



Gender Balance in Computing

Evaluation of Informal Learning: Apps for Good programme

August 2021

Martina Maglicic, Jemuwem Eno-Amooquaye,
Callum O'Mahony and Matthew Holt

Contents

Contents	1
Executive summary	1
Overview of the project	2
Evaluation approach	2
Key findings	3
Recommendations	4
Background	6
Informal Learning	6
Apps for Good programme	7
GBIC partners	7
Methods	8
Evaluating evidence of promise	8
Implementation and process evaluation	12
Key findings	16
Attitudes toward computing	17
Intention to study GCSE computer science	23
Intention to study STEM subjects at GCSE	27
Programme delivery	28
Conclusions and lessons learnt	31
Interpretation of findings	31
Lessons for the programme	32
Lessons for evaluation	33
Reference list	34
Appendices	35
Appendix 1: Logic model	35
Appendix 2: Survey measures	35
Appendix 3: Changes to SCSAS	40
Appendix 4: Model specification	42

Executive summary

Overview of the project

Although there have been increasing numbers of girls and women pursuing a career in computing, there remains a distinct gap between genders, which can be traced back to secondary school subject choices (bcs.org/more/about-us/press-office/press-releases). The Gender Balance in Computing (GBIC) programme has been structured around known and well-researched barriers to girls engaging with computing, including a lack of belonging, lack of relevance and difficulty linking informal computing education to formal GCSE computer science, all of which may hinder engagement and may negatively impact perception.

The GBIC programme has been funded by the Department for Education (DfE), with the Raspberry Pi Foundation (RPF) serving as the primary delivery organisation and the Behavioural Insights Team (BIT) acting as independent evaluators. This report details the evaluation of an intervention in the Informal Learning strand of the programme, where the aim was to encourage the uptake of formal computing education among girls by linking their informal computing learning to GCSEs. Specifically, this intervention was the Apps for Good (AfG) course, a teacher-led programme where children in Year 8 are taught practical coding and app development skills in an informal environment outside of school through a 12-hour course. Informal learning is defined as learning that takes place outside of the school curriculum and often outside of school hours, often in after-school clubs such as a Code Club. The aim of the course was to 1) improve the attitudes towards computing of girls taking the course, as well as their intent to take up 2) computing and 3) other STEM subjects at a GCSE level.

Evaluation approach

The intervention was evaluated using a mixed methods approach. A quantitative evaluation was conducted to investigate whether there was evidence of promise that the intervention achieved its stated outcomes. In parallel, a qualitative implementation and process evaluation was conducted to explain the quantitative findings and determine the processes of implementation and the mechanisms of change.

Evaluating evidence of promise

The quantitative evaluation was focused on whether there was evidence of promise that the intervention impacted:

1. Girls' attitudes to computing
2. Proportion of girls stating that they intend to study computer science at GCSE level

3. Proportion of girls stating that they intend to study any STEM subject (excluding Computer Science) at GCSE level

The primary outcome was measured using the Student Computer Science Attitude Survey (SCSAS), a validated survey tool for assessing attitudes toward computing for school pupils. The evaluation was originally designed as a randomised controlled trial (RCT), whereby schools would be randomly assigned into either the treatment or control group. The control group were to participate in the business-as-usual AfG course, and the treatment group would have undertaken a course which included specific activities linking the informal learning on the course to formal GCSE computer science education. Based on power calculations, a recruitment target of 150 schools was set. Unfortunately due to the COVID-19 pandemic, only 53 schools agreed to take part in the trial, which meant the original design would be under-powered. As such, we agreed with RPF and the DfE to move to a pre-post design to compare the three target measures before the intervention to those after the intervention. All schools received the enhanced course which included the additional activities and we used two surveys to measure the change in pupil outcomes. The first survey was filled out by pupils in the first lesson and the second survey in the last lesson.

Implementation and process evaluation

Alongside the evidence of promise evaluation, a qualitative implementation and process evaluation (IPE) was conducted to answer the following research questions:

1. What were the barriers and facilitators to successful implementation of the intervention?
2. What range of factors helped and hindered girls' engagement with the intervention?
3. What range of factors influenced girls' attitudes towards curricular computing education?
4. What range of factors influenced girls to participate in curricular computing education?

We interviewed an AfG staff member responsible for developing the classroom materials over the phone to understand the design and aims of the programme, and three teachers who delivered the intervention were interviewed over the phone or virtual platforms to better understand the feasibility of delivery, any programme adaptations and the perceived outcomes for girls. In addition to these interviews, BIT researchers spoke with pupils from two schools to better understand how they experienced the intervention and to capture the reasons behind any change, or a lack of change, in their attitudes and intentions to pursue computer science at a GCSE level.

Key findings

Overall, girls' attitudes toward computing improved from baseline to completing the AfG course, whereas boys' attitudes declined. Despite the improvements in attitudes for girls, there is no observed increase in girls' intent to study computing or other STEM subjects at GCSE level. Given the design of the evaluation, we cannot say whether these observed changes were due to the AfG programme or some combination of other factors. One of these

factors may have been related to fidelity, as teachers explained large differences in how they delivered the course and the adaptations that they made. In this way, it is difficult to consider the intervention as a stable concept to be evaluated.

Attitudes toward computing

- We observe a small, positive change in girls' attitudes towards computing between baseline and endline survey measures, while the change for boys was small and negative.
- Teachers reported that girls' attitudes to computing changed when they saw that computing was relevant to real world problems and when they interacted with role models in the computing industry as part of the AfG course
- Girls' reported that their attitudes to computing was helped by the AfG course's focus on creativity and on collaboration

Intent to study GCSE computer science

- We do not observe any change in girls' intent to study computer science at GCSE between the start and end of the course.
- Girls reported that the programme did not necessarily improve their confidence to study computer science at GCSE level
- Teachers felt that pupils needed more exposure to the technical aspects of computing in preparation for GCSE computer science and that the AfG programme may not have provided enough of this
- Teachers and families played an integral role in encouraging girls to study computer science at a GCSE level

Intent to study STEM subjects other than computing

- Girls' intent to study STEM subjects (other than computing) stayed the same throughout the intervention

Processes of implementation

- The course was rarely implemented as intended. Teachers cut and condensed content due to demands on teacher time and limited time in the timetable

Recommendations

1. Girls' attitudes toward computing improved due to a perceived relevance of computing to their everyday lives and a sense of belonging to the industry through role models. In the interviews, neither teachers nor girls mentioned the informal aspect of the course as a mechanism of change, suggesting that this focus may not be as prominent for girls as other factors. Therefore links between informal and formal learning should be more explicit for both pupils and teachers in future.
2. Boys' attitudes towards computing became more negative following the course, suggesting that the mechanisms working to improve the perception of computing for girls may be having the opposite effect for boys. This may be the focus on female role models, or the other mechanisms highlighted by girls including the focus on creativity and on collaborative working. The informal learning interventions will need to find a balance between encouraging improvements in girls without depriving boys of the positive experience of the course.

