SIMETRICA

# Unrestricted Social Impact Assessment: Siemens Curiosity Project

March 14, 2018

The Shepherds Building, Rockley Road, W14 0DA
0203 883 9249
information@simetrica.co.uk
simetrica.co.uk

# Table of Contents

# 1 Executive Summary

We apply the human capital approach to value the benefits associated with the Curiosity Project in terms of the increased likelihood of students to proceed into STEM careers. The benefits of the Curiosity Project are elicited through primary survey data of students and teachers to better understand the impact of long-term (Greenpower, Rollercoaster, STEM Teach First teachers) and short-term (Science Festivals) Curiosity Project programmes on STEM outcomes, specifically on increased earnings associated with an engineering career, against a counterfactual control group of students who did not receive the Curiosity Project interventions. This gives us greater confidence in the robustness of our impact figures.

Our results show that Siemens contribution to the long-term Curiosity Project programmes (the Greenpower, Teach First, and Rollercoaster programmes surveyed in this report), provides a social return on investment of £11.90 for every £1 spent (Benefit Cost Ratio 11.9:1).

For Siemens contribution to the short-term Curiosity Project programmes (the Big Bang Fair in Birmingham, the Cheltenham Science Festival, the Edinburgh Science Festival and the Thorpe Park in Staines surveyed in this report), provides a social return on investment of £2.9 for every £1 spent (Benefit Cost Ratio 2.9:1).

The data available to us and our current estimations suggest that the long-term programmes have a greater impact for the cost of providing them. Based on these findings, it may be recommended that Siemens focus future Curiosity Project funding on maximising the impact of their STEM programmes through longer-term interventions, which have been shown to be most effective from a cost-benefit perspective.

We consider these impact figures to be realistic and defensible in the light of impact evaluations in other sectors.

# 2 Context of the study

Siemens believes that businesses should operate as sustainable, ethical and responsible member of the community. In order to aid this, it is necessary to understand society's challenges and needs, and act in the best interest of coming generations with respect to the economy, the environment, and society. Siemens' Corporate Citizenship activity aims to create sustained tangible and intangible benefits which impact on the company's employees and wider communities. Siemens works to connect with the wider community to ensure the sustainable development of society through technologies, local presence and leadership, and continues to work with schools, communities, and individuals using a range of resources and skills to help develop a more sustainable future.

A key focus for Siemens is the shortfall of people with the necessary skills in science, technology, engineering and maths (STEM) to support the engineering sector in the UK. Despite the increasing number of young people studying STEM subjects, there remains a large skills gap with a shortage of about 20,000 British engineering graduates leaving university each year (Engineering UK, 2017). Siemens is seeking to inspire the next generation of engineers through engaging with and inspiring young people. The Curiosity Project is an ongoing engagement programme by Siemens to broaden existing investment to bring STEM to life in the UK, increase the number of people choosing to study STEM subjects, and develop enthusiasm, passion and expertise for careers in STEM.

Siemens seeks to make a sustainable difference through its work as well as to maximise the positive impact of its investment. It is only possible to achieve these aims if the impact of Siemens's investment is assessed and measured in a rigorous and consistent way. Siemens approached Simetrica to build on Simetrica's 2015 Social Impact Assessment of Siemens' Curiosity Project[1]. The purpose of the continued collaboration is to assess the impact of the investment in STEM with bespoke data collected in the 2016-7 period, and to further develop a method to robustly measure the social impact of investment in STEM that can be used by Siemens and a wider audience to estimate future social impact of investment in STEM.

We have evaluated the following programmes and events from the Curiosity Project:

- Long-term programmes within secondary schools:
    - The Greenpower Educational Trust: A team-building programme which aims to advance education in the subjects of sustainable engineering and technology to

[1] Fujiwara, D., Lawton, R. & Trotter, L. (2015) Social Impact Assessment: Siemens Curiosity Project. Accessible at: https://docs.wixstatic.com/ugd/9ccf1d_f93b7b9328e74fdb9caf9a8424a77aee.pdf

young people. This is achieved through a practical project to design, build and race an electric car. We evaluate the results from the

  - Rollercoaster Challenge: Team building challenge to design and build a working rollercoaster, aimed at encouraging young people, especially young women, to take up careers in engineering.
  - Teach First (STEM teachers): Teach First provides teacher training in the UK focused on the needs of pupils from low-income communities. Siemens supports the specialist training of STEM teachers, aiming to improve scientific literacy and numeracy within schools in the UK.
- Short-term events:
  - Big Bang Fair in Birmingham, Cheltenham Science Festival, Edinburgh Science Festival, Thorpe Park in Staines.

We use the current best-practice approach to evaluation as set out by the UK Government in the form of the HM Treasury Green Book guidance. The Green Book sets out the evaluation guidelines for all central government departments in the UK and it aligns closely with best-practice guidelines issued by international organisations and many other OECD governments. The Green Book recommends Cost-Benefit Analysis (CBA) to understand the impact of programmes and interventions on society in terms of the benefits they create to participants, weighed against the costs of providing the project.

# 3 Methodology

### 3.1 Cost-benefit analysis

In CBA, all the costs and benefits associated with a programme are monetised and compared. This allows organisations to compare costs against outcomes on the same metric. The effectiveness of an intervention is then assessed by calculating the net social value due to the intervention (total social benefits – total social costs). These costs and benefits are calculated over the full life and legacy of the intervention.

In this study we use the Human Capital approach to CBA, which looks at impacts in terms of people's future labour market conditions. We look at the impacts of participation in the Curiosity Project on an individual's quality of life as well as on the public purse in the CBA study. Within this CBA framework there are three critical steps:

1. Determine the set of outcomes to be measured and for whom (the stakeholders).
2. Measure the impact of the project on those outcomes.
3. Valuation of the outcomes and costs for CBA.

Following the principles of CBA, we will measure the impact of the Curiosity Project on STEM outcomes, and value these outcomes in monetary terms, in order to compare them back to the costs of the project. This allows us to ascertain the overall net impact of the Siemens Curiosity Project programmes and the social value that they create.

We note that other methods for social impact assessment and policy evaluation exist, of which two of the most common are Cost-Effectiveness Analysis (CEA) and Social Return on Investment (SROI). However, our preference is for CBA in line with current Green Book Guidelines in the UK, which state that CEA should only be used in situations where CBA cannot be used (such as when outcomes cannot be valued). SROI is not a recognised approach in any OECD government or international organisation guidelines and suffers from a number of major problems and disadvantages in relation to CBA (see Fujiwara (2015) for a full review and critique of SROI).

## 3.2 Human Capital valuation

The specific approach we take within the broad field of CBA is the Human Capital approach. Different policy areas have developed versions of CBA that are specific to that area (Boardman et al., 2011) and the human capital approach has become the typical or accepted method for valuing outcomes in interventions related to education and training. The human capital approach to outcomes (i.e. benefit) valuation in CBA estimates monetary valuations by way of changes in individual earnings affected by the interventions in question. The approach is very flexible and can be (and has been) used in a wide variety of policy settings where earnings may be affected over the long term. It is an alternative valuation approach to stated preference and revealed preference methods.

The human capital approach is particularly relevant where the interventions in question relate directly to impacts on educational attainment and job and earning prospects and hence the human capital approach is recognised in the HM Treasury Green Book (2011, p.14) as a key valuation method for interventions and programmes related to skills training and schooling. Therefore, the human capital approach is ideally suited to the task of valuing outcomes related to the Siemens Curiosity Project in CBA.

## 3.3 Measuring Impact

Before we can assess the value of outcomes using the human capital approach we need to assess the level of outcomes produced by an intervention or programme. In statistics and other quantitative disciplines, the 'gold standard' for measuring the impact of a project or policy is through an experiment where the project has been randomly assigned across two or more groups of people. The process of randomisation ensures that the intervention group (treatment group) and non-intervention groups (control group) are on average identical in all respects prior to the treatment and so any differences in outcomes across the groups after treatment can only be the result of the treatment.

It was out of the scope of this study to use an experiment to measure the outcomes associated with the Curiosity Project because the relationships for delivering the project were already in place from the previous years. Where experiments cannot be employed, best-practice dictates that a quasi-experimental method should be employed instead. Quasi-experimental methods attempt to replicate as closely as possible the mechanics of an experiment but without the need for random assignment of the intervention. They can be used retrospectively and applied to projects that have already been completed. There are a large range of methods that sit within this category of methodology and there are different degrees of rigour involved. The commonality in quasi-experimental methods is the process of controlling for other factors. These methods try and control for (i.e., exclude) all other influences that may be explaining any changes in outcomes that are observed. If all the other potential influences can be controlled for, then the quasi-experiment will be robust. However, the ability to control for all other possible factors is rare and hence quasi-experimental methods tend to be at more risk of bias than experiments.

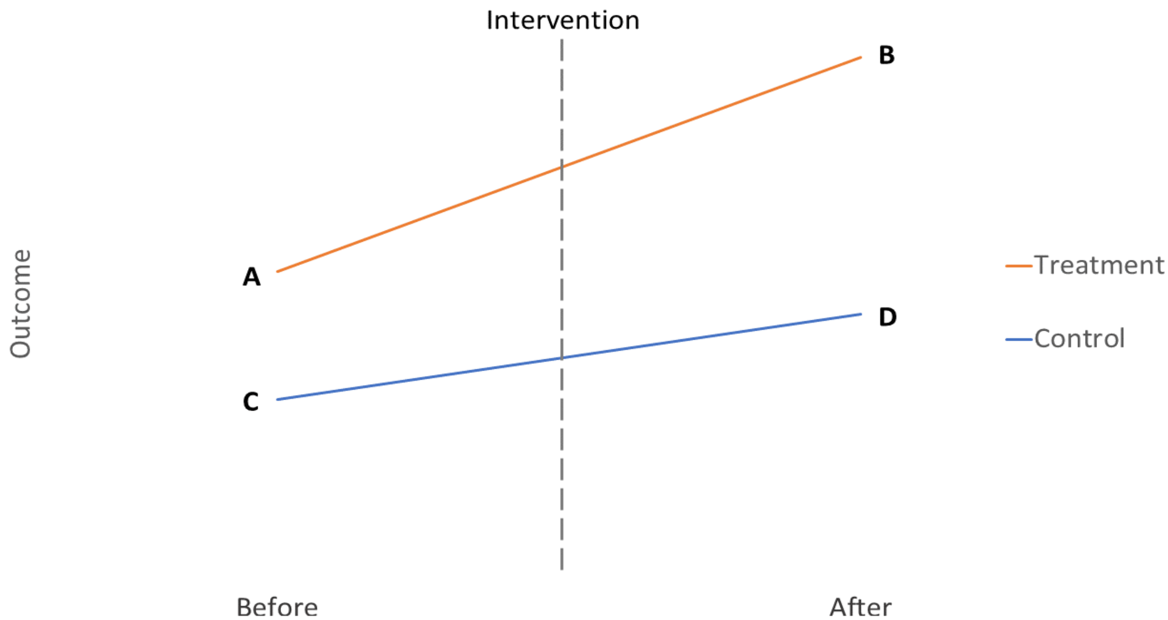
The main potential source of bias in the current study will be that for some programmes it is inevitable that some participants were initially interested in STEM subjects and hence likely to go onto further education or careers in STEM subjects anyway. This means that any observed changes in statements about future career paths cannot be wholly attributed to the Curiosity Project itself. This issue is known as selection bias and is a problem common to all quasi-experimental methods. It will typically result in overstatements of impact estimates.

However, in nearly all cases quasi-experimental methods are more robust than the third category of method known as non-experimental methods. Non-experimental methods are those which do not attempt to replicate the experimental setting or approach in anyway. That is, they make no attempt to control for (exclude) the influence of other factors. Therefore, they are susceptible to significant levels of bias and should not be used to infer cause and effect relationships. For example, a typical method in this category is the before and after study, which simply compares outcomes before and after an intervention without recognition that the observed trends in outcomes could have been due to many other factors than just the intervention itself. Another type of study is anecdotal evidence whereby participants report that the intervention has had a positive impact on them. Non-experimental studies such as these tend to overstate impact and hence are highly problematic for evaluation purposes.

For these reasons, public policy evaluation will use, as a minimum threshold, quasi-experimental methods, but wherever possible experimental methods are used (HM Treasury 2011). Non-experimental methods are generally avoided. Following this best-practice we employ a range of quasi-experimental methods or variants of quasi-experimental methods in the evaluation of the Siemens Curiosity Project.

In Figure 3.a, we set out a hypothetical graph showing the possible points of data collection in any study. In this setting, we have a treatment and control group and data is collected from both groups before and after the intervention. There are further variants of this hypothetical case (e.g., a number of pre-treatment data points could be made), but the current case can easily be modified to fit this and so it is a generalisable case. In this study, the vertical axis refers to enthusiasm towards STEM and the horizontal axis is time.

*Figure 3.a Potential data available in relation to an intervention*




One of the most rigorous quasi-experimental methods is the difference-in-difference (DiD) approach. This uses the trend in the control group as the counterfactual for the trend in the treatment group after controlling for initial differences. The strength of this method comes from the fact that the approach exploits variation over two dimensions: changes over time and changes across groups in the outcome. DiD is our preferred quasi-experimental approach for any study where treatment has not been randomly assigned.

If DiD is not possible or feasible there are other points of data that we can use under the broad quasi-experimental approach. The first option (in no particular order) would be to compare differences in the outcomes between the two groups at one point in time (after the intervention). Here we would compare the levels of the outcome represented by B (for the treatment group) against those represented by D (for the control group). Recognising that these groups are likely to have been different at baseline (because the intervention was not randomised) we can control for differences in key factors and characteristics at baseline, such as age, gender, region, family income, type of schooling etc. This is known as a conditioned cross-sectional study.

In the particular example, in Figure 3.a, we see that there was also an improvement in the control group's outcomes over time and so all of the change represented in the trend AB cannot be assigned to the intervention on its own.

The data collected for the Curiosity Project do not allow for our first-best statistical approach (DiD) – that is none of the programmes had data for all four points (ABCD). We therefore compare the outcomes reported by the intervention group with a non-intervention group at a single point in time (BD). Comparison of outcomes to a control group provides our counterfactual, while we apply corrections to account for the lack of before data (AC) in our analysis (outlined in detail below).

### 3.4 Choice of model

To evaluate the programmes, we are interested in the enthusiasm of respondents towards a career in engineering. We asked respondents to what extent they agree with a set of statements related to STEM and other outcomes. To account for differences in cognitive ability by age, we phrased the statement in two different age-appropriate formats (see Section 4.1).

Respondents had to pick between four options:

- Strongly disagree
- Disagree
- Agree
- Strongly agree

This type of response option is called a Likert scale in psychology. The choices are designed to provide a balanced number of positive and negative option choices.

From a statistical point of view, this kind of variable is referred to as an ordered categorical variable. Categorical variables do not measure a quantity but identify a category to which a respondent belongs. In this case, we categorise each option that the respondent selected over another. Ordered categories indicate that the categories have a natural order, "Strongly disagree" is more negative than "Disagree", which itself more negative than "Agree", etc.

We apply multivariate regression analysis to identify the increased enthusiasm observed among participants, compared to their peers in the reference group. Regression analysis allows us to simultaneously explore a relationship between multiple variables, controlling for many other factors (known as control variables) in the data that may also affect the outcome. This allows us to isolate the association between the intervention and the STEM outcome. Technically regression analyses capture association, rather than impact or causality, since we cannot exclude the full range of unobserved factors that may have a causal effect on people.

We use a model specifically designed for an ordered categorical variable called an ordered logit. Logit, or logistic, models are models made to analyse categorical variables. An ordered logit is a type of logit model specific for ordered categorical (as defined above).

The model measures the shift in enthusiasm observed by various characteristics. This indicates whether respondents who took part in one of the programmes report more positive choices to the Likert question (meaning they agree more than their counterfactual respondents controlling for sociodemographic characteristics).

We control for a range of sociodemographic factors which we anticipate could drive higher pre-existing interest towards doing a career in science or engineering. These may include an individual's age, gender, and indicators of socioeconomic background. By controlling for these factors, we are able to say with greater confidence that involvements in the Curiosity Project Programme is associated with an increase in STEM outcomes, holding constant other factors that may drive enthusiasm in these STEM outcomes.

### 3.5 Correction for optimism bias and long-term impacts

To estimate the benefits associated with the Curiosity Project, we require an estimate of an outcome, in this case the increase in likelihood of people going into engineering careers as a result of involvement in the Siemens Curiosity Project. The data that we have is in the form of post-project self-assessment of participants' enthusiasm towards engineering compared to a control group.

