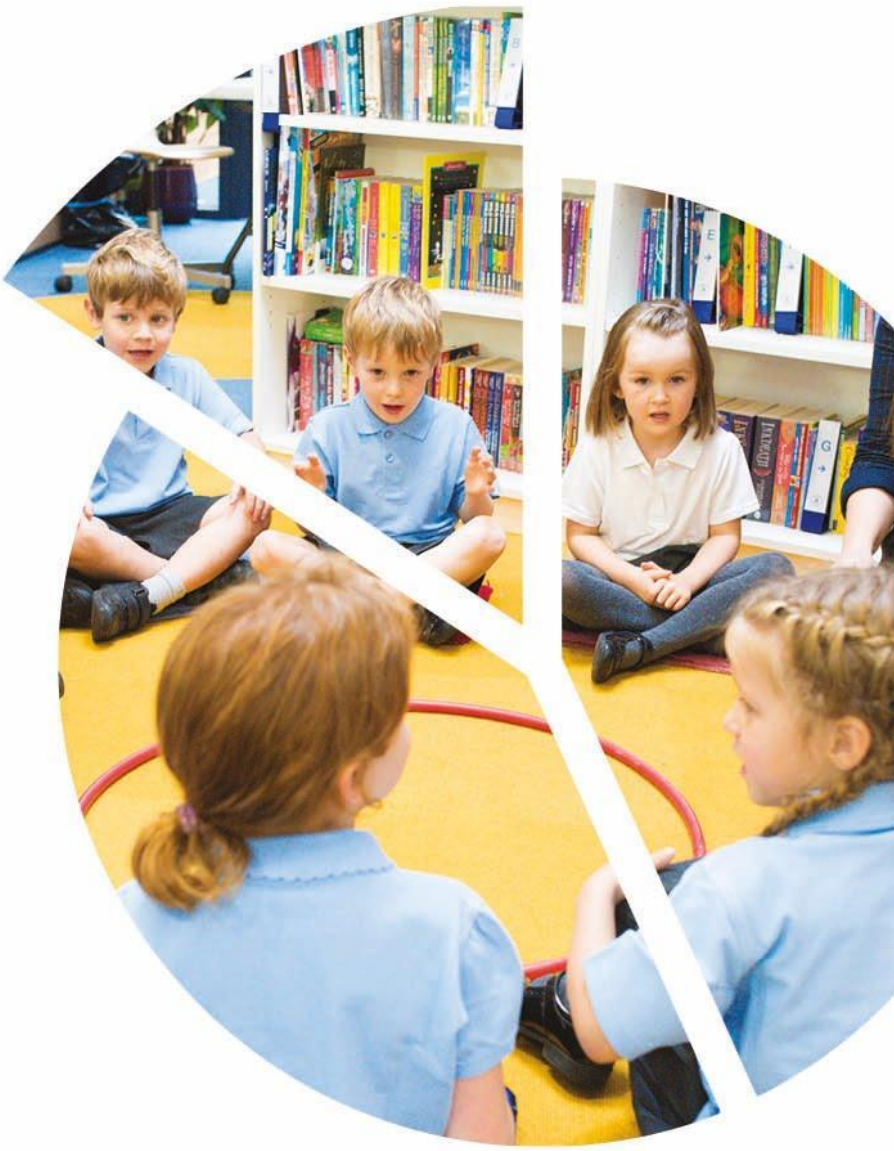
Science Stars participants' grades increased by +10% between the autumn 2021 mock exams and the summer 2022 GCSE exams, whilst control group pupils increased by +12%. Both groups made relatively steady progress over time.



Non-cognitive outcomes

The only 'positive' gains made in non-academic outcomes by Science Stars participants Hartsdown Academy were in their test-anxiety levels which dropped by -6%. This was a big change compared to the control group, for whom anxiety levels increased on average by +10%. In contrast, self-efficacy and metacognition levels remained relatively stable for control group pupils but decreased by -7% and -11% respectively for Science Stars participants. None of these changes were statistically significant. The differences between the percentage changes in the non-cognitive outcomes can be seen in the graph below.





ImpactEd is transforming how schools approach their programmes, embedding an impact culture across the education system."

DAME SUE JOHN, Executive Director, Challenge Partners

Partners and supporters

United Learning – The best in everyone™

THE YOUNG ACADEMY

CHALLENGE PARTNERS

SUPER BEING

nesta

TeachFirst INNOVATION PARTNER 2018/19

LEEDS BECKETT UNIVERSITY

MANCHESTERTSA Teaching School Alliance

Improving pupil outcomes by working with schools to address the evaluation deficit.

Get in touch
hello@impacted.org.uk

ImpactEd