3. Despite improvements in attitudes, there remains a gap between attitudes and intentions for girls pursuing computing or STEM subjects at a GCSE level. While the focus on creative and collaborative elements were enjoyed by girls, teachers felt that there was not enough time spent on the technical aspects of the course which are better served to prepare girls for the demands of formal computing education. While engagement is an important step for girls to consider taking GCSE computer science, a balanced focus on engaging activities and technical skills was perceived by teachers as important in improving girls' confidence to pursue the subject.
4. Interviewed teachers reported that teachers and families were indicated to have played a role in encouraging girls to pursue formal computing education, however this was seemingly through either role modelling over the lifetime or a strategic and sustained plan from the teachers. The role of teachers and the family seems like a sensible focus of other interventions trying to link informal, or recreational, computing with formal learning through GCSEs.
5. Ultimately, an RCT would be useful to estimate the impact of the additional linking activities on girls' attitudes and future study plans. While the qualitative evaluation can shed light on what girls and teachers perceive to be the driving mechanism of changes in attitudes, an RCT would be able to determine whether the additional reflective activities are a worthwhile deviation from the standard AfG course. Furthermore, an RCT and more in-depth qualitative research, preferably not during a pandemic, would be helpful in better understanding whether and why boys may have perceived the course in a more negative light.

Background

Computing has a decades-old problem with gender imbalance with limited reliable evidence of what works in closing the gap (Royal Society, 2017). The Gender Balance in Computing (GBIC) programme has been structured around known and well-researched barriers to girls engaging with computing (Childs, 2021). The review of existing literature published by Childs in 2021 revealed a number of common barriers to girls selecting computing as a subject: a mismatch of teaching approaches to student learning styles; a lack of encouragement to studying computing; a lack of familial and other role models in computing; a perceived lack of relevance of computing to students; and a disconnect between extra-curricular computing activities and subject choice. The intervention being evaluated in this report focused on the latter, addressing the gap between informal and formal learning pathways.

Informal Learning

The link between existing informal learning, where pupils take part in computing activities outside of the school curriculum or school hours, and subject choice is not well-established. Furthermore, the influence of participation in after-school programmes is not well understood (Straw, Bamford & Styles, 2017). Nevertheless, after-school programmes already have a good gender balance and a set of female participants who have a social network of other girls who participate (<https://www.appsforgood.org/impact>; Leadbetter, Hazeldean & Quinlan, 2018). Therefore, understanding how to use participation in such programmes as a way of increasing girls' pursuit of formal computing education is key for improving the gender balance in computing.

These settings also typically offer opportunities for contextualised learning, a focus on creativity, and are gender neutral (<https://www.appsforgood.org/impact>; Leadbetter, Hazeldean & Quinlan, 2018). This intervention investigated whether an increased focus on subject choice in existing informal learning provision - through modules delivered during school time (but outside of class time) or outside of school time - increased interest or participation in curricular computing education. The materials draw on the behavioural science concept of endowed progress. By emphasising that girls have already made progress towards a goal (in this case, studying computing), they may be more likely to ultimately achieve that goal (Nunes & Dreze, 2006). Additionally, if a girl has a lack of self-confidence and doesn't see herself as someone who is "good at computing", then letting her know that she is already succeeding (and enjoying) elements of computing education should lead to an increase in computing confidence, which should in turn translate to an increase in the number of girls studying computing in formal settings.

The logic model (see Appendix 1) which frames the programme of interventions highlighted a number of barriers preventing girls from participating in formal computing education. A key barrier relevant to this trial is that *female pupils who participate in informal learning (in this case, the Apps for Good programme) may not recognise this learning as computing* and therefore participation in informal learning may not optimally translate into choosing to study computing in formal education (Means et al., 2017).¹ The intervention evaluated in the current trial aimed to address this barrier.

Apps for Good programme

The current trial focused on informal learning, which includes after-school programmes such as the Apps for Good (AfG) courses. AfG courses are teacher-led programmes where children in Year 8 are taught practical coding and app development skills in an informal environment outside of school through a 12-hour volunteer-led course. AfG courses currently have a wide reach in primary and secondary schools.

Intervention materials were developed by AfG with close discussion with RPF and were made available online to teachers delivering the AfG course. Teachers accessed the content online and delivered it face-to-face to pupils participating in the extracurricular AfG course, although due to COVID-19, some course delivery moved online. AfG course teachers were a mix of IT/computing teachers already working in the schools or volunteer teachers recruited specifically for the AfG course.

This project evaluated a variation of the programme, which focused on linking informal and formal learning. This linking is woven into the course as a starter activity, mid-way activity, and plenary (with each activity taking 20 minutes). In total, the extra material takes about one hour to complete. RPF instructed AfG teachers to follow material in the designated order, however understood that in occasional instances, AfG teachers would present pupils with choices in terms of material, or move through material in a different order.

GBIC partners

This project joins the National Centre for Computing Education (run by a consortium comprised of STEM Learning, the British Computer Society [BCS], and the Raspberry Pi Foundation [RPF]) with the Behavioural Insights Team (BIT), combining the extensive experience of organisations who have computing at the core of their mission with expertise in designing and evaluating interventions. The funding body for this programme as a whole is the Department for Education (DfE), and BIT fulfils the role of an independent and external evaluator.