We note that a simple increase in interest in engineering does not necessarily mean that people will actually go on to obtain a career in the subject field. A wide range of other, unobservable, factors may be at play, which make an individual more or less likely to continue into an engineering career. This will also need to be accounted for in the analysis. Evidence shows that the correlation between intention and behaviour is low (less than 20%) (Fujiwara et al., 2015) and this is in instances where people are asked to state their intention concerning a specific activity in the near future. In our case, where people are asked about their intention to go into an engineering career, the actual behaviour is far into the future and there are considerable barriers to achieving this goal, even if people do intend to carry out the behaviour (such as ability, labour market conditions and luck). Therefore, we expect an increase in enthusiasm to become an engineer to only be very marginally associated with actually becoming an engineer in later life (much lower than 20%). In the absence of additional guidance from the literature we use a conservative estimate of **behaviour change** and assume that 5% of people who state that they are enthusiastic about pursuing a career in engineering in their teenage years would actually choose to pursue a career in engineering. We believe people are less likely to proceed to an engineering career based on participation in a short-term project (which has a momentary or temporary impact on their life), compared to a long-term project (which we expect to impact on more aspects of their life over a prolonged period), and apply a 1% behaviour change correction to the short-term impact evaluation.

In all analysis, we control for **optimism bias.** There is a demonstrated, systematic tendency for people to be overly optimistic when predicting the benefits of a project. This is known as optimism bias and it usually affects capital development projects but can also be inherent to other forms of programmes. In the current study students may be overly optimistic about their likelihood of going onto further education or careers in STEM subjects and hence optimism bias may become an issue wherever projections or forecasts on behalf of the beneficiaries or teachers are made. We therefore apply a reducing factor for optimism bias in all cases where forecasts are made in the data. For the optimism bias calibration factor, we keep 80% of the effect identified in the long-term interventions[2] and 60% for short-term interventions. We are more conservative with short-term interventions as there might be a possible focusing bias due to asking people about their enthusiasm for engineering at an event focused specifically on this topic.

Table 3.1 shows the percentage of the identified effect size that we are confident attributing to the project, taking in to account the various biases above.

*Table 3.1 Effect size kept after correction of bias for long-term or short-term interventions*

| Long-term interventions | Short-term interventions |
|---|---|
| **4%** | **0.6%** |

## 3.6 Valuation

There are two types of values that must be considered when assessing social impact for the purposes of cost-benefit analysis. The first, known as **primary benefits**, concern values directly held by individuals. This is the value of the positive effect that the outcome has on people's wellbeing directly. This could be an improvement in health, a reduction in crime or finding work, all of which will improve people's wellbeing and quality of life (QoL).

A second type of values are **secondary benefits**. These refer to values to society more widely, which may indirectly benefit (and hence have value for) individuals. These are usually measured as impacts on public services and on the public purse (Exchequer). For instance, for a health improvement the secondary benefits will be a reduction in state health-related expenditure.

Primary and secondary benefit values are accounted for in the human capital approach, which recognises that an improvement in job prospects and earnings due to training will improve the quality of people's lives directly, and also will have impacts on the public purse in terms of, for

[2] based on an average of figures used in Green Book Supplementary Guidance: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/191507/Optimism_bias.pdf

example, increased income tax receipts and less unemployment benefit payments. Secondary benefit values are usually already measured in monetary or financial metrics and there are a number of UK databases that can be used to estimate them.[3]

Both primary and secondary benefit values are crucial for measuring social value in a full holistic sense and both should be included in social impact measurement methods like CBA. This study will seek to measure both types of value to the greatest extent possible under the framework of the human capital approach to CBA.

# 4 Data

### 4.1 Survey instruments

Multiple surveys were designed to collect primary data on each of the Curiosity Project interventions. Each survey was carefully developed in collaboration with Siemens to:

- ensure that it complies with child protection principles;
- is carefully written to coincide with the specific context of the intervention;
- is adapted to the respondent's age and current education level, and
- include a core set of identical questions across all interventions to be analysed.

All students were asked the following questions:

- In which region are you currently living?[4]
- What school year are you currently enrolled in?[5]
- What is your gender?
- Do you get any help or support at school due to a health or behavioural issue or a disability?
- What month and year were you born?
- Do you go on a family holiday away from home every year (e.g., caravan holiday, Butlins holiday camp, package holiday abroad)?
- How often do you do something related to science or engineering outside the classroom (excluding school trips)? This might be attending a science club at school, watching a science documentary or going to a science museum.

---

[3] https://www.gov.uk/guidance/transport-analysis-guidance-webtag

[4] Note that for surveys taken in a class, the students did not have to answer those two first questions as the teacher would have provided us with that information for the whole class already.

[5] See above

Quasi-experimental methods require that we identify a set of factors that could also influence the outcome(s) of interest and that we collect the relevant data to control for these factors (for more details, please refer to the previous section on measuring impact). By asking these questions, we are able to capture other influences that may be driving observed changes in outcomes. This set of standard questions were asked identically to all respondents, regardless of age.

The second core set of questions asked was in relation to education prospects, school subjects of interest, and agreement/disagreement with various statements about science and engineering. The purpose of these questions was to capture potential changes in pupil outcomes associated with participation in the Curiosity Project.

These core questions had potential comprehension problems for younger respondents due to their complexity or unfamiliarity.[6] To ensure respondents fully understood the concepts in the questions, we displayed different worded versions of the questions depending on their age:

- 11-14-year-olds: Language graded to the reading and comprehension capacities at this age level
- 15-18-year-olds: Language not graded (adult-level)

Our analysis focuses on pupils' interest in doing an engineering career. When asking the respondents "To what extent do you agree with the following statements?", the statement related to this outcome was framed:

- For 11-14-years-old: "Engineering/science is an interesting career"
- For 15-18-years-old: "I would consider a career in science or engineering"

This study draws on three main sources of data:

1) **Long-term interventions surveys**, for which we developed an online teacher survey, linked to a student online survey. The teacher surveys were sent by email to teachers or team leaders of an intervention group (in the case of Greenpower), asking them general information about the class or team, like school grade, gender composition, etc. Once completed, they were provided a web link to a student survey to be forwarded to their students. This web link allowed us to carry over class-level data elicited in the teacher survey and include this within the student data. It also provided information on the age of the child, feeding into the graded young/old branches of the survey.

[6] For a complete list of all the survey questions, see "Survey Instruments" in Annex.

For the Rollercoaster programme, we also developed a standalone pad version of the student survey which did not require the prior teacher-linked survey to be performed (more suitable for one-off events where children may attend with parents/guardians). It contained the necessary screening questions to appropriately display the survey according to the respondent's age.

2) **Short-term interventions survey**. We designed a single multi-event survey instrument that could be delivered on-site using a pad. It contained the necessary screening questions to appropriately display the survey according to the respondent's age. The core set of questions was identical to the long-term surveys.
3) Finally, we designed a **control group survey** asked to an online panel of pupils across the United Kingdom. The control group sample was provided by Toluna, a respected provider of online survey panel respondents. The questions were written to allow comparability with the intervention surveys. Survey questions were branched into younger/older graded questions (as above) based on respondent age.

Long-term interventions cover a number of programmes, such as Greenpower, Rollercoaster, and Teach First STEM Teachers. These programmes have in common that they took place over several months to a year. For the purpose of analysis, we combine all these programmes under the banner **'long-term programmes'**. Due to low sample sizes for some of the individual long-term programmes, we are unable to disaggregate analysis to assess the impact of each programme individually.

Short-term programmes, like Science Festivals, are events that took place during just a few days and the students attended one of those days for a few hours. They are combined under the banner **'short-term programmes'**. Again, it is not possible to separate out the impacts attributable to individual Science Festivals.

The control group survey provided us with a total sample of 674 respondents. The control groups in both cases represent a subset of the control group sample, chosen to match as closely as possible the population of pupils targeted by those interventions (i.e. if the treatment question is asked only to year 7 to 11 and the intervention only takes place in England, the control group is composed of 7 to 11 year-olds living in England). This is because the data collected for long-term and short-term Curiosity Project programmes differed by region or age group. This approach to sub-setting the control group increases the validity of the analysis and the statistical confidence that the counterfactual group represents an indication of what would have happened without the Siemens-funded interventions.

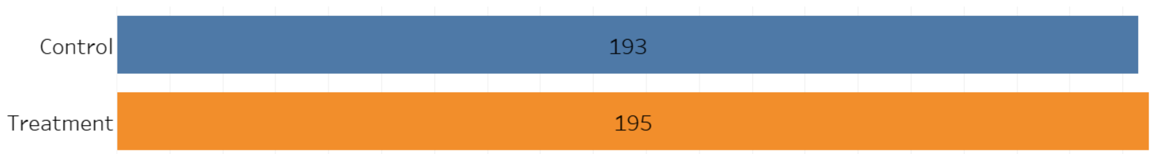
As part of the human capital impact evaluation of the Curiosity Project, we focus predominantly on ensuring robustness through the core principles of research design – Construct, Internal, and Statistical validity - to ensure that we provide an informative estimate for a specific subset of the population (see Annex: Research design and validity for an introduction of this concepts). In our research design, we have made the choice to adapt the content of the surveys to the age of the respondents, dividing our sample into two age groups. This ensures greater internal validity in our analysis. However, as we outline below, it does mean that our impact estimates are based on only those age groups for which sufficient sample size is available (the younger age group in the case of long-term programmes and the older age group for short-term).

## 4.2 Sample

### *4.2.1 Long-term programmes*

Our sample for the analysis of long-term interventions is composed of a total of 388 respondents. Among those respondents, 195 took part in one of Siemens programmes and will constitute our treatment group, and 193 responded an online survey and constitute the control group.

*Figure 4.a Long-term: Sample size*

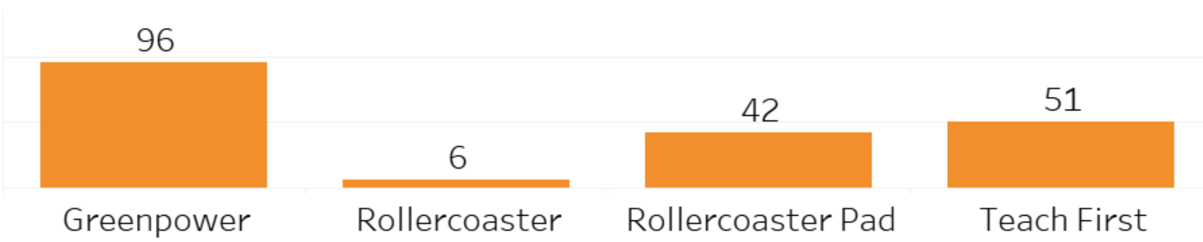



The treatment group is composed of participants from three Curiosity Project programmes:

- 96 took part in a Greenpower activity.
- 48 took part in a Rollercoaster programme (6 answered an online survey forwarded by their teacher and 42 answered the survey on a pad during an event).
- 51 are in a class with a Teach First teacher.

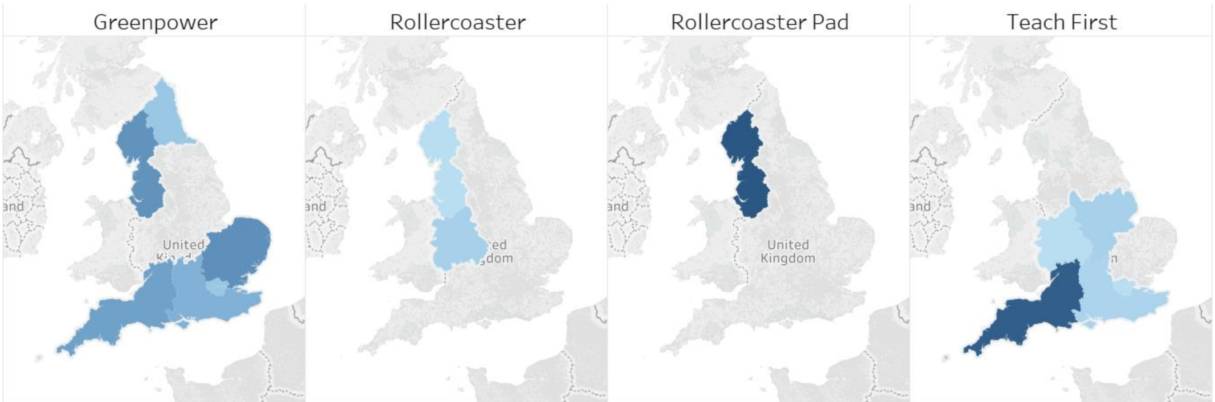
Below we outline the descriptive statistics for the key demographic characteristics of the sample and the survey questions asked of the long-term treatment and control groups.

*Figure 4.b Long-term: Treatment Sample*



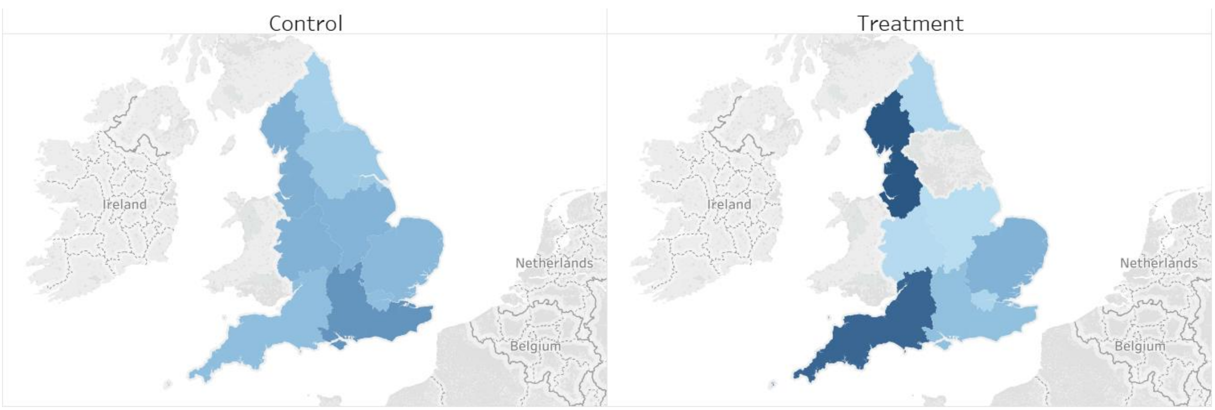

The Curiosity Project programmes took place in various locations across England, providing a good representative spread of regions across England (See Figure 4.c). All regions of England are represented in our control group, to ensure representative of the Siemens impact evaluation against a national counterfactual (See Figure 4.d). Full tables are provided in Annex.

*Figure 4.c Long-term: Treatment sample location*




Note: The shades of blue indicate the number of respondents in each region. The lightest blue representing n=1 respondent and the darkest n=42.

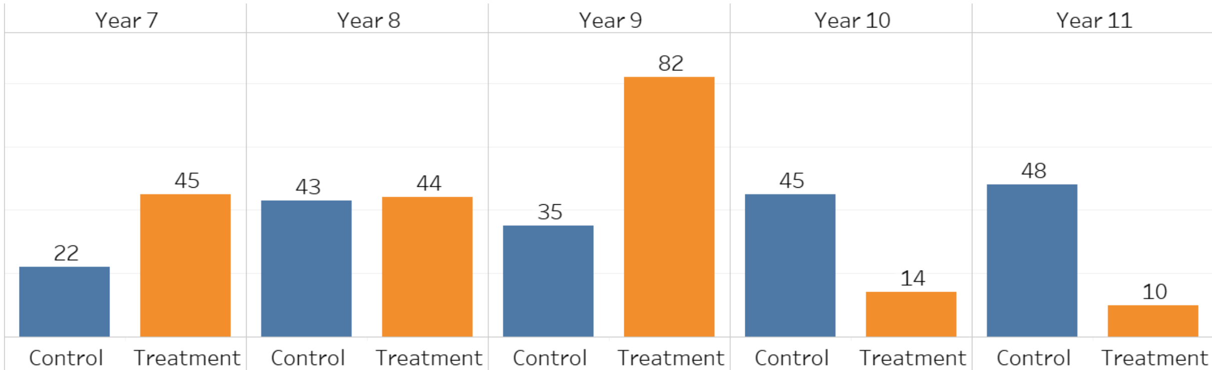
*Figure 4.d Long-term: Treatment and control sample location*




Note: The shades of blue indicate the number of respondents in each region. The lightest blue representing n=5 respondents and the darkest n=66.

Figure 4.e shows the composition of the treatment and control group by school year. Six respondents in the treatment group reported being in year 12 or 13. Recall that the survey questions are graded by age (younger/older graded language). Therefore, the sample size for the older school years (years 12-13) was too low to allow impact evaluation. Consequently, impact assessment of the long-term interventions is performed only on pupils from year 7 to 11.