¹ For example, inclusive STEM high schools in the US aim to address this barrier by admitting underrepresented groups on the basis of interest rather than competitive examination, and by including curricular and extracurricular STEM experience; underrepresented pupils attending such schools were more likely to pursue advanced STEM coursework and express an interest in STEM careers (Means et al. 2017).

Methods

The AfG course was evaluated using a mixed methods approach. A pre-post survey of three outcome measures was conducted and this was complemented by a qualitative implementation and process evaluation, which aimed to explore the mechanisms of change and to complement the quantitative survey findings. This section addresses the methods used as well as the limitations of our approach.

Evaluating evidence of promise

An evaluation was conducted to determine if there was evidence of promise that the intervention was affecting the target outcomes.

Research questions:

The impact evaluation was focused on whether we observe change in:

1. Attitudes to computing
2. The proportion of girls stating that they intend to study computer science at GCSE level
3. The proportion of girls stating that they intend to study any STEM subject (excluding computer science) at GCSE level

Research design:

To evaluate the effectiveness of the intervention, we used a pre-post design to compare the three target measures before the intervention to those after the intervention. Due to us not being able to meet our recruitment targets, all schools were assigned to the treatment group.

We used one survey to measure the effect of the intervention on pupil outcomes, before the intervention and afterwards (see Appendix 2). The survey was filled out by pupils in the first lesson and the same survey was completed by pupils in the last lesson. The aim was to determine if pupils felt differently about computing after the lessons they'd attended. Both surveys were filled out by pupils on a computer or phone in class with a teacher nearby to help answer any questions. Each survey took about 10 minutes to complete. We asked about the pupil's personal information (e.g. name) to facilitate matching baseline and endline surveys, if they wanted to study different STEM subjects at GCSE level and to agree or disagree with a range of statements about computing. Questions did not differ between baseline and endline surveys. All pupils were asked to fill out the survey, but we excluded data from boys in our main analysis. Instead, the exploratory analysis looks at boys' outcomes.

An initial run of this trial was launched in December 2019 and disrupted by COVID-19 during March 2020. Baseline and outcome survey data were collected for a subset of pupils (this is summarised for female pupils in Table 3 below).

To minimise the effect of attrition on our already reduced sample, we kept data from the first trial and included them in the analysis for the primary (attitudes towards computing) and one of the secondary outcomes (stated intention to study computer science at GCSE level). This was also designed to help mitigate external time effects, since the sample was based at two different time points. This included everyone who was in the control group in the pilot study too, as the only difference between the control and the treatment group was additional material (around 1 hour split across three lessons), which had minimal effect on our estimates. We were not able to include the pilot data in the analysis of the other secondary outcome (stated intention to study any STEM subject except computer science at GCSE level) as this question included an almost entirely different set of answer choices in the pilot.

Outcome Measures:

The primary research question is whether the intervention affects attitudes to computing. This is measured using the outcome measures outlined below, which were collected at both baseline and endline:

Outcome Measures	Data to be collected
Primary: General attitudes towards computing	Overall score on the Student Computer Science Attitudes Survey (SCSAS; score range: 0-100) ²
Secondary: Stated intention to continue studying computer science at GCSE level	Single item survey of whether the pupil intends to continue studying computer science at GCSE level, with possible responses including “Yes”, “No”, or “ <i>I don’t know</i> ”. The intent of the evaluation is to separate “positive” and “non-positive” responses, so this was binarised into: <ul style="list-style-type: none"> • 1 = “Yes” • 0 = “No”, “<i>I don’t know</i>”
Secondary: Stated intention to continue studying any STEM subject (excluding Computer Science) at GCSE level	Binary indicator in which <ul style="list-style-type: none"> • 1 = intends to study any STEM subject at GCSE, excluding Computer Science (only a “yes” response counts as “intending to study” — a “don’t know” response is not considered positive) • 0 = does not intend to study any STEM subject at GCSE level (excluding Computer Science)

The Student Computer Science Attitude Survey (SCSAS) has been developed to measure attitudes towards computing (Haynie and Packman, 2017) (see Appendix 2 for the full survey content and Appendix 3 for the summary of adaptations. It contains 25 questions and has 5

² Adapted from Haynie and Packman (2017). Available at: <https://csedresearch.org/tool/?id=156>. See Annex B for details of adaptation.

subcategories (5 questions per subcategory): confidence, interest, belonging, usefulness and encouragement. Survey item answers are taken on a five-point Likert scale ranging from strongly disagree to strongly agree. Within each subcategory, the 5 items are scored from 0 (strongly disagree) to 4 (strongly agree), and summed to create subscores. Thus, each 5-item subscore has a potential range of 0-20, and the total score has a potential range of 0-100, with 100 representing a very positive attitude towards computing. For the secondary outcome measures, pupils self-reported their intention to study a range of STEM subjects.

Only girls were included in the main analysis, but we repeated the analysis with boys and non-binary pupils in an exploratory fashion (see Appendix 4 for details of analytical approach). As outlined above, we included data from the initial pilot in the analysis, except for the secondary outcome 'stated intention to study STEM subject' as this question had been altered to such an extent that answers were not comparable anymore.

Sample:

Recruitment of schools was conducted by RPF (as opposed to the research team), as it was important that the relationship with the schools was developed by RPF to support implementation and it also reflected real-world implementation in which research was not taking place. This arrangement also supported maintaining independence in the research as the research team would have had no contact with the schools prior to randomisation if randomisation had happened. All secondary schools in the UK were eligible for this trial, with the exception of schools that participated in another GBIC trial in the same academic year (specifically, the Teaching Approach i1a RCT or the Belonging i3 RCT). This was to avoid contamination between trials. RPF was responsible for anticipating and managing such 'conflicts' between trials during recruitment.

To enter the sample, the secondary school must have volunteered to introduce the AfG course. This may undermine the external validity of the findings, because those schools that volunteer are likely to be the most enthusiastic, and this may interact with the treatment effect to compound any effects. These schools may also already have higher levels of engagement and so make it harder to increase the outcome.