*Figure 4.e Long-term: School year*




Treatment and control groups for the long-term programmes are of similar gender composition, with a slight predominance of males in both group, 53% in the control group and 54% in the treatment (see Figure 9.a in Annex). As an indicator of socioeconomic status, we asked if the respondent takes a family holiday away from home every year. 78% in the control group and 83% in the treatment group go on holiday every year, while respectively 20% and 12% do not go every year[7] (See Figure 9.b in Annex). Participants were asked "Do you get any help or support at school due to a health or behavioural issue or a disability?" as an indicator of pre-existing health/behavioural issues that could affect their likelihood to enrol on the project and pursue STEM outcomes, and which should be controlled for as best practice (See Figure 9.c in Annex). Finally, to control for pre-existing interest and enthusiasm towards STEM, we asked participants if they regularly take part in other science activities (See Figure 9.d in Annex).

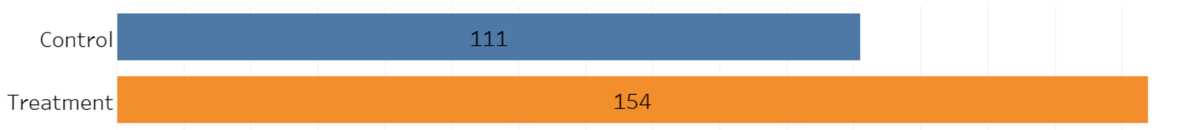
For more detailed descriptive statistics, see Table 9.1 in Annex.

#### *4.2.2 Short-term programmes*

Our sample for the analysis of short-term interventions is composed of a total of 265 respondents. Among these respondents, 154 attended a Science Festival or other short-term event supported by Siemens and will constitute our treatment group. The online control group is composed of 111 respondents, matched by age to the treatment group.

[7] Note that 1.6% of respondents in the control group and 5.1% in the treatment group reported "Don't know / Rather not say"

*Figure 4.f Short-term: Sample size*




The treatment group is divided in various subgroups depending on which of the events/Science Festivals they attended:

- 38 attended Big Bang Fair in Birmingham.
- 31 attended Cheltenham Science Festival.
- 12 attended Edinburgh Science Festival.
- 73 attended Thorpe Park in Staines.

Below we outline the descriptive statistics for the sample demographic characteristics and survey questions asked of the short-term treatment and control groups.

*Figure 4.g Short-term: Treatment Sample*


The Science Festivals took place in Birmingham, Cheltenham, Edinburgh and Thorpe. To account for those who travelled to attend the Science Festivals, we consider the region of origin and not just the city location of the Science Festival in our analysis (Figure 4.h). The control group was subset to include only regions where one of the Science Festivals occurred (Figure 4.i).

*Figure 4.h Short-term: Treatment sample location*

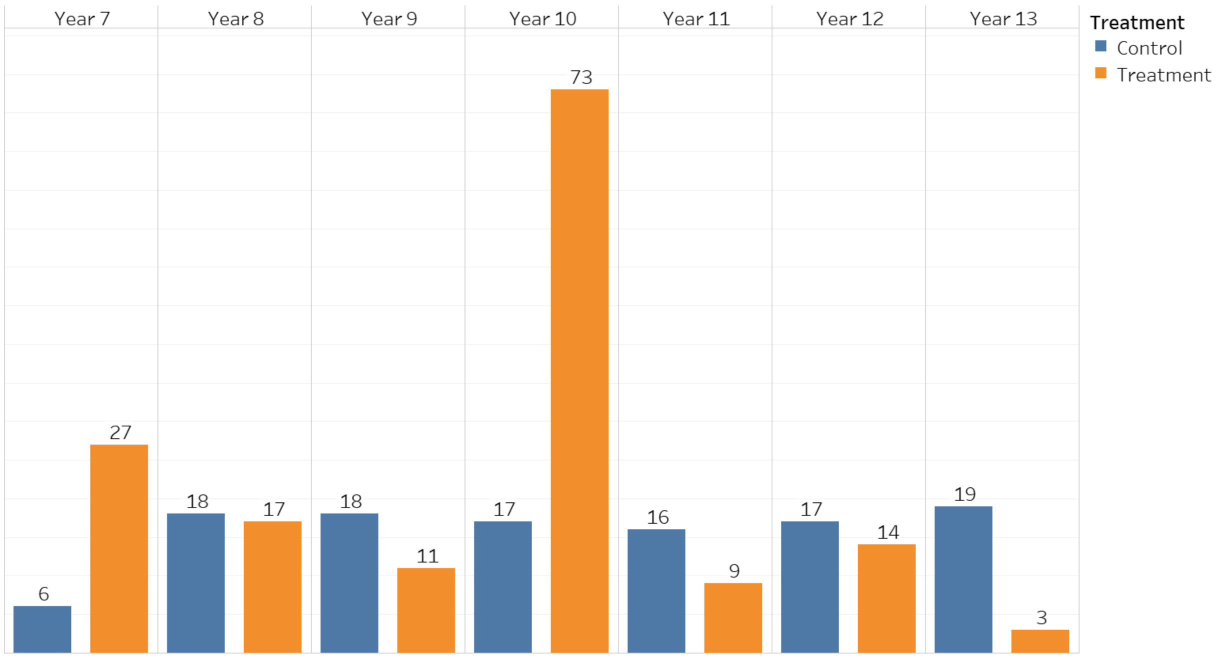

Note: The shades of blue indicate the number of respondents in each region. The lightest blue representing 12 respondents and the darkest 73.

*Figure 4.i Short-term: Treatment and control sample location*


Note: The shades of blue indicate the number of respondents in each region. The lightest blue representing 12 respondents and the darkest 73.

Figure 4.j shows the composition of the treatment and control group by school year. Recall that the survey questions are divided by younger and older grade. The sample size for the younger school years (years 7-9) was too low to allow robust impact evaluation of short-term programmes. Consequently, impact assessment of the short-term interventions is performed only on pupils from year 10 to 13.

*Figure 4.j Short-term: School year*




Treatment and control group for short-term programmes are of similar gender composition, with a slight predominance of males in both groups (60% in the control group and 51% in the treatment) (See Figure 9.e in Annex). 78% in the control group and 83% in the treatment group go on holiday every year, while respectively 20% and 14% do not go every year[8] (See Figure 9.f in Annex). Participants were asked "Do you get any help or support at school due to a health or behavioural issue or a disability?" (See Figure 9.g in Annex). Finally, to control for pre-existing interest and enthusiasm towards science, we asked participants if they regularly take part in other science activities (See Figure 9.h in Annex).

For more detailed descriptive statistics, see Table 9.2 in Annex.

# 5 Results

## 5.1 Long-term programmes

### *5.1.1 Methodology*

Students were asked about their opinion on various statements. To each statement, respondents had to pick between four options:

[8] Note that 1.8% of respondents in the control group and 3.2% in the treatment group reported "Don't know / Rather not say"

- Strongly disagree
- Disagree
- Agree
- Strongly agree

To evaluate the programmes, we are interested in the enthusiasm of respondents towards a career in engineering. Depending on their age, we asked the question in a different form:

- For 11-14-years-old: "Engineering/science is an interesting career"
- For 15-18-years-old: "I would consider a career in science or engineering"

We apply multivariate regression analysis to identify the increased enthusiasm observed in the treatment group, compared to the reference group, and controlling for other confounding factors (full description in Section 3.4). The set of controls were selected as:

- Age
- School year
- Gender
- Going yearly on holiday
- Has a disability
- Take part in similar science activities
- Region

#### *5.1.2 Analysis*

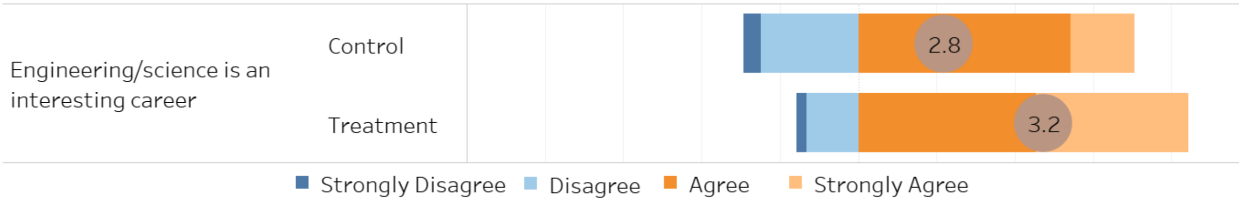
Figure 5.a provides a graphical display of the results of our **primary outcome of interest: enthusiasm towards an engineering career.**

The four response options are scored from 1 to 4, with 1 being "Strongly disagree" and 4 being "Strongly agree". We calculate the average 'score' for each statement (the number within the circle in the figure).[9]

---

[9] We acknowledge this score to be imperfect as we might have reason to believe that the "distances" between those four points to not being equal. For example, the "distance" between "Strongly agree" and "Agree" may be shorter than the distance between "Agree" and "Disagree". However, it is a useful addition to the comparison of general agreement.

*Figure 5.a Enthusiasm towards a career in science or engineering: Long-term programmes*




A higher proportion of the treatment group agree with the statement "Engineering/science is an interesting career" (the stack of orange bars is longer). The treatment group also has a higher mean Likert Score (3.2 compared to 2.8 among the control group), which indicates that **those who participated in the long-term elements of the Curiosity Project had stronger enthusiasm for an engineering career overall, compared to a control group of their peers in the wider population.**

Our regression model estimates the statistical difference in level of agreement (in percentage points) attributed to being part of the treatment group, controlling for other factors. We also provide information on other statements asked as part of the broader Curiosity Project evaluation, but which do not contribute to the human capital impact model.

*Table 5.1 Difference in agreement for statements (in percentage points, ordered logit): Long-term programmes*

| Statement for younger students | Percentage point difference in agreement (Agree & Strongly Agree) |
|---|---|
| Engineering/science is an interesting career | +13 (***) |
| I have a better idea of what engineering is and how scientists and engineers make a difference to my life | +12 (***) |
| Mathematics is one of my best subjects | 9 |
| I have always done well in mathematics | 8 |
| I enjoy the class and would like to take part in similar science/technology/maths activities | +6 (**) |
| Foreign languages are an interesting career | 3 |
| I could use what I have learned | +3 (*) |
| I get good marks in mathematics | 3 |
| What I have learnt in class helped me understand something I didn't | +3 (**) |
| I have learned new things | 1 |

*Notes: Proportions are calculated using ordered logit regression controlling for age, school year, gender, going yearly on holiday, has a disability, take part in similar activities and region. *** <1% significance; ** <5% significance; * <10% significance*

For the purposes of our human capital impact evaluation, we are interested in the relationship between taking part in a long-term Curiosity Project programme and students' reported enthusiasm towards engineering and science. This association is significant and positive, **associated with a +13**

**percentage-point increase**, after controlling for a number of other possible drivers of enthusiasm in engineering.

*Box 5-a Effect of gender on the programme*

The results in Table 5.1 suggests that taking part in the long-term Curiosity Project programmes is positively and significantly associated with higher enthusiasm towards a career in science or engineering across the whole sample, but that being female doesn't have any significant interaction with the treatment (either positive or negative). In other words, the impact of the programme is not statistically different for male and female participants. For more details, see Annex: Gender interaction model.

## 5.2 Short-term programmes

### *5.2.1 Methodology*

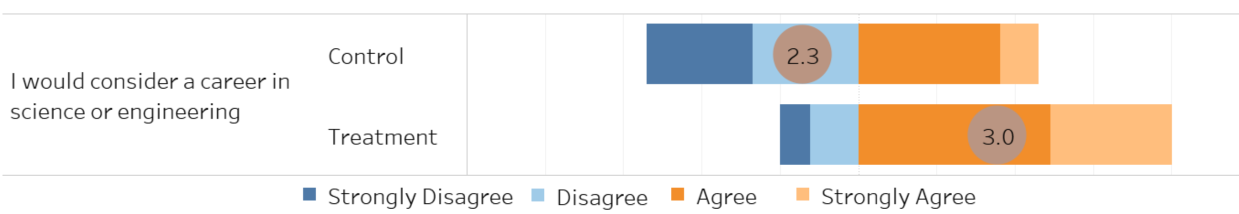
Students were asked about their opinion on various statements, on a four-point 'Likert' scale (as in the long-term model, above). Likert scales are commonly used for measuring attitudes by asking people to respond to a series of statements about a topic, in terms of the extent to which they agree with them, and so tapping into the cognitive and affective components of attitudes (full explanation in Section 3.4).

To evaluate the programmes, we are again interested in the enthusiasm of respondents towards a career in engineering (divided by age, as above). The econometric model included the same set of controls (excluding region).

### *5.2.2 Analysis*

Figure 5.b provides a graphical display of the results of our primary outcome of interest: enthusiasm towards an engineering career.

*Figure 5.b Enthusiasm towards a career in science or engineering: Short-term programmes*




A higher proportion of the short-term treatment group express agreement to the statement "Engineering/science is an interesting career" (the stack of orange bars is longer). The treatment group also has a higher mean Likert Score (3.0 compared to 2.3 among the control group), which indicates that **those who participated in the short-term elements of the Curiosity Project reported stronger enthusiasm for an engineering career overall, compared to a control group of their peers in the wider population.**

Our model estimated the difference in level of agreement (in percentage points) that could be attributed to being part of the treatment group, controlling for other factors. We also provide information on other statements asked as part of the broader Curiosity Project evaluation, but which do not contribute to the human capital impact model.

*Table 5.2 Difference in agreement (in percentage points, ordered logit)*

| Statement for older students | Percentage point difference in agreement (Agree & Strongly Agree) |
|---|---|
| I would consider a career in science or engineering | +26 (**) |
| I now understand how what I learnt today can be used in science or engineering | +32 (**) |
| I have enjoyed the day and found the fair engaging | +26 (***) |
| I would consider a career using foreign languages | 9 |
| The fair has increased my knowledge of the subject and I could describe what I learned | 8 |
| I could use what I learned to solve a problem or apply it to a new situation | 3 |
| I now understand what science or engineering is and how science and engineering contributes to society | 1 |

*Notes: Proportions are calculated using ordered logit regression controlling for age, school year, gender, going yearly on holiday, has a disability and take part in similar activities. *** <1% significance; ** <5% significance; * <10% significance*

For the purposes of our human capital impact evaluation, we are interested in the relationship between taking part in a short-term Curiosity Project programme and students' reported enthusiasm towards engineering and science. This association is significant and positive, **associated with a +26 percentage-point increase**, after controlling for a number of other possible drivers of enthusiasm in engineering.

In the section below, we will apply the impact estimates obtained for the respective age groups to the whole sample and use this to estimate the overall impact of the Curiosity Project. While there may be some limitations to the external validity of this approach (the results obtained for younger age groups may not apply identically to the older age group, and vice versa), we prefer to guarantee internal validity rather than produce impact estimates of limited predictive power based on low sample size (recall Section 4.1).

# 6 Cost-Benefit Analysis

## 6.1 Estimating benefits

Using the framework of the human capital approach, the primary benefits associated with the Curiosity Project are estimated through data on **increased earnings associated with an engineering career.**

To understand the impact of the Siemens Curiosity Project in encouraging people to take up a career in engineering, we looked into the benefits of choosing an education in engineering over another subject. Our reference case is a counterfactual that the pupil would have chosen any other subject to study in further education. Accordingly, we apply the wage difference between engineering and all other subjects in our CBA model.

We use data from the Engineering UK report (2017) which provides estimated mean salary for UK graduates in full-time employment six months after graduation. Salary information is available for people studying to first degree, postgraduate degree, or doctorate. Wage premia are based on 2014/15, which we uprate to 2016 values using the Bank of England inflation calculator (Table 6.1).

*Table 6.1 Wage premia by highest educational level (Engineering UK 2017)*

| | Engineering annual salary | All subjects annual salary | Wage premium (annual) | Proportion of qualifications in 2015 in the UK | Lifetime benefits |
|---|---|---|---|---|---|
| Doctorate | £34,718 | £34,160 | £559 | 1% | £2,594 |
| Taught postgraduate | £29,177 | £26,556 | £2,621 | 9% | £19,263 |
| First degree | £26,595 | £22,313 | £4,282 | 10% | £35,612 |
| **Expected lifetime benefits (all qualification levels)** | | | | **100%** | **£5,230** |

*Source: Engineering UK report 2017, p. 165 (values uprated to 2016 figures). Source for proportion of qualifications in 2015 in the UK: Annual Population Survey, 2015. Wage premia calculated as (wage for engineering – wage for all subjects). Lifetime benefits calculated as wage premia * (10 years – number of additional years education) (applying HM Treasury 3.5% Net Present Value).*

We assume that the impact of the Curiosity Project only has an effect on the pupil's subject of interest and does not directly influence their level of education (in other words, pupils in the Curiosity Project are as likely to proceed to postgraduate or doctoral studies as pupils in the wider population). We use average national population levels of education in the UK in 2015 as an indicator of the level of education that pupils are likely to attain (fourth column Table 6.1). Based on this data, 20% of the population are predicted to study to degree level or above, and thereby generate benefits to be accountable in our CBA. This approach allows us to partly reduce some of the bias in our data, since we do not assume that all Curiosity Project attendees would proceed to higher education.

Table 6.1 provides earning premia for educational attainment for a single year. The primary benefits of improved earning in engineering careers will be felt over the course of an individuals' lifetime. However, other factors also affect an individual's earning potential later in their career, including continued personal development and in-work training. For this reason, we estimate career benefits across ten years, rather than their whole lifetime. We calculate the flow of benefits for each higher education degree level. We estimate the benefits associated with each qualification at the time that an individual obtains their highest qualification level (given that a postgraduate masters require one extra year of study, and doctorates four years more than a first degree, these two scenarios will generate fewer total years of benefits ).[10] We apply HM Treasury Green Book (2011) discounting rates at 3.5% to calculate net present value (NPV).

In sum, an individual who intends to pursue a career in engineering can expect an average wage premium of £5,230 (averaged across all qualification levels) (final line Table 6.1).[11]

## 6.2 Overall benefits of the programmes

97,545 students were involved in long-term Curiosity Project programmes.[12] Regression analysis shows that involvement in the long-term Curiosity Project interventions is associated with a +13 percentage-point increase in enthusiasm to pursue an engineering career, compared to the national population control group.

117,932 students have been involved in short-term Curiosity Project programmes. Regression analysis shows that involvement in the short-term Curiosity Project Science Festivals interventions is associated with a +26 percentage-point increase in enthusiasm to pursue an engineering career, compared to the national population control group.

The Curiosity Project programmes target students aged 11 to 18 years old. For discounting purposes, we will use 15, the center of that age category, as a reference point. Our estimated net present value benefits are calculated at age 22 and will require to be discounted an extra 7 years.

We calculate the benefits of the change in enthusiasm for engineering using the **engineering career wage premia methodology.**

*Table 6.2 Estimated benefits associated with enhanced enthusiasm for engineering career*

| | Total | Difference in | Bias | Average | Discounting | Project |
|---|---|---|---|---|---|---|

[10] Note that we assume the cost of an education in engineering to be no different than studying another subject

[11] 1% will study to doctorate level, with a net present value of £2,594 over ten years. 9% will study to postgraduate level with a NPV of £19,263. Finally, 10% will stop at a first degree and get a NPV of £35,612.

[12] See **Error! Reference source not found.** for the detail.

| | population (number of participants aged <18) | | proportion of students enthusiastic about engineering (T vs. C) | | correction (proportion of impacted population) | | lifetime wage premium (£) | | | | lifetime net present value benefits (£) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Long-term programmes | 97545 | × | 13% | × | 4% | × | £5,230 | ÷ | $1.035^7$ | = | £2,085,103 |
| Short-term programmes | 117932 | × | 26% | × | 0.6% | × | £5,230 | ÷ | $1.035^7$ | = | £756,268 |

*T = Treatment group (Curiosity Project participants); C = Control group. Net present value (NPV) calculated using Green Book discounting rates at 3.5%. Bias correction measures outlined in full in Table 3.1. Average lifetime wage premium* Table 6.2.

*Table 6.3 Individual lifetime net present value benefits*

| | Project lifetime net present value benefits (£) | | Total population | | Individual lifetime net present value benefits (£) |
|---|---|---|---|---|---|
| Long-term programmes | £2,085,103 | ÷ | 97545 | = | £21.38 |
| Short-term programmes | £756,268 | ÷ | 117932 | = | £6.41 |

Lifetime net present value benefits calculated for long-term Curiosity Project programmes amount to £2,085,103. This equates to an average lifetime benefit of £21.38 per person involved in long-term programmes.

Lifetime net present value benefits calculated for short-term Curiosity Project programmes amount to £756,268. This equates to average lifetime benefits of £6.41 per person involved in short-term programmes.

The benefits produced by the Curiosity Project is zero for those years while beneficiaries are still in education (in other words financial benefits only accrue to the individual when they reach employment). This makes it difficult to estimate a single annual benefit figure of the project (because the benefits will vary depending on what year of a beneficiary's life we are estimating). We can however calculate equivalent yearly benefits across the individual's next ten years by applying HM Treasury Green Book corrections for Net Present Value.

To obtain the **equivalent yearly benefits** over the 10 years following participation in the Curiosity Project we use the following formula:

$$Equivalent\ Yearly\ Benefits = \frac{Net\ Present\ Value * \left(1 - \frac{1}{1+\propto}\right)}{1 - \frac{1}{1+\propto}^{10}}$$

$$where\ \ \propto = 3.5\%\ (the\ discount\ rate\ )$$

The equivalent yearly benefits of the long-term Curiosity Project programmes amount to £242,237.

The equivalent yearly benefits of the short-term Curiosity Project programmes amount to £87,860.

## 6.3 Total impacts

In this section we combine the impact results above to estimate the total lifetime costs and benefits associated with the long-term and short-term Curiosity Project Programmes, we surveyed as part of the Siemens 2016/2017 impact assessment report. We provide details of total project costs and benefits and average costs and benefits per student for the short and long-term programmes. With this data we are able to estimate the net benefits for each strand of the Curiosity Project (the total benefits after project costs have been considered). We are also able to estimate the benefit cost ratio (BCR), which tells us the total benefits produced for each £1 invested in the programme.

*Table 6.4 Cost & benefits: Long-term Curiosity Project programmes*

| Long-term projects | Lifetime benefits |
|---|---|
| **Benefits** | |
| Total benefits | £2,085,103 |
| Average benefits (per student) | £21.38 |
| **Costs** | |
| Total projects' costs including opportunity costs | £175,468 |
| Average costs (per student) | £1.80 |
| **Cost-benefit analysis** | |
| Net benefits of the projects | £1,909,635 |
| Net benefit (per student) | £19.58 |
| Benefit-cost ratio | 11.9 |

*Table 6.5 Cost & benefits: Short-term Curiosity Project programmes*

| Short-term projects | Lifetime benefits |
|---|---|
| **Benefits** | |
| Total benefits | £756,268 |
| Average benefits (per student) | £6.41 |
| **Costs** | |

| | |
|---|---|
| Total projects' costs including opportunity costs | £262,321 |
| Average costs (per student) | £2.22 |
| | |
| **Cost-benefit analysis** | |
| Net benefits of the projects | £493,947 |
| Net benefit (per student) | £4.19 |
| Benefit-cost ratio | 2.9 |

The CBA represents the overall cost-benefit results for society accounting for impacts on individuals and the exchequer. We calculate net benefits of the long-term projects (overall social benefits minus costs) by accounting for total costs as project costs, plus an additional opportunity cost of 8% (Jenkins and Kuo 2007). This is based on the average social discount rate used to calculate net present value of investment projects, which captures the potential profits that could have been made through an alternative use of funds (HM Treasury, 2011).

- **Net benefits of the long-term Curiosity Project programmes are £19.58 per student summing to a total of £1,909,635 over ten years. The benefit-cost ratio is 11.9.**
- **Net benefits of the short-term Curiosity Project programmes are £4.19 per student summing to a total of £493,947 over ten years. The benefit-cost ratio is 2.9.**

### 6.4 Summary and caveats to the analysis

We note a number of important caveats in our study and approach. First, this study has been dictated to a large extent by the type of data that was collected during the different programmes. We were not able to assess the Curiosity Project impact using an experiment or conduct pre-intervention data collection and this has restricted our ability to arrive at a causal estimate. Therefore, the possibility that other factors may partly explain some of the observed changes or impacts must be acknowledged when reading and interpreting the results. For instance, one individual will have higher existing levels of 'science capital', in terms of knowledge, interest, scientific literacy and science/technology-related social contacts.[13] In other words, for some projects it is inevitable that participants that were initially interested in STEM subjects and hence likely to go onto further education or careers in STEM subjects anyway got involved in the project. These possible factors are likely to result in an upward bias since some of the students involved in the projects will have selected in to the project. This would result in our estimates being overstatements of the impact and social value of the Curiosity Project. For this reason, we take conservative estimates wherever possible.

[13] http://www.sciencemuseum.org.uk/educators/special_projects/enterprising_science_project.aspx

Second, outcome variables as measured in a number of the Siemens-supported programmes are self-reported measures of enthusiasm around educational or career choices towards STEM subjects. In such cases, there are a number of causal leaps to be made between such an assessment and the final achievement of a STEM career or education. In these cases, we have made predictions and estimations using best available data, including, for example, official progression rates from GCSE to further education, and estimates of general population enthusiasm for STEM careers, and applied bias reduction measures where appropriate.

Third, CBA (and indeed other methods of policy evaluation) requires that all of the impacts (both positive and negative) be accounted for in the evaluation. Our study has focussed on an important but narrow set of outcomes related to the Curiosity Project. Although progress into education and careers in STEM subjects are probably the key outcomes of interest, there may be other positive outcomes associated with the Curiosity Project, such as improved social relationships, improved general confidence and a heightened awareness and interest in education more generally. This could lead, for example, to some people going onto further education in a non-STEM subject rather than quitting school. Although this type of outcome would not result in more people in STEM-related careers, it will still be a positive benefit for society. These outcomes have not been assessed in this analysis since the necessary data on factors such as social relationships were not collected as part of the surveys or were not collected in a usable format. Since there are likely to be very few, if any, negative effects of the Curiosity Project, we believe that the outcomes captured in this study represent an important but partial assessment which, all else constant will produce conservative estimates of overall social value. This, however, is not such a problem here as the main outcomes have been assessed and because it helps to counterbalance the uncertainties and possible overstatements in the other parts of the analysis.

# 7 Conclusion and next steps

To summarise the headline figures from the Siemens Curiosity Project evaluation: For £175,468 spent by Siemens on its long-term Curiosity Project programmes surveyed in this report Siemens produced a net benefit of approximately £1,909,635. The total Benefit-Cost Ratio shows that Siemens investment in these long-term Curiosity Project programmes provides a social return on investment of £11.90 for every £1 spent (BCR 11.9:1). Net benefits per student are £19.58 for long-term Curiosity Project programmes.

For £262,321 spent by Siemens on its short-term Curiosity Project programmes surveyed in this report Siemens produced a net benefit of approximately £493,947. The total Benefit-Cost Ratio shows that Siemens investment in these long-term Curiosity Project programmes provides a social return on investment of £2.9 for every £1 spent (BCR 2.9:1). Net benefits per student are £4.19 for short-term Curiosity Project programmes.

The data available to us and our current estimations suggest that the long-term programmes have a greater impact for the cost of providing them. We note that from the raw regression results, short-term projects were associated with a higher (+26 percentage-point) increase in enthusiasm for engineering, while the long-term projects we associated with only a +13 percentage point increase. However, given that we anticipated higher levels of focusing and optimism bias to occur when an individual is surveyed via a pad at a Science Fair or other event, we applied higher bias corrections to the short-term programmes. As a result, the set of benefits attributed to the Curiosity Project was higher for long-term than short-term programmes in the final impact evaluation.

Based on these findings, it may be recommended that Siemens focus future Curiosity Project funding on maximising the impact of their STEM programmes through longer-term interventions, which have been shown to be most effective from a cost-benefit perspective.

We consider these impact figures to be realistic and defensible in the light of impact evaluations in other sectors. In previous impact evaluation reports we have taken BCR below 40:1 to be realistic (e.g. Sainsbury Social Impact Report 2017).

Building on the previous iteration of the Curiosity Project impact assessment (2016-2017), this evaluation aimed at correcting some of the limitations of the previous study. Improvement from the previous model was threefold:

- Use of control group to strengthen the robustness of the results by considering a counterfactual of what would have happened without the Curiosity Project (previous iterations used deadweight values, which is a less sensitive way of accounting for counterfactuals (both approaches are provided for by HM Treasury Green Book);
- More precision in biases applied to short and long-term programmes to account for the greater framing and optimism biases that are expected to operate on short-term projects like Science Festivals.
- More realistic counterfactuals around highest level of qualification based on national averages from UK-wide datasets.

There are still ways in which Siemens can enhance the robustness and sensitivity of the impact evaluation for the Curiosity Project. Going forward there are three key areas of research to develop and improve the evaluation framework:

i. Impact measurement

Data collection techniques can be improved to maximise our ability to measure cause and effect relationships (i.e. impact) in a rigorous way. Firstly, future analysis should replicate the study here by focussing on quasi-experimental methods. In this case future surveys should be aimed at collecting data over time (pre and post surveys) and for different groups (i.e., points ABCD in Figure 3.a Potential data available in relation to an intervention Figure 3.a) and for good sample sizes that have been chosen at random (random sampling) if not all of the participants are included in the survey. This will permit a much higher level of rigour in the impact analysis through use of methods such as DiD analysis allowing us to much better understand causal impacts over time.

One of the limitations of the current study was the low sample size obtained via the web link survey. This is partly attributed to the late running of the survey towards the end of the summer term. This was due to difficulties of obtaining data protection assurances from all data collection platforms. In the future, response rates could be improved by preparing respondents (informing them that there will be surveys in the coming school year) and by surveying students earlier in the year, when exam preparation and end of term fatigue will be lower.

For even greater levels of causal inference, the potential and feasibility of running some experiments should be assessed. Here different interventions or different parts of an intervention can be assigned randomly across a sample of students to test the true causal effect on outcomes. Experiments are becoming increasingly cheaper to administer and also a number of methods have been developed to address and solve for the well-known ethical issues that may arise in some study settings.

ii. Programme outcomes

The main outcomes evaluated in this report were based on enthusiasm for engineering careers. The development of pre/post surveys (see above) would enable us to evaluate other outcomes like actual educational attainment and selection of STEM subjects (at GCSE/A-level/university level).

The retrospective nature of the present study meant that it was not possible to assess impact on these outcomes in this report, but these outcomes would be amenable to monetary valuation. This could include measurement of primary benefit values related to improvements in people's wellbeing, which is a more substantial task and would require additional data collection and analysis. This will provide a more holistic assessment of the social value of the Curiosity Project and a more rigorous and informative CBA study.

iii. Valuation

The collection of a wider range of outcomes will permit a fuller assessment of social value because it will permit the inclusion of a wider range of benefits in the CBA calculation. For example, future surveys could include measures and questions on social relationships and friendships developed as a part of the Curiosity Project. These outcomes can be valued separately and incorporated into the CBA study alongside the earnings premia benefits that have been calculated here.

The methodologies applied here and those suggested for future research can be applied to all areas of community investment within Siemens and thus future research could also focus on using these methods in other Siemens projects.

# 8 Annex: Research design and validity

The validity of causal inference studies, or more generally, of any research design is built upon four pillars:

**Construct validity** concerns to the degree to which the particulars of a study - such as the samples, outcomes and settings - are adequate representations of the constructs that we are interested in. For example, we may be interested in measuring outcomes such as the development and educational progress of a teenager. In this case, we must consider whether an assessment of the impacts of the intervention on their GCSE exam scores suffice, or whether we measure educational progress and attainment in more ways than just through GCSE scores.

**Internal validity** concerns the extent to which we can conclude that the estimated relationship between the intervention and the outcome(s) is a causal one. The confidence with which we can state that the intervention caused the outcome(s) of interest is based on three theoretical criteria: whether the cause precedes the effect, whether the cause is correlated with the effect, and whether there is no plausible alternative explanation for the observed correlation. This can be overcome through careful study designs and advanced statistical estimation methods.

**Statistical validity** relates to the statistical methods used in the causal analysis. This will cover things such as whether standard errors have been estimated correctly and whether the functional form of the model is correct. Statistical validity is independent of internal validity in that we can have a method with very high levels of internal validity but which has been employed poorly and thus is not statistically valid. A robust causal inference study is one that is both internally and statistically valid.

**External validity** is about whether the estimated cause and effect relationships are generalisable to other populations and settings outside of the study in question. This is important in situations where the effects of the intervention impact in a heterogeneous (different) way on different population groups. Having a sample that is not representative of the target population or using indicators which capture collinear effects (such as local area effects) may undermine the external validity of a study.