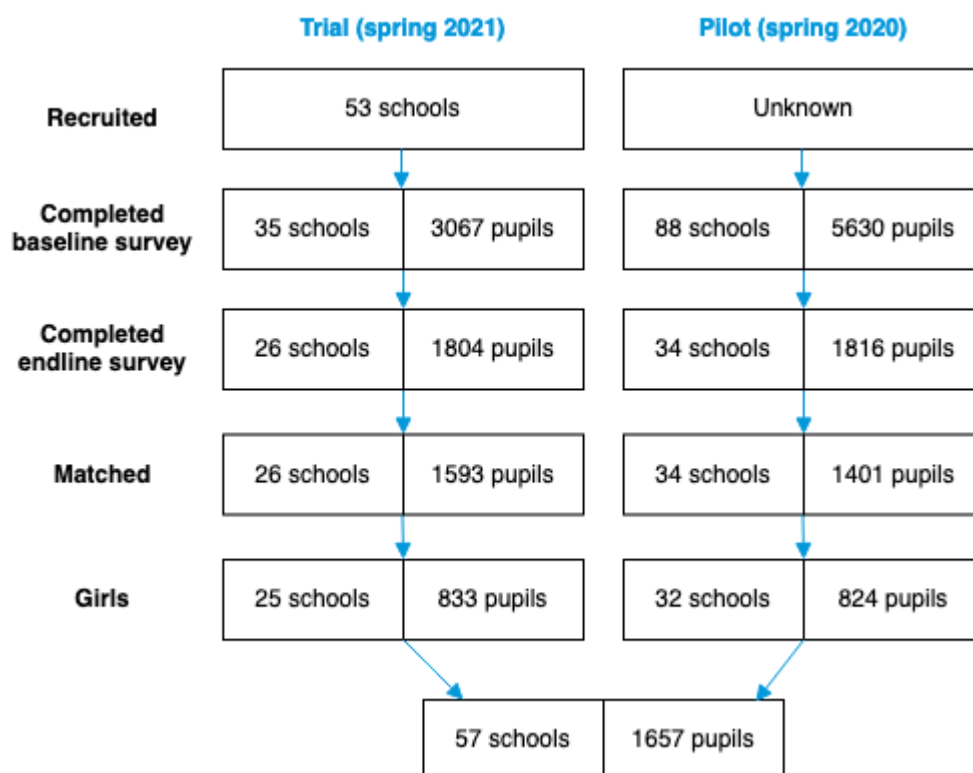
In order to preserve the fidelity of the intervention, only the schools that could offer the full 12-week programme to a single cohort of pupils were able to enter the trial. Schools must also have female pupils (all-boys schools were excluded).

All pupils in this trial were in Year 8. As per DfE preferences, the trial included both boys and girls, but only data from girls will be analysed for primary and secondary analyses. Exploratory analysis of data from boys was conducted to estimate any incidental effects of the intervention among this group.

Figure 1 below breaks down the collected sample at different stages of the trial and shows at which stages participants and schools dropped out, as well as the final sample size. The high rate of attrition between baseline and endline survey in both instances is likely due to

disruption of the coronavirus pandemic. In spring 2020, the pandemic had just started and data collection was impacted by a lockdown, while in spring 2021 there was a third wave of infections and many pupils were at home self-isolating. Furthermore, schools had to produce Centre Assessed Grades for GCSEs and A Levels during this time, and this was cited by schools to RPF as one of the reasons for attrition. The final analytical sample consists of 1657 girls. The sample of boys and non-binary pupils used in exploratory analyses contains 1337 pupils.

Figure 1. Sample breakdown.



Implementation and process evaluation

Alongside the evaluation of evidence of promise, a qualitative implementation and process evaluation (IPE) was conducted. The IPE was concerned with the mechanisms of success and the diversity of implementation and programme delivery.

Research questions:

1. What were the barriers and facilitators to successful implementation of the intervention?
2. What range of factors helped and hindered girls' engagement with the intervention?
3. What range of factors influenced girls' attitudes towards curricular computing education?
4. What range of factors influenced girls to participate in curricular computing education?

Research design:

We note up front that we were not able to conduct the research exactly as planned, due to challenges posed by Covid safety measures. Throughout this section, we explain the planned activities and how they were modified due to Covid restrictions.

Qualitative methods were appropriate as they allowed us the flexibility to explore the context of each school and the teachers and pupils within them who participated in the programme. We planned and implemented a case study design, conducting qualitative research with teachers and pupils from the same school. We envisaged conducting research activities in four case study schools, however were only able to successfully recruit three.

Sampling, recruitment and methods:

Across the participating schools, case study schools were selected to represent range and diversity, both in terms of the school makeup as well as the teacher and pupil experience of the intervention.

Case study school criteria

Primary sampling criteria for schools included i) region and ii) proportion of free school meals (FSM). School's average number of pupils eligible for FSM were retrieved from government official national information about schools. We were able to achieve some diversity in the proportion of FSM, however only schools in the North consented to interviews and site visits.

Staff criteria

For teachers within the case study schools, primary criteria included gender, and secondary criterion included teaching experience, however this was removed from the sampling criteria due to recruitment issues (explained below). We also interviewed the one AfG staff member involved in co-developing the additional resource tool.

Pupil criteria

We similarly aimed to recruit a sample of eight year 8 pupils (across the same four case study schools; two pupils per school) for individual interviews. We aimed to interview a small number of boys (approximately two across the four schools) to deepen our understanding of any potential backfire effects of the intervention, as well as similarities and differences in their experience of the intervention compared to girls, however were only able to interview boys from a single school. This is because one school could not give us access to pupils due to most of the class self-isolating at the time of the visit, and the other school being an all girls school.

For pupils, primary criteria included gender and the pupil's confidence with computing, as advised by the teacher prior to inviting pupils to interview. The sampling of pupils occurred following the teacher interview, or at least following conversations with the teacher who was able to identify students with different levels of confidence within their class.

The impact of COVID-19 on class absences meant that we were not in a position to be too rigid with our sampling. Contacted teachers explained that self-isolating students meant that classrooms were rarely at capacity, and in some extremes, completely empty. Instead, we relaxed our criteria in order to balance the need for range and diversity with the realities of

conducting qualitative research in schools during a global pandemic. Table X below details the achieved sample of case study schools.

Table 1. Achieved case study sample.

School	Profile	Teacher	Pupils	Data collection
S01	-Located in the North -Below average FSM -State school	-Single classroom teacher -Male	-Unable to interview pupils due to COVID absences	-Interview recorded and transcribed
BS03	-Located in the North -Independent school -FSM data unavailable -Single sex school	-Single classroom teacher -Female	-Female pupil with high computing confidence -Female pupil with low computing confidence	-Individual interviews -Conducted virtually with note taker
S03	-Located in the North -State school -Higher than average FSM	-Single classroom teacher -Male	-Female with high confidence -Male with inconsistent confidence -Male with low confidence	-Group interview -Conducted virtually and transcribed

School recruitment

We initially reached out to selected schools via the email contacts provided by RPF, however received few responses. After several attempts, we asked RPF to reach out to schools on our behalf, leveraging their relationship with school networks to legitimise the request for participation in this evaluation. Following agreement to participate, BIT staff contacted the classroom teachers delivering the intervention, discussed the practicalities of a school visit or whether research should be conducted online, and scheduled a date for participation.

Staff interviews

Individual in-depth semi-structured interviews were planned to be conducted with a teacher of treatment classes in each case study school to explore their experience of the intervention and any factors that influenced their ability to implement the intervention with their pupils during the term. We were able to conduct individual 45-minute phone interviews with a teacher in all 3 case study schools.

An individual in-depth semi-structured interview was also conducted, as planned, with the AfG staff member involved in the provision of the additional online resource, to explore their experience of co-developing the additional resource tool and working with schools that are implementing the additional resource in practice. The interview took place over the phone in June 2021 and lasted about 45-minutes.

Pupil interviews

For each case study school, we had initially planned to visit schools and observe a lesson of the AfG course, before conducting interviews with pupils who attended that lesson. When this was deemed infeasible due to Covid restrictions, we then planned to conduct individual semi-structured interviews via a video platform with two pupils in treatment classes per case study school, for a total of eight pupil interviews. Ultimately, we were able to interview five pupils through a combination of individual interviews and a small group interview, the latter of which having been considered more logistically appropriate by the teacher. The interviews were conducted toward the end of the programme implementation in July 2021. In one case study school, we followed the intended plan and conducted 30 minute interviews with two female pupils. For another case study school, after speaking to the teacher it was deemed more appropriate to conduct a small group interview, as the teacher felt the pupils would be more comfortable speaking. This group consisted of three pupils and lasted around 45 minutes. For the final case study school, we were not able to interview pupils due to school closures and pupils self-isolating following COVID-19 outbreaks.

Analysis:

Interview transcripts and field notes were managed using the Framework Approach³. This involved summarising transcripts and notes into a matrix organised by themes and sub-themes (columns) as well as by individual cases (rows). The managed data was then interpreted with the aim of identifying and categorising the range of phenomena present in each of the sampling groups. We conducted case and theme analysis to focus on providing rich **descriptions** of participant experience whilst looking for **explanations, linkages and typologies** within and across participant groups.

Limitations:

The IPE had several limitations for consideration when interpreting the findings:

1. The case study approach means that findings should not be generalised across all participants, but rather understood as conveying some of the range and diversity of participant experiences.
2. Due to COVID-19, we were unable to fulfil our intended sample of schools, instead adopting a convenience sampling approach, focusing on schools that were able to take part in the research activities during the pandemic.
3. We were also unable to interview as many pupils as intended, however by working with teachers we were still able to speak to pupils with different genders and different levels of confidence with computing.
4. We were unable to conduct field visits and observe the lessons in action due to COVID-19 restrictions and class absences. Instead, fieldwork was conducted online and relied on teacher and pupil accounts to convey implementation, mechanisms and outcomes for pupils.

³ Ritchie, J., Lewis, J., Nicholls, C. M., & Ormston, R. (Eds.). (2013). *Qualitative research practice: A guide for social science students and researchers*. sage.

Key findings

This evaluation was conducted on a pre-post basis. As such, we cannot be certain that any differences were caused directly by the programme; they may have been caused by other events that occurred at the same or similar time, or simply due to the passage of time or the fact that at the point of endline measurements, the students were a little older and more experienced than at the baseline.

With that limitation, we found:

- **A statistically significant but small improvement in girls' engagement**, as measured by the SCSAS score. This increased by 0.7 points on the 100-point scale post-intervention. This difference is significant at the 5% level according to a multilevel linear regression analysis, controlling for school year and with random intercepts and random treatment slopes at the school level, using a t-test and the Satterthwaite approximation for the degrees of freedom. Full details of this model and test can be found in appendix 4.
- **This increase was focused in the “Confidence” and “Encouragement” subscales of the SCSAS**, showing increases of 0.3 points out of 20 and 0.5 points out of 20 respectively (both significant at the 1% level using the same test as above). The “Interest” subscale decreased slightly, by 0.2 points out of 20 (significant at the 5% level using the same test as above). There was no significant change in the other subscales.
- **No significant change in the percentage of girls who indicate that they intend to study computer science at GCSE**. This was assessed using the same model and test as above, except that the regression was logistic rather than linear as this outcome is binary.
- Similarly, we observed **no significant change in the percentage of girls who indicated that they intend to study non-computer-science STEM subjects at GCSE** using the same model and test as previous.

Overall, the findings demonstrated an improvement in girls' attitudes towards computing but not in their intent to pursue computer science or other STEM subjects at a GCSE level. The findings also suggest that boys reported a more negative attitude toward computing following the AfG course, meaning that the mechanisms that girls find engaging and rewarding may have the opposite effect on boys. However, we caution that due to the design of the evaluation, we cannot be certain that these changes in pupil attitudes and intentions can be solely attributed to the AfG course.

This section provides both quantitative and qualitative findings related to each of the three primary outcomes: 1) Attitudes toward computing, 2) Intention to study computer science at a GCSE level and 3) Intention to study STEM subjects at a GCSE level. We have also included

some findings relating to programme implementation from the perspectives of the teachers to provide context and suggestions for improvement for future course iterations.