# 9 Annex: Tables and figures

## 9.1 Data

### *9.1.1 Long-term programmes*

*Figure 9.a Gender*

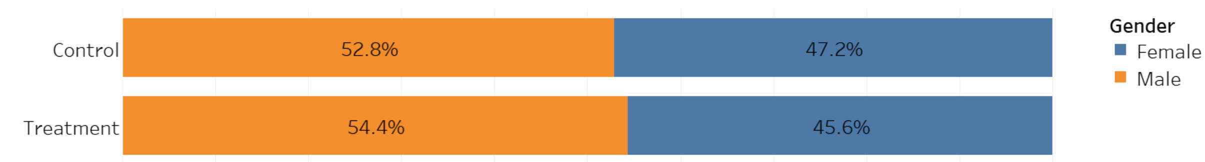
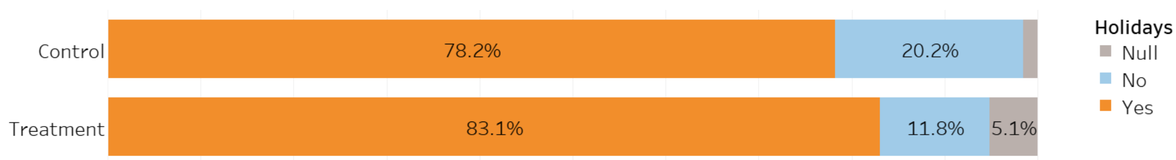
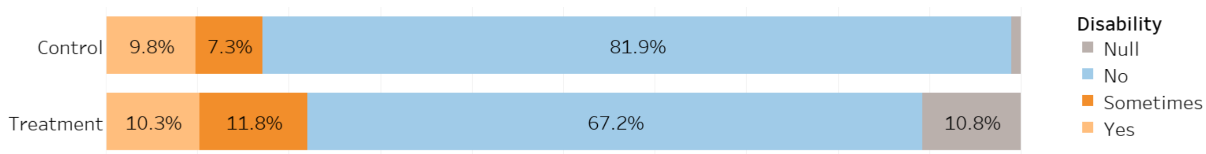



*Figure 9.b Go on family holiday every year*




Note: 1.6% of the respondents in the control group and 5.1% in the treatment group reported "Don't know / Rather not say"

*Figure 9.c Disability*




Note: 1% respondents in the control group and 10.8% in the treatment group reported "Don't know / Rather not say"

*Figure 9.d Taking part in other science activities*

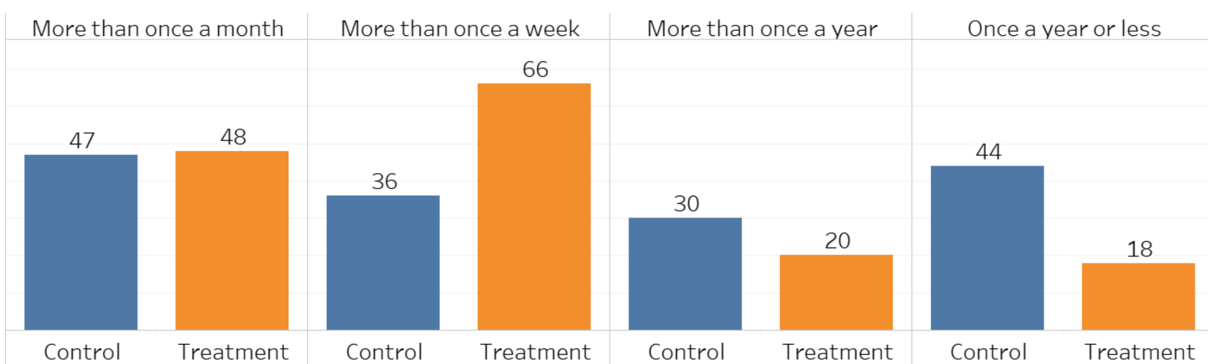



*Table 9.1 Descriptive statistics for long-term interventions data*

| Characteristic | Treatment (T) | | Control (C) | | Diff. (T-C) | Statistical significance (p-value) |
|---|---|---|---|---|---|---|
| | N | Average | N | Average | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| Statement - Engineering/science is an interesting career | 187 | 3.29 | 128 | 2.91 | 0.38 | 0 |
| Statement - Foreign languages are an interesting career | 187 | 2.52 | 128 | 2.49 | 0.03 | 0.76 |
| Statement - What I have learnt in class helped me understand something I didn't | 187 | 3.37 | 128 | 3.28 | 0.09 | 0.22 |
| Statement - I have a better idea of what engineering is and how scientists and engineers make a difference to my life | 187 | 3.21 | 128 | 2.88 | 0.34 | 0 |
| Statement - I enjoy the class and would like to take part in similar science/technology/maths activities | 187 | 3.21 | 128 | 3.06 | 0.15 | 0.09 |
| Statement - I have learned new things | 187 | 3.49 | 128 | 3.39 | 0.1 | 0.15 |
| Statement - I could use what I have learned | 187 | 3.33 | 128 | 3.23 | 0.1 | 0.18 |
| Statement - I get good marks in mathematics | 187 | 3.19 | 128 | 3.2 | 0 | 0.98 |
| Statement - Mathematics is one of my best subjects | 187 | 3.03 | 128 | 2.94 | 0.09 | 0.41 |
| Statement - I have always done well in mathematics | 187 | 3.01 | 128 | 2.95 | 0.05 | 0.61 |
| Age | 195 | 13.9 | 193 | 13.99 | -0.1 | 0.46 |
| school_year==Year 7 | 195 | 0.23 | 193 | 0.11 | 0.12 | 0 |
| school_year==Year 8 | 195 | 0.23 | 193 | 0.22 | 0 | 0.95 |
| school_year==Year 9 | 195 | 0.42 | 193 | 0.18 | 0.24 | 0 |
| school_year==Year 10 | 195 | 0.07 | 193 | 0.23 | -0.16 | 0 |
| school_year==Year 11 | 195 | 0.05 | 193 | 0.25 | -0.2 | 0 |
| Female | 195 | 0.46 | 193 | 0.47 | -0.02 | 0.77 |
| disab==No | 174 | 0.75 | 191 | 0.83 | -0.07 | 0.08 |
| disab==Sometimes | 174 | 0.13 | 191 | 0.07 | 0.06 | 0.06 |
| disab==Yes | 174 | 0.11 | 191 | 0.1 | 0.02 | 0.63 |
| Holidays | 185 | 0.88 | 190 | 0.79 | 0.08 | 0.03 |
| excur==Once a year or less | 152 | 0.12 | 157 | 0.28 | -0.16 | 0 |
| excur==More than once a year | 152 | 0.13 | 157 | 0.19 | -0.06 | 0.16 |
| excur==More than once a month | 152 | 0.32 | 157 | 0.3 | 0.02 | 0.76 |
| excur==More than once a week | 152 | 0.43 | 157 | 0.23 | 0.2 | 0 |
| region==North West | 195 | 0.34 | 193 | 0.13 | 0.21 | 0 |
| region==North East | 195 | 0.04 | 193 | 0.06 | -0.02 | 0.47 |
| region==East Midlands | 195 | 0.03 | 193 | 0.12 | -0.09 | 0 |
| region==East of England | 195 | 0.12 | 193 | 0.11 | 0.01 | 0.66 |
| region==Yorkshire and the Humber | 195 | 0 | 193 | 0.07 | -0.07 | 0 |
| region==Greater London | 195 | 0.05 | 193 | 0.1 | -0.05 | 0.05 |
| region==South East | 195 | 0.09 | 193 | 0.19 | -0.1 | 0 |
| region==South West | 195 | 0.3 | 193 | 0.1 | 0.2 | 0 |
| region==West Midlands | 195 | 0.03 | 193 | 0.12 | -0.09 | 0 |

Notes: Long-term intervention sample composed of grades 7-11. Younger version dependent variables.

### *9.1.2 Short-term programmes*

*Figure 9.e Gender*

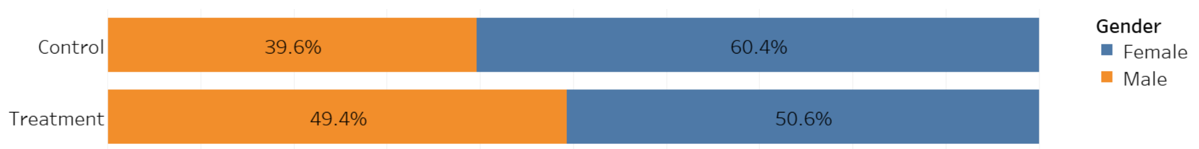
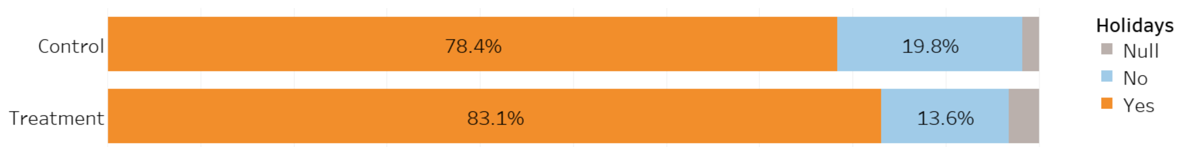
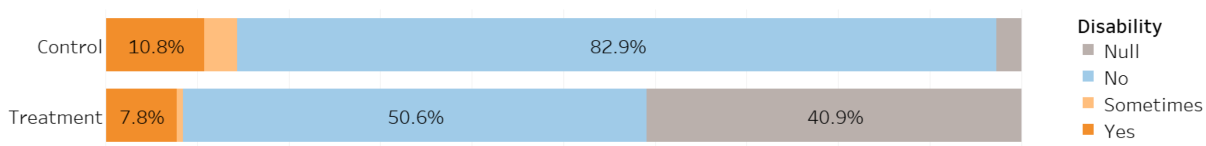



*Figure 9.f Go on family holiday every year*




Note: 1.8% of the respondents in the control group and 3.2% in the treatment group reported "Don't know / Rather not say"

*Figure 9.g Disability*




Note: 3.6% in the control group and 0.6% in the treatment group reported "Sometimes". 1% respondents in the control group and 10.8% in the treatment group reported "Don't know / Rather not say"

*Figure 9.h Taking part in other science activities*

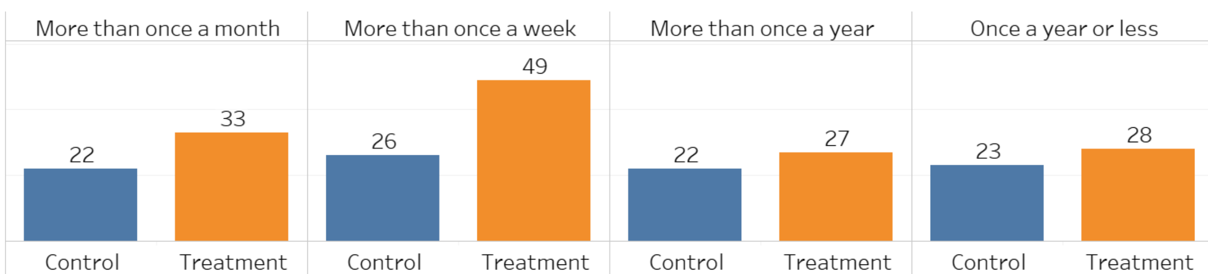



*Table 9.2 Descriptive statistics for short-term interventions data*

| Characteristic | Treatment | | Control | | Diff. | p-value |
|---|---|---|---|---|---|---|
| | N | Average | N | Average | | |
| Statement - I would consider a career in science or engineering | 99 | 2.99 | 60 | 2.62 | 0.37 | 0.02 |
| Statement - I would consider a career using foreign languages | 99 | 2.25 | 60 | 2.3 | -0.05 | 0.76 |
| Statement - I now understand how what I learnt today can be used in science or engineering | 99 | 3.1 | 27 | 3 | 0.1 | 0.49 |
| Statement - The fair has increased my knowledge of the subject and I could describe what I learned | 99 | 3.07 | 27 | 3.11 | -0.04 | 0.77 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Statement - I could use what I learned to solve a problem or apply it to a new situation | 99 | 2.92 | 27 | 3.22 | -0.3 | 0.03 |
| Statement - I now understand what science or engineering is and how science and engineering contributes to society | 99 | 3.15 | 60 | 3.08 | 0.07 | 0.56 |
| Statement - I have enjoyed the day and found the fair engaging | 99 | 3.49 | 27 | 3 | 0.49 | 0 |
| Age | 154 | 15.05 | 111 | 15.07 | -0.03 | 0.91 |
| Female | 154 | 0.51 | 111 | 0.6 | -0.1 | 0.12 |
| school_year==Year 7 | 154 | 0.18 | 111 | 0.05 | 0.12 | 0 |
| school_year==Year 8 | 154 | 0.11 | 111 | 0.16 | -0.05 | 0.22 |
| school_year==Year 9 | 154 | 0.07 | 111 | 0.16 | -0.09 | 0.02 |
| school_year==Year 10 | 154 | 0.47 | 111 | 0.15 | 0.32 | 0 |
| school_year==Year 11 | 154 | 0.06 | 111 | 0.14 | -0.09 | 0.02 |
| school_year==Year 12 | 154 | 0.09 | 111 | 0.15 | -0.06 | 0.12 |
| school_year==Year 13 | 154 | 0.02 | 111 | 0.17 | -0.15 | 0 |
| disab==No | 91 | 0.86 | 108 | 0.85 | 0.01 | 0.92 |
| disab==Sometimes | 91 | 0.01 | 108 | 0.04 | -0.03 | 0.24 |
| disab==Yes | 91 | 0.13 | 108 | 0.11 | 0.02 | 0.66 |
| excur==Once a year or less | 137 | 0.2 | 93 | 0.25 | -0.04 | 0.44 |
| excur==More than once a year | 137 | 0.2 | 93 | 0.24 | -0.04 | 0.48 |
| excur==More than once a month | 137 | 0.24 | 93 | 0.24 | 0 | 0.94 |
| excur==More than once a week | 137 | 0.36 | 93 | 0.28 | 0.08 | 0.22 |

Notes: Short-term intervention sample composed of grades 7-13. Older version dependent variables.

## 9.2 Results

### *9.2.1 Long-term programmes*

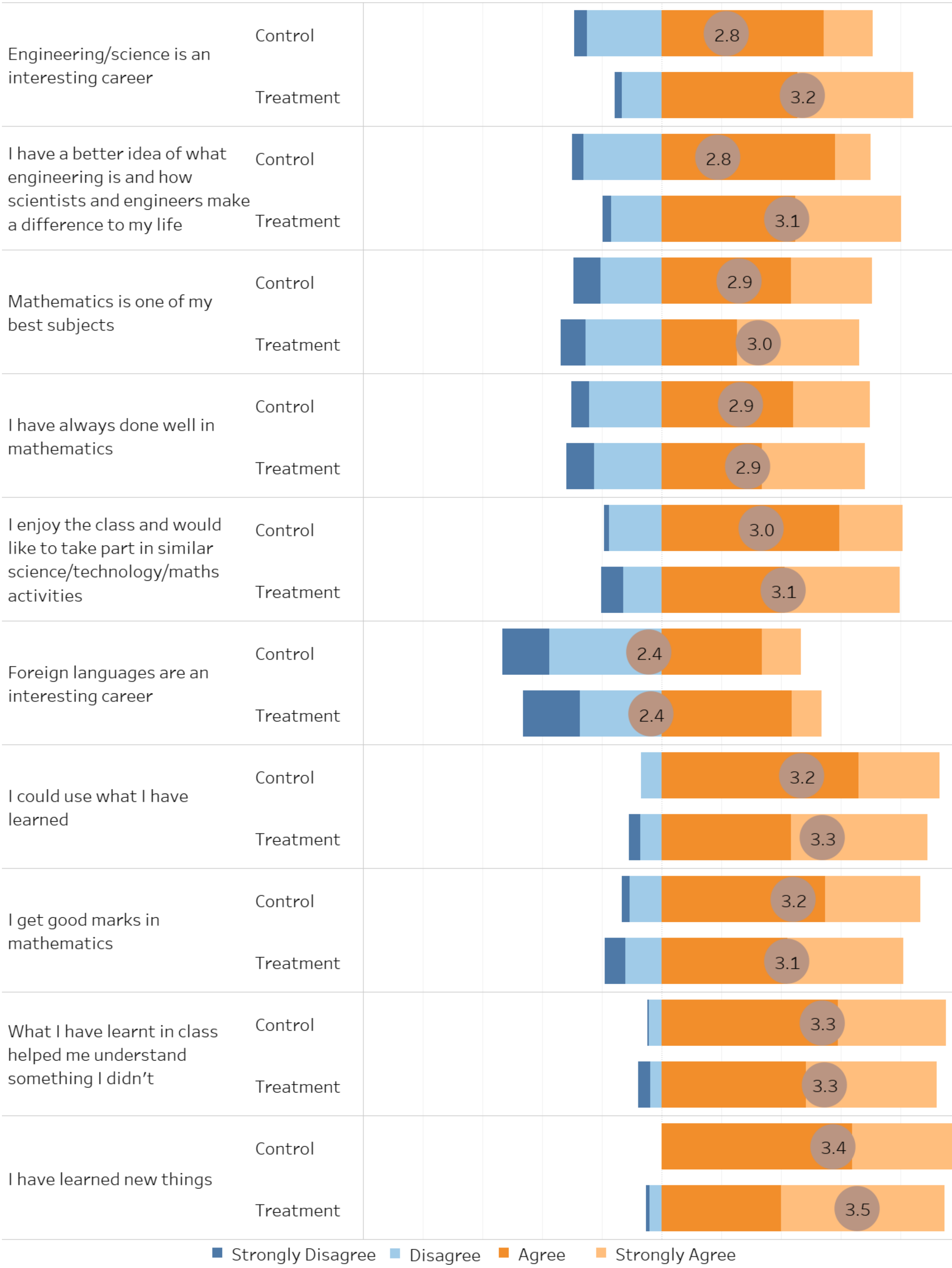
*Table 9.3 Level of agreement with various statements*