1. Attitudes toward computing

Key findings:

As measured by the survey,

- We observed a small but positive change in girls' attitudes, while boys' attitudes became more negative.
- Overall, girls' attitudes towards computing were neutral.

According to interviews with teachers and pupils,

- Girls' attitudes to computing changed when they saw that computing was relevant to real world problems and when they interacted with role models in the computing industry.
- Girls' attitudes to computing was helped by the AfG course's focus on creativity and on collaboration

Girls' improvements in attitudes to computing

Pupil's attitudes towards computing were measured using the Student Computer Science Attitude Survey (SCSAS) The minimum composite score was 0 (indicating a very negative attitude) and the maximum was 100 (meaning a very positive attitude).

We found that girls' scores on the SCSAS increased by 0.7 points from a baseline average of 44.4 points, indicating a slightly more positive attitude towards computing (see Figure 2). However, scores from boys and non-binary pupils showed a weak but significant decrease of 1.2 points from a baseline of 52.7 points. Although we cannot draw causal conclusions from this finding, it is possible that the intervention may have contributed to improving girls' attitudes but also worsening boys' attitudes in the process. However, these changes are quite small; they are equivalent to each girl increasing her answer to one question by one point on average. Findings from the implementation and process evaluation support this possibility and we present these findings in more detail below.

Our exploratory analysis, which looked at each of the five subscales, further revealed that this increase was likely driven by girls feeling more confident and more encouraged to study computing. Interestingly, we also found that interest in computing slightly dipped at the end of the intervention. Feeling of belonging and usefulness remained the same as at the start of the lessons. For the full breakdown, see Figure 3.

Figure 2. Girls' scores on the Student Computer Science Attitude Survey, before and after the intervention. The bar on the left shows the average SCSAS score before the intervention began, and the bar on the right after it had ended. Orange lines indicate confidence intervals.

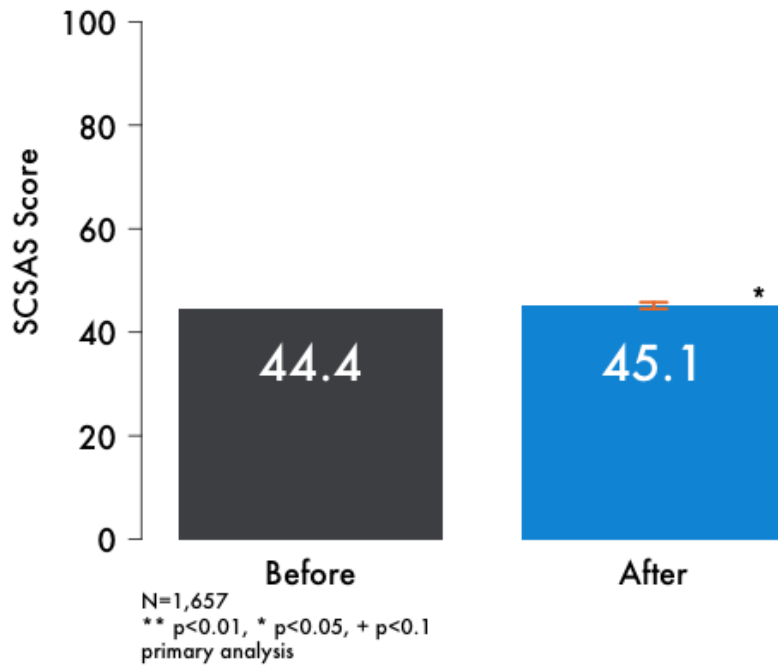
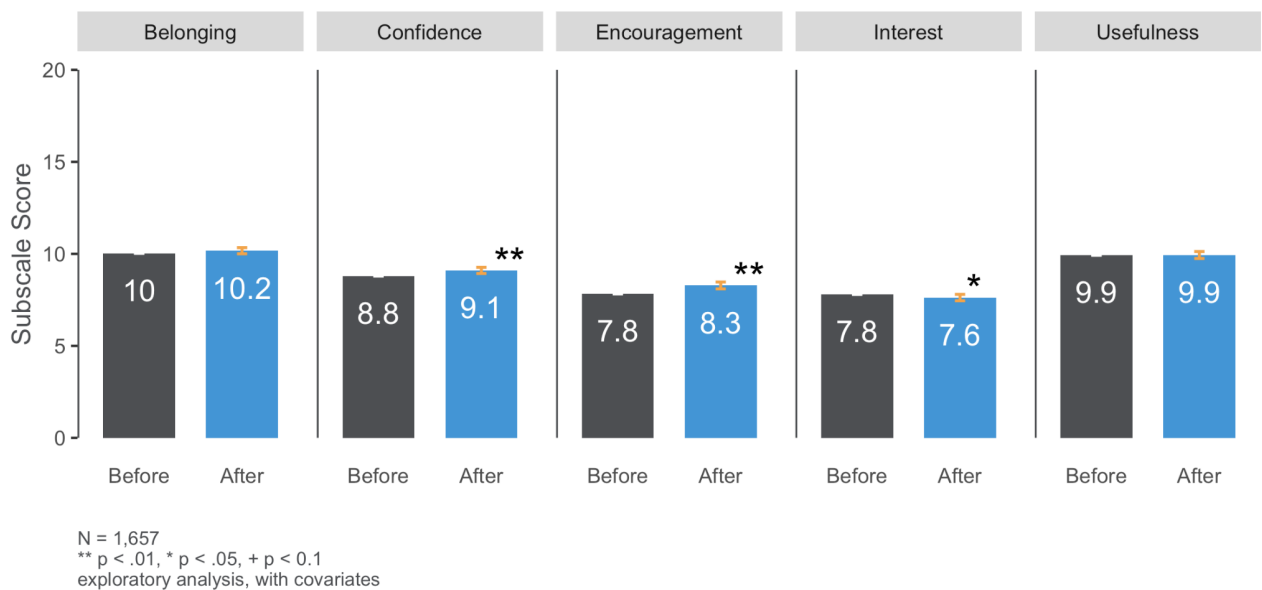


Figure 3. Girls' scores on the Student Computer Science Attitude Survey subscales, before and after the intervention. Panels specify the subscale and orange lines indicate confidence intervals.



Factors influencing attitudes toward computing

Findings from the qualitative IPE highlighted several key factors that either promoted or hindered positive attitude changes for girls thinking about computing. Within the course itself, teachers and female pupils discussed the importance of relevance, role modelling, collaboration and creativity in engaging girls in the AfG course and in helping them expand

their view of computing. However, the activities designed to link informal learning to formal education were not considered by teachers or pupils to have contributed to the changes in attitudes.

1. Relevance to real world problems

The teachers stated that the main appeal of the course was the relevance to pupils' daily lives and the problems they face in the real world. This improved attitudes toward computing by shifting the perception of computing away from gaming and abstract coding to the use of computers in solving real world problems. Teachers stated that girls had a perception of computing as detached from the real world in the realm of gaming and hacking, and that girls typically use computing less than boys in their daily lives, in particular outside of formal education settings.

"It's more of an issue about, the girls are not playing around with computers and stuff so much outside of the school environment. I mean, yes, sure, they'll play Candy Crush Saga and stuff like that, but they're not, in their own time so much, picking up a Raspberry Pi and programming it for fun, and that sort of thing" (BS03).

In contrast, teachers felt that the relevance to real world problems in the community made computing more purpose-driven and appealing to girls. One teacher felt that it was important for girls not only to know how to use computers, but also to develop an understanding of how computers work given the increasingly technologically driven nature of workplaces these days. The AfG staff member highlighted the importance of providing context to computing if girls are to see computing as relevant, stating that female pupils will not be motivated to build an app for the sake of it or to develop a program for the sake of it. Instead, the AfG staff member suggested that girls' attitudes toward computing improved when they learned how computers could be used to help people.

"We've been really keen in Apps for Good to try and overcome, to have it at more - very much focused on the problem solving, that you are helping people, you are solving problems. You are not just being a geek and playing with the hardware side of it" (AfG).

This notion was supported by interviewed teachers, one of whom stated that the girls in their class who could see the real world application of computing in their community were more likely to see themselves as computer scientists.

“You do something like Apps For Good and they go, ‘Yes, I can build an app that can rescue pets in my region and that’s computing,’ and suddenly, it’s relevant. I think that’s helped” (S03).

This teacher stated that they had observed an improvement in girls’ attitudes to computing due to perceived relevance, stating that they had spoken with girls who went home and discussed their app with their family and in some cases, members of their community. The teacher described the community relevance of computing as *“a bit of a breakthrough for some of them” (S03).*

Teachers also suggested ways that the AfG programme could be perceived as even more relevant by girls, in particular the idea of using completed apps early in the programme to demonstrate what is possible to develop over the course.

“Get some finished App Lab apps that are pretty useful and demo them and get them to play with them, because I think not seeing what’s possible held them back at the beginning. If I’m doing it again, I think we need to see much more what’s possible” (S03).

This teacher stated that in future they would begin by showing the pupils working apps and let them play with successful apps from the App Lab studio, believing this would help debunk the myth that computing is to do with space exploration and flying cars, and that they can use it to help their community.

2. Role modelling from industry speakers and staff

Role modelling from industry experts improved the legitimacy of the course for female pupils. Pupils spent time with a range of experts in a range of industries involving computing, including developers from Kahoot! and Siemens. Schools were able to request an expert by completing a short form, whereby the team at AfG would then match an expert to the session and they would join the classroom via Google Hangouts, Zoom, Microsoft Teams etc. The students would then pitch their ideas and the expert would give them some feedback and ideas on how to improve. Teachers viewed these interactions as helping pupils feel important and to view the course as important. Teachers also recognised the importance of the attendance of female experts, with one teacher stating that their female students were able to see themselves being represented.

“There were lots of different experts, but mostly, they were female, which in itself is great. The girls saw themselves represented on the screen by these experts” (S03).

While representation was viewed as helpful in motivating pupils to engage, the actual profession of the expert was a critical factor in influencing pupil perceptions. One teacher compared the two experts who worked with two different groups, noticing the difference in reception for the expert who worked directly in computing and app development and the one who worked as a teacher.

“The first group who had the expert who was not a teacher were really wowed by it and it was fantastic. The second group who had the teacher were not quite so wowed by this fact, so that's just something to bear in mind” (BS03).

The teacher observed pupils disengaging from the activities involving the teacher expert, and stated that this consideration of what constitutes an expert and what may be deemed an industry role model should be considered moving forward. Where pupils were engaged and motivated by the expert, teachers recommended moving onto the creative and technical aspects of the programme, in order to capitalise on the momentum gained by expert interactions. In the AfG delivery model, teachers are meant to deliver two lessons on research following the industry expert session. One teacher felt that following the industry expert with sessions on research contributed to attention loss and a loss in enthusiasm to use the tool and develop a product.

“It would have been nice if after that we could've just got into the tool and started creating a product, but no, we had to do two more lessons of researching this or researching that, and I felt like their attention was lost when we had such good attention after the industry expert. If we could've gone and produced the product then, there may be more enthusiasm and I think that was lost” (BS03).

3. Creativity

Interviewed pupils found the creativity of app development to be an engaging aspect of the AfG course. Pupils learned about the process of developing an app, including the process of ideation and of refinement, and then of presenting the finished product. Pupils stated that they enjoyed the freedom to do what they thought was best, however one pupil did state that they enjoyed the creative side as it was easier than doing more technical work.

“It's easier than doing written work. In written work there's a lot of techniques and a lot of things you need to include. If you're doing creative work you can do what you think is best” (Pupil).

In this way, while pupils may have been more engaged with the course, they may not have ended up more prepared for the technical requirements of GCSE computer science. This is discussed in greater detail in the following section addressing pupil intent to study computing.