| | Treatment | | | | Control | | | |
|---|---|---|---|---|---|---|---|---|
| | Strongly disagree | Disagree | Agree | Strongly agree | Strongly disagree | Disagree | Agree | Strongly agree |
| Engineering/science is an interesting career | 3 | 19 | 85 | 80 | 3 | 28 | 74 | 23 |
| Foreign languages are an interesting career | 5 | 25 | 82 | 75 | 3 | 29 | 77 | 19 |
| What I have learnt in class helped me understand something I didn't | 14 | 44 | 52 | 77 | 9 | 26 | 57 | 36 |
| I have a better idea of what engineering is and how scientists and engineers make a difference to my life | 14 | 39 | 66 | 68 | 6 | 30 | 56 | 36 |
| I enjoy the class and would like to take part in similar | 11 | 17 | 80 | 79 | 1 | 19 | 79 | 29 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| science/technology/maths activities | | | | | | | | |
| I have learned new things | 30 | 50 | 86 | 21 | 16 | 52 | 41 | 19 |
| I could use what I have learned | 7 | 13 | 78 | 89 | 0 | 8 | 83 | 37 |
| I get good marks in mathematics | 11 | 19 | 80 | 77 | 2 | 14 | 69 | 43 |
| Mathematics is one of my best subjects | 5 | 7 | 88 | 87 | 1 | 7 | 75 | 45 |
| I have always done well in mathematics | 2 | 9 | 72 | 104 | 0 | 0 | 78 | 50 |

*Note: Statements displayed to 11-14-years-old.*

*Figure 9.i Likert graph of statements*



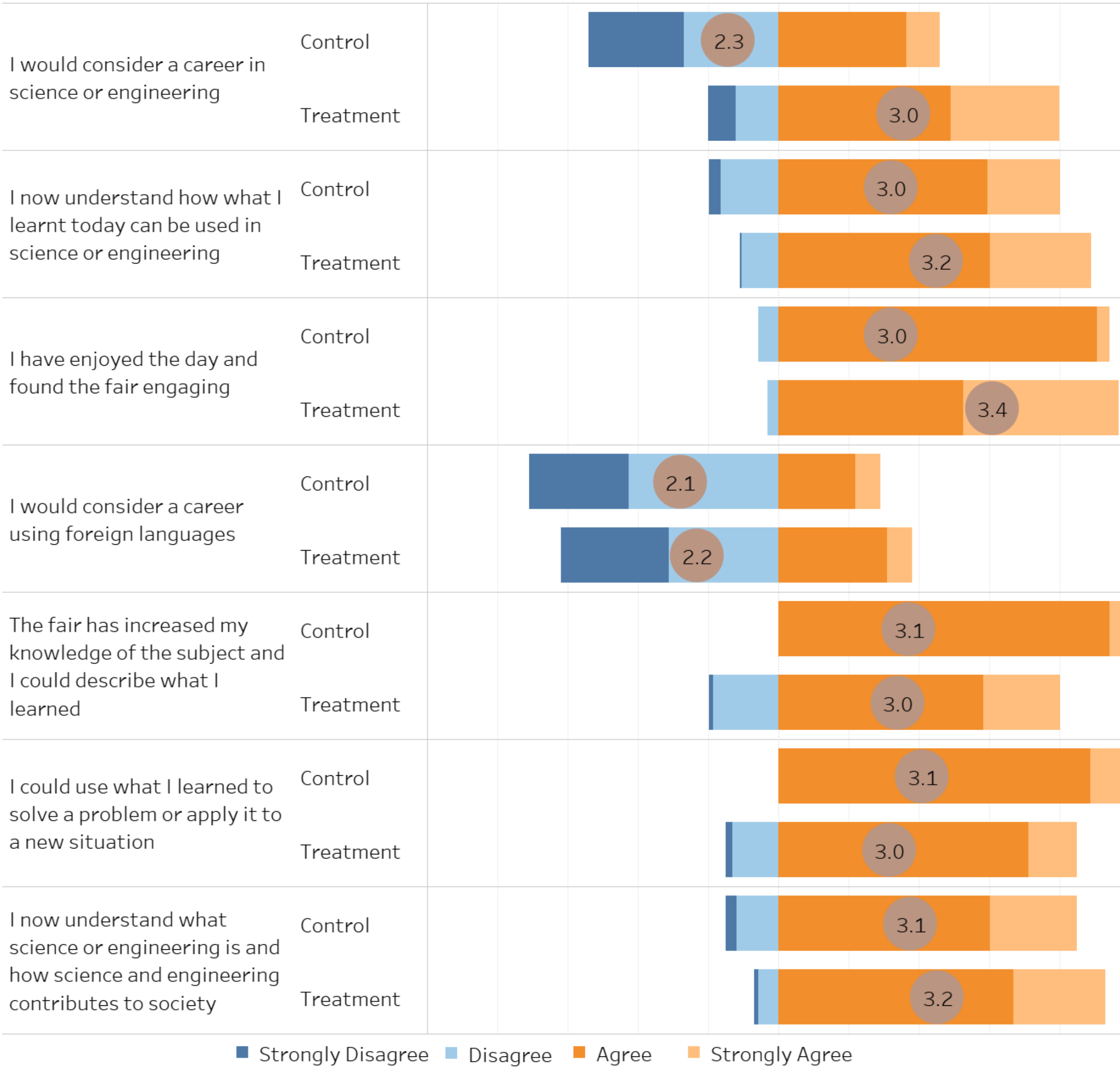
### 9.2.2 *Short-term programmes*

*Table 9.4 Level of agreement with various statements*

| | Treatment | | | | Control | | | |
|---|---|---|---|---|---|---|---|---|
| | Strongly disagree | Disagree | Agree | Strongly agree | Strongly disagree | Disagree | Agree | Strongly agree |
| I would consider a career in science or engineering | 8 | 17 | 42 | 32 | 12 | 13 | 21 | 14 |
| I now understand how what I learnt today can be used in science or engineering | 1 | 14 | 58 | 26 | 1 | 3 | 18 | 5 |
| I have enjoyed the day and found the fair engaging | 0 | 2 | 46 | 51 | 0 | 3 | 21 | 3 |
| I would consider a career using foreign languages | 24 | 35 | 31 | 9 | 14 | 23 | 14 | 9 |
| The fair has increased my knowledge of the subject and I could describe what I learned | 2 | 14 | 58 | 25 | 0 | 0 | 24 | 3 |
| I could use what I learned to solve a problem or apply it to a new situation | 3 | 17 | 64 | 15 | 0 | 0 | 21 | 6 |
| I now understand what science or engineering is and how science and engineering contribute to society | 2 | 8 | 62 | 27 | 3 | 8 | 30 | 19 |

*Note: Statements displayed to 15-18-years-old.*

*Figure 9.j Likert graph of statements*




## 9.3 Cost-benefit analysis

*Table 9.5 Flow of present value benefits for each degree level at age 22*

| | Doctorate | Taught postgraduate | First degree |
|---|---|---|---|
| Year 1 | | | £4,137 |
| Year 2 | | £2,446 | £3,997 |
| Year 3 | | £2,364 | £3,862 |
| Year 4 | | £2,284 | £3,732 |
| Year 5 | £470 | £2,207 | £3,605 |
| Year 6 | £454 | £2,132 | £3,483 |
| Year 7 | £439 | £2,060 | £3,366 |
| Year 8 | £424 | £1,990 | £3,252 |

| | | | |
|---|---|---|---|
| Year 9 | £410 | £1,923 | £3,142 |
| Year 10 | £396 | £1,858 | £3,036 |
| **Total** | **£2,594** | **£19,263** | **£35,612** |

# 10 Annex: Gender interaction model

*In this annex, we will focus on the possible role that gender plays on the impact of long-term Curiosity Project programmes.*

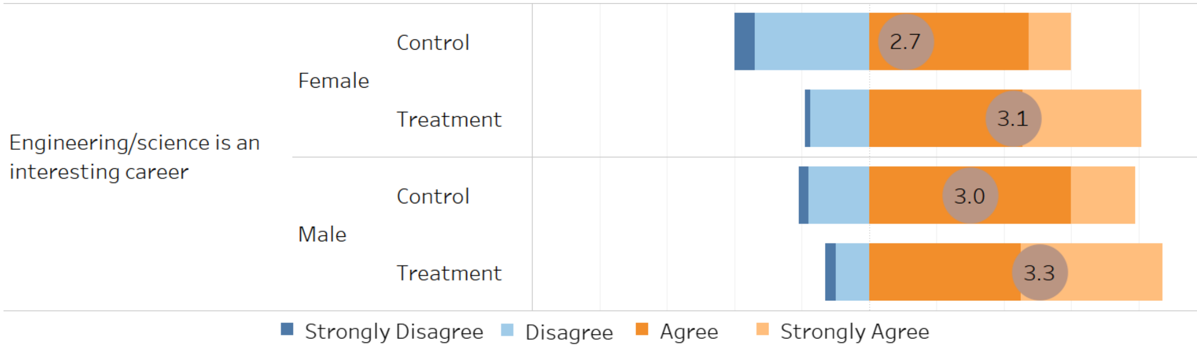
The 2017 Engineering UK Report stated that *"Efforts to attract girls and women into engineering are falling short: today less than 1 in 8 of the engineering workforce is female; boys are 3.5 times more likely to study A-level Physics (in England, Wales and Northern Ireland) than girls; and five times more likely to gain an engineering and technology degree."*.

If males in society are more likely to go into science education and careers (for cultural, social or any other reasons), then this trend could also affect our data. It may be that male student' 'self-selected' into the Curiosity Project programmes because of a higher initial interest in science, engineering, or maths. Conversely, it might be more difficult to recruit female students, meaning there will be a negative selection effect against female participants.

We explored differences in the proportion of males within our treatment groups compared to the national control group. The gender composition between the control and treatment groups is nearly identical (with around 47% female in both), with this difference being statistically insignificant (see Table 9.1). We therefore conclude that the long-term Curiosity Project programmes succeeded in recruiting both male and female participants and do not suffer from gender unbalance.

We explored the difference in levels of enthusiasm between the Curiosity Project treatment group and the national control group for both male and female participants. Note that it may be that we observe a higher level of enthusiasm for men consistently in both the treatment and control group, because of higher exiting levels of interest. The question is whether the Curiosity Project has a more beneficial effect on female participants when compared to females in the national control group.

*Figure 10.a Enthusiasm towards a career in science or engineering between gender: Long-term programmes*




In Figure 10.a above, we present the level of enthusiasm for both treatment and control respondents and each gender. We see enthusiasm for a career in engineering or science is lower overall for females than males in both the treatment group and the control group. On average the treatment group has higher enthusiasm for a career in engineering or science.

To test this hypothesis, we used the same regression model as applied in [section analysis] and added an interaction term between gender and the treatment variable. This will allow to test if the impact of the programme has a differentiated effect for male and female participants.

*Table 10.1 Interaction model with gender*

| Variable | Coeff. |
|---|---|
| Treatment | 1.331*** |
| Female | -0.780* |
| Treatment # Female | 0.410 |
| Age | -0.154 |
| Year 7 | 0 |
| Year 8 | -0.0716 |
| Year 9 | 0.0654 |
| Year 10 | -0.0634 |
| Year 11 | 0.292 |
| No | 0 |
| Sometimes | -0.667 |
| Yes | -0.522 |
| No | 0 |
| Yes | 0.318 |
| Once a year or less | 0 |
| More than once a year | -0.223 |
| More than once a month | 0.652 |
| More than once a week | 1.190*** |
| North West | 0 |
| North East | 0.369 |

| | |
|---|---|
| East Midlands | 0.453 |
| East of England | 0.142 |
| Yorkshire and the Humber | -0.497 |
| Greater London | 0.413 |
| South East | 0.257 |
| South West | -1.077** |
| West Midlands | 0.281 |
| cut1 | -5.660** |
| cut2 | -3.321 |
| cut3 | -0.0899 |
| Observations | 234 |

*Notes: Odds ratio calculated using an ordered logit regression. *** <1% significance; ** <5% significance; * <10% significance*

The results in Table 10.1 suggests that taking part in the long-term Curiosity Project programmes is positively and significantly associated with higher enthusiasm towards a career in science or engineering across the whole sample, but that being female doesn't have any significant interaction with the treatment (either positive or negative). In other words, the impact of the programme is not statistically different for male and female participants.

In conclusion, no gender specific effects were identified from the long-term Curiosity Project date, suggesting that the benefits of the Curiosity Project programme are felt equally by both male and female students, after controlling for initial differences in STEM enthusiasm.

# 11 Annex: Exchequer CBA

We estimate benefits to the exchequer via the additional income tax and social contributions implied by the increased earnings of an engineering career.

Using the OECD "Benefits and Wages: Tax-Benefit calculator", we obtain estimates for income taxes and social contributions in 2015.

*Table 11.1 Income taxes and social contributions: Highest education level*

| Degree level | Subject | Salary | Fiscal base | Income Tax | Social Contributions |
|---|---|---|---|---|---|
| **Doctorate** | **Engineering** | £34,126 | £33,790 | £4,638 | £3,088 |
| | **All subjects** | £33,577 | £33,431 | £4,566 | £3,044 |
| **Taught postgraduate** | **Engineering** | £28,679 | £28,398 | £3,560 | £2,441 |
| | **All subjects** | £26,103 | £25,882 | £3,056 | £2,139 |
| **First degree** | **Engineering** | £26,141 | £25,882 | £3,056 | £2,139 |
| | **All subjects** | £21,932 | £21,928 | £2,266 | £1,664 |

*Source: All values presented are in 2015 prices. OECD Benefits and wage: Tax-Benefit calculator. http://www.oecd.org/els/soc/benefitsandwagestax-benefitcalculator.htm*

Exchequer premia of an engineering career are uprated to 2016 values using the Bank of England inflation rate of 1.7% for 2015/16.

*Table 11.2 Premium to the exchequer: Highest education level*

| Degree level | Income Tax premium | Social Contributions premium | Total premium |
|---|---|---|---|
| Doctorate | £73 | £45 | £118 |
| Taught postgraduate | £513 | £307 | £820 |
| First degree | £804 | £483 | £1,287 |

The flow of discounted benefits is calculated based on the expected time frames of each degree level. We apply Green Book discounting rates at 3.5% to calculate net present value (NPV).

*Table 11.3 Flow of present value secondary benefits for each degree level at age 22*

| | Doctorate | Taught postgraduate | First degree |
|---|---|---|---|
| Year 1 | | | £1,243 |
| Year 2 | | £765 | £1,201 |
| Year 3 | | £740 | £1,161 |
| Year 4 | | £715 | £1,122 |
| Year 5 | £99 | £690 | £1,084 |
| Year 6 | £96 | £667 | £1,047 |
| Year 7 | £93 | £645 | £1,012 |
| Year 8 | £90 | £623 | £977 |
| Year 9 | £87 | £602 | £944 |
| Year 10 | £84 | £581 | £912 |
| **Total** | **£548** | **£6,027** | **£10,703** |

Following the same approach taken in the previous section, we calculate the average annual Exchequer contributions at the time of graduating at £1,590[14].

*Table 11.4 Estimated benefits associated with enhanced enthusiasm for engineering career*

| | Total population | Difference in proportion of students enthusiastic | Bias correction (proportion of impacted | Average wage premium (£) | Discounting | Lifetime net present value benefits (£) |
|---|---|---|---|---|---|---|

[14] 1% will do a doctorate and generate a net present value of £464 in secondary benefits over ten years. 9% will do a post graduate and can expect a NPV of £5,446. Finally, 10% will stop at a first degree and get a NPV of £11,078

| | | | about engineering (T vs. C) | | population) | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Long-term programmes | 97545 | × | 13% | × | 4% | × | £1,590 | ÷ | $1.035^7$ | = | £663,708 |
| Short-term programmes | 117932 | × | 26% | × | 0.6% | × | £1,590 | ÷ | $1.035^7$ | = | £229,846 |

*T = Treatment group (Curiosity Project participants); C = Control group. Net present value (NPV) calculated using Green Book discounting rates at 3.5%.*

The Exchequer CBA represents the cost-benefit results for the public purse accounting for impacts on the Exchequer. We calculate total costs as project costs, plus an additional opportunity cost of 8% (Jenkins and Kuo 2007). We also calculate net secondary benefits of the project (overall Exchequer benefits - costs).

*Table 11.5 Cost & secondary benefits: Long-term Curiosity Project activities*

| Long-term projects | Lifetime benefits |
|---|---|
| **Secondary benefits** | |
| Total secondary benefits | £633,708 |
| Average secondary benefits (per student) | £6.50 |
| **Costs** | |
| Total projects' costs including opportunity costs | £175,468 |
| Average costs (per student) | £1.80 |
| **Cost-benefit analysis** | |
| Net benefits of the projects | £458,240 |
| Benefit-cost ratio | 3.6 |

*Table 11.6 Cost & secondary benefits: Short-term Curiosity Project activities*

| Short-term projects | Lifetime benefits |
|---|---|
| **Secondary benefits** | |
| Total secondary benefits | £229,846 |
| Average secondary benefits (per student) | £1.95 |
| **Costs** | |
| Total projects' costs including opportunity costs | £262,321 |
| Average costs (per student) | £2.22 |
| **Cost-benefit analysis** | |
| Net benefits of the projects | - £32,475 |
| Benefit-cost ratio | 0.9 |

Net secondary benefits of the long-term Curiosity Projects are £458,240 over ten years. The benefit-cost ratio is 3.6.

Net secondary benefits of the short-term Curiosity Projects are - £32,475 over ten years. The benefit-cost ratio is 0.9.

# 12 Annex: Educational Materials

## 12.1 Context of the study

Siemens has created a suite of comprehensive Key Stage 2, 3 and 4 teachers' guides, Schemes of Work, lesson plans and practical activities for teachers to use.

Each Scheme of Work (topic) is supported by a corresponding supporting lesson plan PowerPoint presentation ready for use on the interactive whiteboard and designed to aid classroom delivery. Each presentation is inclusive of relevant video and photographic stimulus.

In each Scheme of Work students are presented with a problem to solve. Using the supporting lesson plan PowerPoint presentations, Student Support and Work Sheets teachers are able to provide ideas and stimulus drawn from contemporary contexts. Pupils are challenged to use primary and secondary research skills and respond in a variety of ways by investigating, researching and presenting ideas.

The topics are presented in ways that can be customised by the teacher according to the required focus. draw upon the content of maths, science and technology curricula and are designed to encourage students to see how these areas are interrelated.

In line with our evaluation of the broader Curiosity Project above, the programmes using Siemens Educational Materials were either long-term or short-term.

- **Long-term interventions** – those using Educational Materials in regular lessons over an entire school term/year - we collected pre and post survey data at two points in time, before the summer term (mostly in January 2018) and at the end (June-July 2018).
- **Short-term interventions** – where Educational Materials were applied in a single day of the school term - data was collected only once, as a retrospective survey at the end of the day that the Educational Materials were used.

Our outcomes of interest are a change in the pupils' responses to agreement questions about science, maths, technology and engineering (in line with those used to evaluate Curiosity Projects in the above report).

### 12.2 Long-term

To evaluate the long-term interventions using Educational Materials, we used a sample of 47 respondents. These pupils were in school years 7 or 8, from four different English regions and with various levels of pre-existing interest towards science and indicators of social diversity. Table 12.1 presents the characteristics of the long-term sample in more detail.

*Table 12.1 Long-term educational materials sample characteristics*

| Characteristics | Percent |
|---|---|
| Female | 38.30% (18/47) |
| School year: Year 7 | 51.06% (24/47) |
| School year: Year 8 | 48.94% (23/47) |
| Get help due to disability: No | 57.50% (23/40) |
| Get help due to disability: Sometimes | 25.00% (10/40) |
| Get help due to disability: Yes | 17.50% (7/40) |
| Go on a family holiday away from home every year | 82.93% (34/41) |
| Freq. of science activities outside of school: Don't know / Rather not say | 42.55% (20/47) |
| Freq. of science activities outside of school: Once a year or less | 14.89% (7/47) |
| Freq. of science activities outside of school: More than once a year | 4.26% (2/47) |
| Freq. of science activities outside of school: More than once a month | 12.77% (6/47) |
| Freq. of science activities outside of school: More than once a week | 25.53% (12/47) |
| Region: North West | 29.79% (14/47) |
| Region: East Midlands | 29.79% (14/47) |
| Region: East of England | 19.15% (9/47) |
| Region: South West | 21.28% (10/47) |

Using multivariate analysis (as outlined in detail in Section 3.4), we analysed the change over time in agreement with multiple statements of STEM interest, controlling for the demographic factors presented in Table 12.1. The one difference between this model compared to that presented in Section 3.4 is in the point of comparison (reference group) used. In the main section of the report as described in Section 3.4, we compared outcomes of a treatment group to outcomes from a similar control group gathered from the national population. For Educational Materials in the long-term, we do not use a control group but instead compare outcome levels within the same sample over time, using pre and post surveys collected at two points in time.

Table 12.2 presents the percentage point difference in agreement that we could attribute to the intervention alone, removing any effect that could be due to inner difference between the two groups (such as age, gender, school year, going yearly on holiday, having a disability, taking part in similar activities or region):

- The long-term use of Educational Materials is significantly associated with a 25-percentage point increase in agreement with the following statement: "I have a better idea of what engineering is and how scientists and engineers make a difference to my life".
- The analysis finds, however, that the use of long-term Educational Materials intervention is not significantly associated with an increased level of enthusiasm towards a career in science or engineering.

This result is suggestive that the long-term Educational Materials intervention does have a STEM impact, in terms of improving pupils' understanding of science and engineering, albeit that we do not detect any corresponding effect on our key outcome of interest (enthusiasm towards a career in STEM), as analysed in the Curiosity Project Evaluation (above).

*Table 12.2 Difference in agreement (in percentage points, ordered logit)*

| Statement for younger students | Percentage point difference in agreement (Agree & Strongly Agree) |
|---|---|
| Engineering/science is an interesting career | 6 |
| Foreign languages are an interesting career | 0 |
| What I have learnt in class helped me understand something I didn't | 1 |
| I have a better idea of what engineering is and how scientists and engineers make a difference to my life | 25 (*) |
| I enjoy the class and would like to take part in similar science/technology/maths activities | -2 |
| I have learned new things | 5 |
| I could use what I have learned | 2 |
| I get good marks in mathematics | 7 |
| Mathematics is one of my best subjects | 11 |
| I have always done well in mathematics | 11 |

*Notes: Proportions are calculated using ordered logit regression controlling for age, school year, gender, going yearly on holiday, has a disability, take part in similar activities and region. *** <1% significance; ** <5% significance; * <10% significance*

The results in Table 12.2 reports changes in STEM outcomes which attributable to the Educational Materials programme. We may also be interested in viewing the levels of agreement for each statement and how they change over time. The Figures below provide more descriptive information on pupil's responses to each of the survey questions. The figures are presented in an easy to understand format, but do not capture impact in the way that Table 12.2 does. Consequently, the differences we see in the figures below are not as statistically robust, as they may be due to inner differences between the two groups (which we can controlled for in Table 12.2, but not in the figures below).

Figure 12.a, Figure 12.b and Figure 12.c show the observed differences in agreement with each survey question. They provide descriptive visualisation of the levels of agreement/disagreement of three key statements related to educational outcomes in the Educational Materials pre and post survey:

- Between pre and post survey, there is an increase in the proportion of pupils who agree and who strongly agree with the statement: "I have a better idea of what engineering is and how scientists and engineers make a difference to my life".
- We observe an increase in the proportion of pupils who strongly agree that "Engineering/science is an interesting career", but a decrease in the proportion who only agree to this statement, and no change in the proportion who disagree/strongly disagree. This suggests that for some of those who previously had some agreement that science/engineering is an interesting career in the pre-survey, the use of Educational Materials over the long-term increased their interest in science/engineering as a career (resulting in a higher proportion of strongly agree responses in the post survey). Recall that these results are descriptive, and not robust causal analysis.
- We find little change in the level of agreement that "Foreign languages are an interesting career" between pre and post, which suggests that Siemens Educational Materials are impacting on STEM outcomes, while non-STEM outcomes remain unaffected, which accords with our prior expectations.

*Figure 12.a Agreement to the statement "I have a better idea of what engineering is and how scientists and engineers make a difference to my life"*

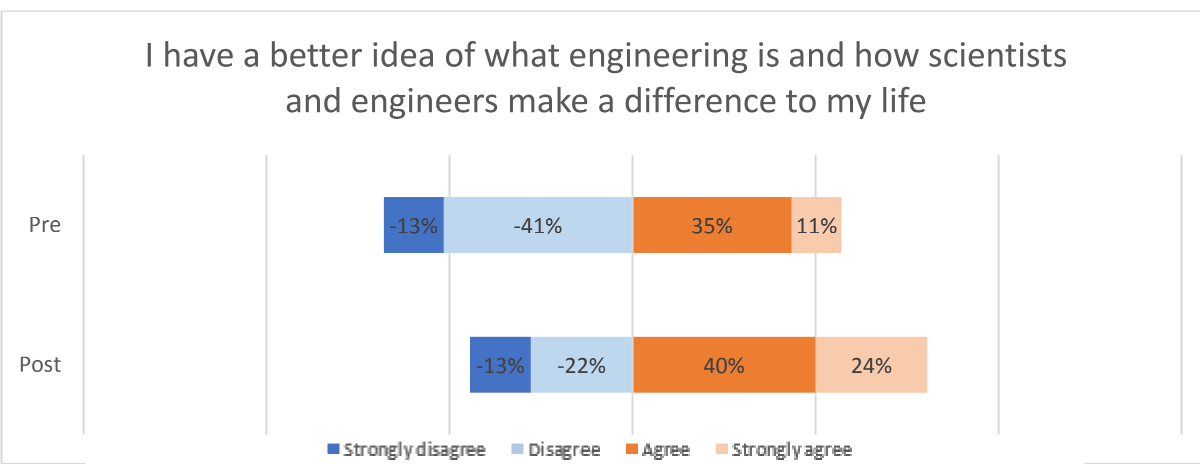


*Figure 12.b Agreement to the statement "Engineering/ science is an interesting career"*

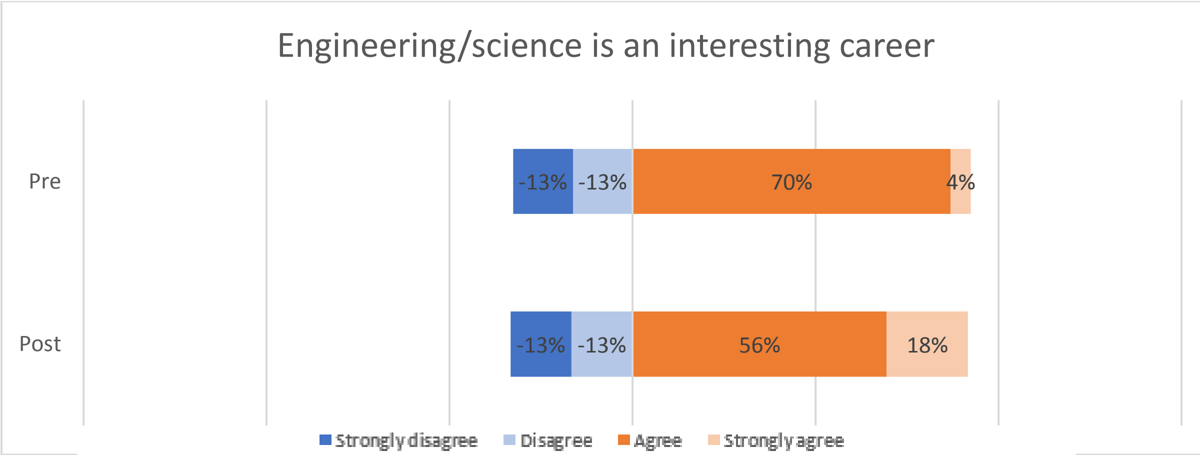
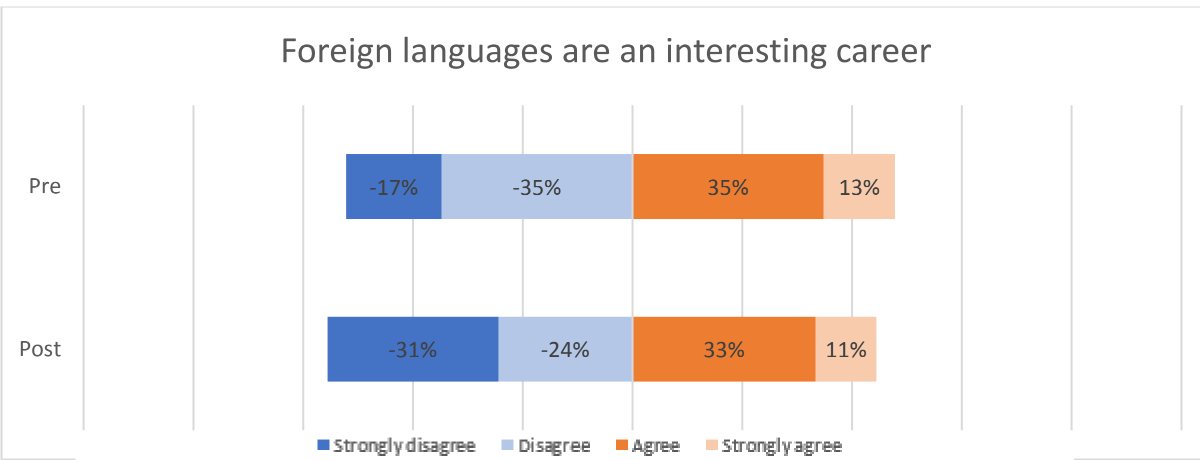



*Figure 12.c Agreement to the statement "Foreign languages are an interesting career"*




## 12.3 Short-term

To evaluate the short-term Educational Materials interventions, we used a sample of 77 respondents. 60 respondents received the short-term Educational Materials intervention, while the 17 others are a subset of our control group selected to match as closely as possible to our treatment group. These pupils were in school years 7, 8 or 9, from four different English regions and with various levels of pre-existing interest towards science. Table 12. presents the characteristics of the short-term sample in more detail. The final column indicates whether the difference between the outcomes of treatment and control samples are statistically significant.

*Table 12.3 Short-term educational materials sample characteristics*

| Characteristics | Treatment | | Control | | Difference (Treat.-Cont.) | Significant difference?[15] |
|---|---|---|---|---|---|---|
| | N | Average | N | Average | | |
| Female | 60 | 0.82 | 17 | 0.41 | 0.4 | Yes |
| School year: Year 7 | 60 | 0.4 | 17 | 0.59 | -0.19 | |
| School year: Year 8 | 60 | 0.38 | 17 | 0.24 | 0.15 | |
| School year: Year 9 | 60 | 0.22 | 17 | 0.18 | 0.04 | |
| Go on a family holiday away from home every year | 56 | 0.8 | 17 | 0.82 | -0.02 | |
| Freq. of science activities outside of school: Don't know / Rather not say | 60 | 0.27 | 17 | 0.12 | 0.15 | |
| Freq. of science activities outside of school: Once a year or less | 60 | 0.25 | 17 | 0.12 | 0.13 | |
| Freq. of science activities outside of school: More than once a year | 60 | 0.1 | 17 | 0.29 | -0.19 | Yes |
| Freq. of science activities outside of school: More than once a month | 60 | 0.23 | 17 | 0 | 0.23 | Yes |
| Freq. of science activities outside of school: More than once a week | 60 | 0.15 | 17 | 0.47 | -0.32 | Yes |
| Region: North West | 60 | 0.02 | 17 | 0 | 0.02 | |
| Region: North East | 60 | 0.13 | 17 | 0 | 0.13 | |
| Region: Greater London | 60 | 0.5 | 17 | 0.06 | 0.44 | Yes |
| Region: West Midlands | 60 | 0.35 | 17 | 0.94 | -0.59 | Yes |

For short-term Educational Materials interventions we do not have pre and post surveys with which to compare change over time. We must instead find a reference group against which to compare educational outcomes. There are two potential reference groups against which to compare the results of Educational Materials:

- A group from the national population who received no STEM-specific educational intervention (Table 12.4, column 2);
- A group of participants of who received a comparable intervention, specifically the short-term Curiosity Projects (Science Fairs) (Table 12.4, column 3).[16]

[15] Significance in the table is used to specify if the difference in mean outcome levels is statistically significant at the 95% significance level.

[16] Note that the statements asked of younger and older age groups are not directly comparable. For example, while the treatment groups saw the following statement "I have an idea of what engineering is and how scientists and engineers make a difference to my life", the control group saw "I have a better idea of what engineering is and how scientists and engineers make a difference to my life". If the control group has to react to an absolute statement, the treatment group

For analysis purposes we applied multivariate regressions (as described in Section 3.4), to identify the statistical association between the Educational Materials intervention and our key outcome variables (agreement with STEM statements) controlling for a set of demographic factors.

We are interested in outcome questions asked to the younger age group as all pupils in the short-term Educational Materials interventions surveyed were in school years 7 to 9. In Table 12.4 (column 3) we report the outcome levels of respondents in the younger age bracket involved in short-term Curiosity Project interventions, which allows us to compare between Educational Materials and Science Fairs.[17]

Compared to the short-term general population control group (school years 7-9), on average students surveyed as part of the Educational Materials short-term interventions report a significantly lower level of agreement for the four outcome statements:

- What I have learnt [today/in school] helped me understand something I didn't understand before;
- I have learned new things;
- I could use what I have learned today;
- I have [a better/an] idea of what engineering is and how scientists and engineers make a difference to my life.

Compared to the short-term Curiosity Project interventions (Science Fairs), on average students surveyed as part of the Educational Materials short-term interventions report a significantly lower level of agreement for the same four outcome statements.

In no cases do the short-term Educational Materials interventions produce significantly higher levels of agreement for any of the outcome questions compared to either the general population control group of the control group of pupils involved in Curiosity Project short-term interventions.

This suggests that the short-term uses of Educational Materials are less effective than other short-term interventions provided by the Siemens Curiosity Project (Science Fairs) in improving STEM outcomes like understanding of what engineering is as well as their willingness to take part in similar activities again. This aligns with other research, such as the British Science Association (BSA)

---

answers to a relative one. This was not an issue in the main report due to the focus on the "Engineering/science is an interesting career" which is an absolute statement and was therefore readily comparable. The two treatment groups are exposed to only marginally different statements making these comparisons more appropriate.

[17] Note that the Science Fair results reported in this section are based on the younger pupil age group, and are not comparable to those reported in Section 5.2.

evaluation of CREST– a hands-on, extra-curricular STEM project – which showed that participants achieved half a grade higher on their best science GCSE result, compared to a matched control group.[18]

Although the evidence may be suggestive that extra-curricular STEM interventions have greater impact than in-school interventions, we caveat these results by noting that the analysis is based on a small sample size, and that the effects of short-term on-day interventions are likely to be small and therefore difficult to detect, increasing the risk that our analysis may pick up noise in the data, leading to spurious negative results.

*Table 12.4 Difference in agreement (in percentage points, ordered logit)*

| Statement for younger students (Grade 7-9) | Percentage point difference in agreement (Agree & Strongly Agree) of | |
|---|---|---|
| | Short-term Educational Materials compared to general population control group | Short-term Educational Materials compared to Science Fairs |
| Engineering/science is an interesting career | -7 | -6 |
| I would consider a career using foreign languages | -11 | -9 |
| What I have learnt [today/in school] helped me understand something I didn't understand before | -19 (**) | -14 (**) |
| I have learned new things | -21 (***) | -18 (***) |
| I could use what I have learned today | -12 (*) | -12 (**) |
| I have [a better/an] idea of what engineering is and how scientists and engineers make a difference to my life | -23 (**) | -22 (**) |
| I enjoyed the [lesson/event/class] and would like to take part in similar activities | 0 | -5 |

*Notes: Proportions are calculated using ordered logit regression controlling for age, school year, gender, going yearly on holiday, take part in similar activities and region. *** <1% significance; ** <5% significance; * <10% significance*

Figure 12.d, Figure 12.e, Figure 12.f and Figure 12.g provide descriptive visualisation of the levels of agreement/disagreement with four key statements related to educational outcomes between the short-term Educational Materials and short-term Curiosity Project interventions (Science Fairs). Again, the Figures below provide more descriptive information on pupil's responses to each of the survey questions, but do not capture impact in the way that Table 12.4 does.

[18] https://edtechnology.co.uk/Article/research-emphasizes-impact-of-extra-curricular-stem

- A higher proportion of short-term Education Materials pupils disagree with the statement: "I have a better idea of what engineering is and how scientists and engineers make a difference to my life" compared to pupils surveyed at Science Fairs, while a lower proportion short-term Education Materials pupils agree or strongly agree to the statement. Recall that these results are descriptive, and not robust causal analysis.

- A higher proportion of short-term Education Materials pupils agree with the statement that "I enjoyed the [lesson/event/class] and would like to take part in similar activities" compared to pupils surveyed at Science Fairs, but a lower proportion of Education Materials pupils strongly agree with that statement. Overall, more have agreed or strongly agreed among visitors of the science fairs than among short-term beneficiaries of the Education Materials.

- A higher proportion of short-term Education Materials pupils agree with the statement that "Engineering/science is an interesting career" compared to pupils surveyed at Science Fairs, but a lower proportion of Education Materials pupils strongly agree with that statement. The overall agreement (Agree & Strongly agree) is higher among the Science Fairs pupils than in the Education Materials pupils.

- We find little difference in the level of agreement that "Foreign languages are an interesting career" between short-term Educational Materials and Science Fairs, which suggests that our surveys are detecting changes in the expected STEM outcomes, while non-STEM outcomes remain unaffected, which adds additional confidence to the differences detected between short-term Educational Materials and Science Fairs interventions.

*Figure 12.d Agreement to the statement "I have a better idea of what engineering is and how scientists and engineers make a difference to my life"*

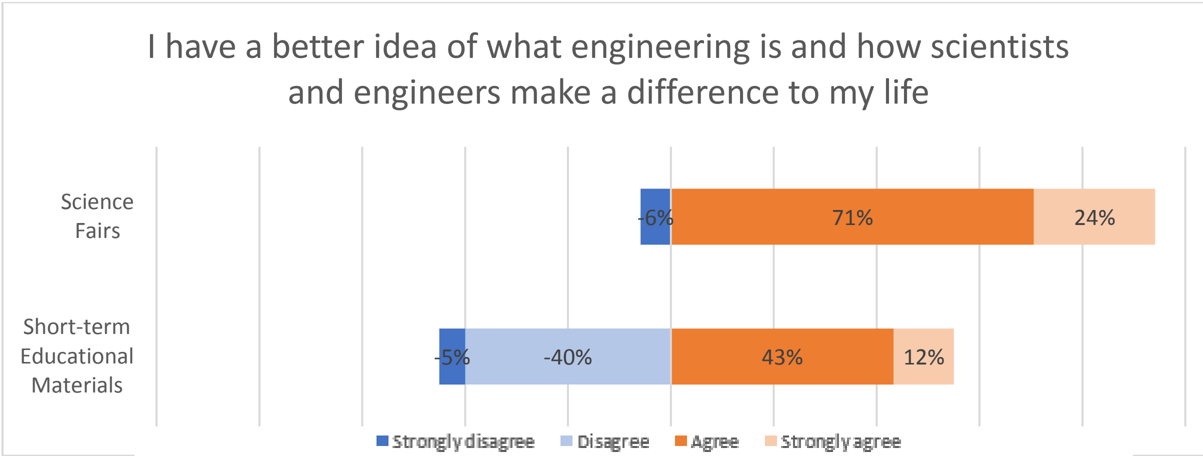



*Figure 12.e Agreement to the statement "I enjoyed the lesson and would like to take part in similar activities"*

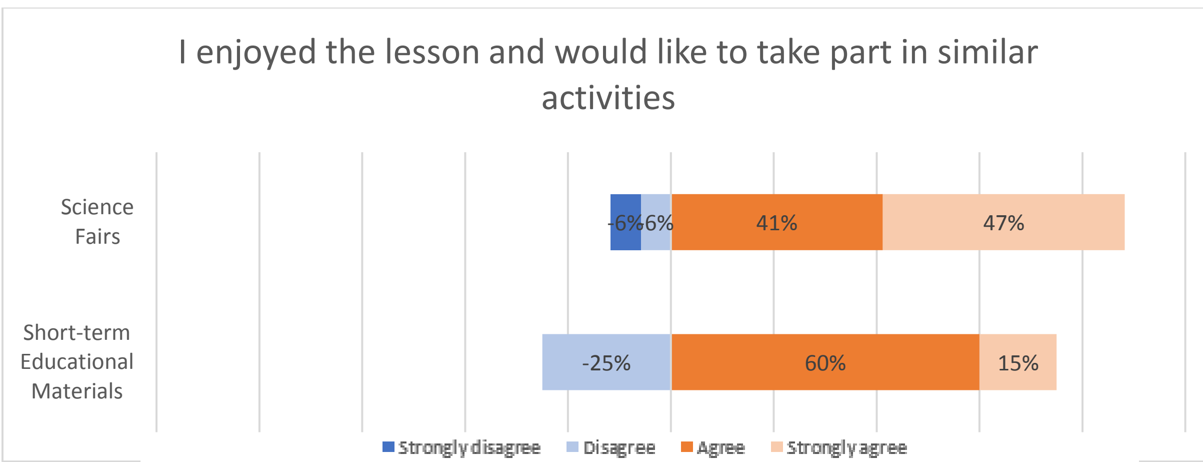



*Figure 12.f Agreement to the statement "Engineering/ science is an interesting career"*

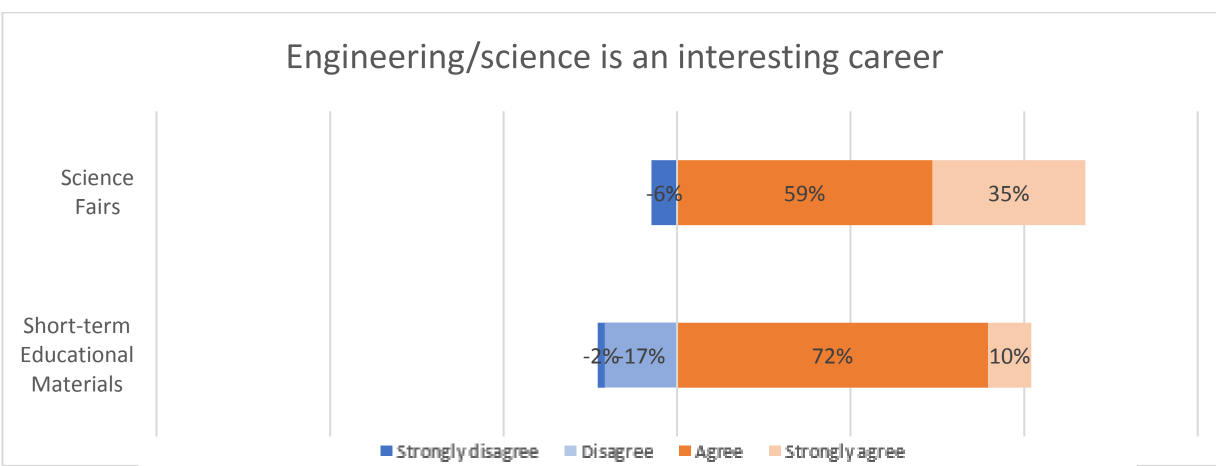
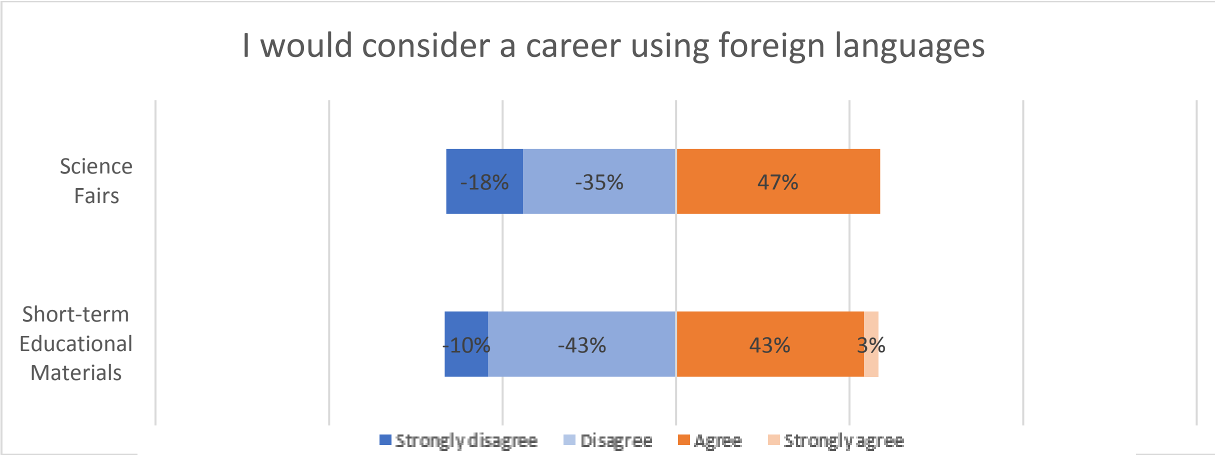


*Figure 12.g Agreement to the statement "I would consider a career using foreign languages"*




## 12.4 Conclusion

The data available to us and our current estimations suggest that the long-term use of Educational Materials is positively associated with STEM outcomes, but that there is no positive effect detected for short-term applications of EM. We could not find any positive and significant association of the short-term Education Materials intervention with any of the outcomes of interest. The long-term Educational Materials intervention was associated with a higher (+25 percentage-point) agreement with the following statement: "I have a better idea of what engineering is and how scientists and engineers make a difference to my life".

These findings suggest that the longer-term application of Educational Materials might be a more effective way for Siemens to increase pupil's interest in STEM-related subjects.

## 12.5 Additional tables

*Table 12.5 Long-term Likert table*

| Characteristics | Pre | | | | Post | | | |
|---|---|---|---|---|---|---|---|---|
| | Strongly disagree | Disagree | Agree | Strongly agree | Strongly disagree | Disagree | Agree | Strongly agree |
| Engineering/science is an interesting career | 6 | 6 | 25 | 8 | 6 | 6 | 32 | 2 |
| Foreign languages are an interesting career | 14 | 11 | 15 | 5 | 8 | 16 | 16 | 6 |
| What I have learnt in class helped me understand something I didn't understand before | 5 | 2 | 24 | 14 | 5 | 5 | 25 | 11 |
| I have a better idea of what engineering is and how scientists and engineers make a difference to my life | 6 | 10 | 18 | 11 | 6 | 19 | 16 | 5 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| I enjoy the class and would like to take part in similar science/technology/math activities | 6 | 14 | 18 | 7 | 5 | 16 | 16 | 9 |
| I have learned new things | 3 | 0 | 27 | 15 | 4 | 4 | 24 | 14 |
| I could use what I have learned | 3 | 8 | 22 | 12 | 4 | 6 | 26 | 10 |
| I get good marks in mathematics | 6 | 6 | 26 | 7 | 3 | 17 | 19 | 7 |
| Mathematics is one of my best subjects | 8 | 12 | 18 | 7 | 10 | 18 | 9 | 9 |
| I have always done well in mathematics | 8 | 11 | 14 | 12 | 8 | 17 | 12 | 9 |

*Table 12.6 Short-term Likert table*

| Characteristics | Science Fairs | | | | Educational Materials | | | |
|---|---|---|---|---|---|---|---|---|
| | Strongly disagree | Disagree | Agree | Strongly agree | Strongly disagree | Disagree | Agree | Strongly agree |
| Engineering/science is an interesting career | 1 | 10 | 43 | 6 | 1 | 0 | 10 | 6 |
| I would consider a career using foreign languages | 6 | 26 | 26 | 2 | 3 | 6 | 8 | 0 |
| What I have learnt today helped me understand something I didn't understand before | 2 | 17 | 34 | 7 | 1 | 2 | 9 | 5 |
| I have learned new things | 1 | 18 | 24 | 17 | 1 | 1 | 8 | 7 |
| I could use what I have learned today | 5 | 11 | 35 | 9 | 0 | 1 | 13 | 3 |
| I have a better idea of what engineering is and how scientists and engineers make a difference to my life | 3 | 24 | 26 | 7 | 1 | 0 | 12 | 4 |
| I enjoyed the lesson and would like to take part in similar activities | 0 | 15 | 36 | 9 | 1 | 1 | 7 | 8 |