4. Collaboration

The collaborative nature of the AfG programme was a key factor in improving girls' attitudes to computing, allowing girls to work with more confident classmates and providing them a role in the programme process. One teacher discussed an awareness of research that girls learn by being able to work on things together and to help one another, making group work the ideal learning environment, and this was supported by positive pupil accounts of teamwork while undertaking the AfG programme:

“Probably that it was fun to do a project in a team in computing specifically - usually I’m not good at computing but because you’re working in a group, other people could help and I felt better about it”. (Female pupil)

As well as the opportunity to learn from one another and to teach one another, teachers also observed the importance of collaborative work in providing a variety of roles for pupils to take on during the group process. This variety meant that girls with different skills, different preferences and different attitudes were able to find something they enjoyed doing and felt confident doing within a computing context.

“The opportunity to take different roles and have somebody who's more of a creative, you know, creates the visuals. Somebody who stands up and talks, not everybody wants to do that. Somebody who sticks a bit of code underlying it, and I think that's quite helpful to be able to make use of the different skills in the room” (BS03).

Despite the opportunity to learn provided by collaboration, teachers also commented on the difficulty in managing group dynamics and ensuring pupils are challenged in a collaborative learning environment. One teacher reflected that they needed to be stricter about the roles being distributed, meaning that pupils took on the roles with which they were most comfortable, without challenging themselves to either improve their technical ability or their presentation skills. The teacher stated that the amount of content in the course was problematic for managing group dynamics, as they were preoccupied with getting through the lesson and did not feel they could spend the time mixing up roles over the course of the programme.

COVID-19 also impacted pupils' ability to collaborate. Teachers described the effects of rising case numbers and girls' absences, meaning that they were not able to access the software to create the app, and pupils found group work difficult with some in school and some at home. To remedy this, one teacher organised meetings on Google Meets and used breakout rooms to try and recreate classroom collaboration online, however commented that using the app development software in school was still the most effective use of time.

2. Intention to study GCSE computer science

Key findings:

As measured by the survey:

- We do not observe any change in girls' intent to study computer science at GCSE

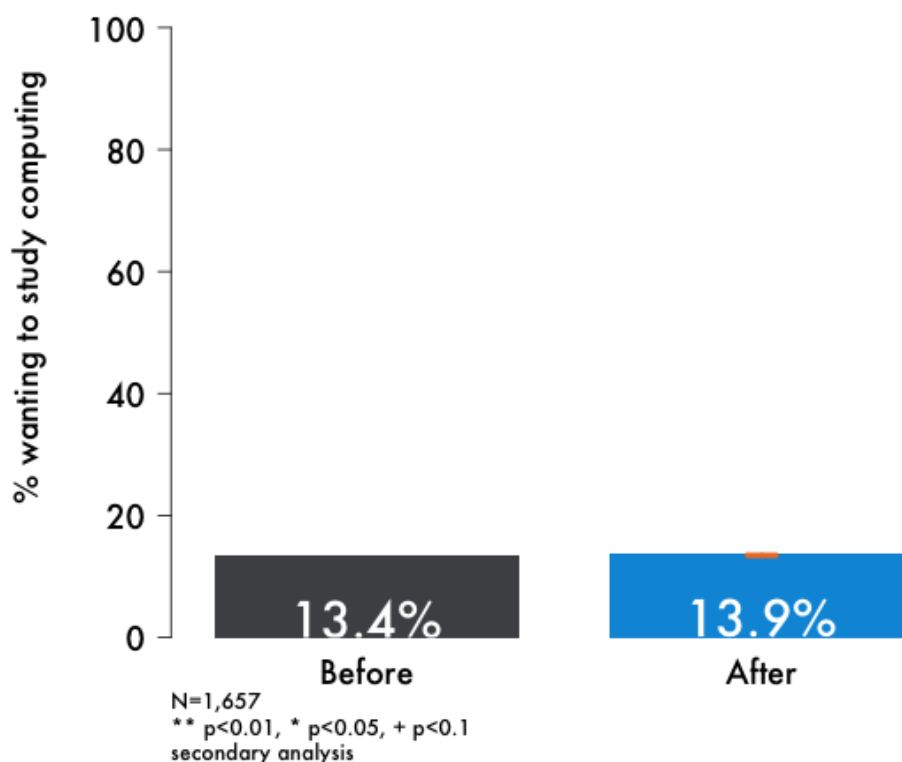
According to interviews with teachers, AfG staff and pupils

- Pupils stated that the programme did not necessarily improve girls' confidence to study computer science at GCSE level
- Teachers felt that the programme needed to focus more on the technical aspects of computing to influence intent to pursue GCSE computer science
- Teachers and families may play an integral role in encouraging girls to study computing

Girls' intention to study computing

We find that the proportion of girls intending to study computer science at GCSE level did not significantly change throughout the intervention (see Figure 4). Although we see a slight increase of 0.5 percentage points from a baseline rate of 13.4%, this effect is not significant. For boys and non-binary pupils we find that the proportion intending to study computing decreased by 2.7 percentage points compared to the start of the intervention but this is also not significant. However, baselines for boys and non-binary students were much higher, with 39.1% intending to study computing.

Figure 4. **Proportion of girls that intend to study computer science at GCSE level, before and after the intervention.** The bar on the left shows the average likelihood that a girl would intend to study computing before the intervention began, and the bar on the right after it had ended. Orange lines indicate confidence intervals.



Factors influencing intention to study computer science

Both participation in the course and influences outside of the informal learning programme seemingly influenced girls' intention to study computer science at a GCSE level. Teachers and girls viewed confidence as a significant barrier or enabler to studying computing. Teachers believed that the course, while enjoyable, did not prepare girls for the technical challenges of GCSE computer science. The other two main influences on girls' intent to study were direct support from teachers over a prolonged period of time, as well as ongoing support from their families.

1. Female confidence in computing

Both teachers and female pupils highlighted a lack of confidence as the main reason they may not intend to study computer science at GCSE level. Interviewed pupils with lower confidence did not see themselves as computer scientists, instead viewing people who do computing as different from themselves.

“To be honest, they seem really clever, I just don't find computing is something that comes naturally to me” (Female pupil).

In contrast, a female pupil with high confidence stated that they enjoy computing, as they get to understand how things work and all of the different aspects of what it takes to create an app. Teachers who have regular conversations with their pupils stated that girls often enjoy

computing at Key Stage 3, but feel that it is too difficult to pursue at GCSE level due to the importance of doing well for their future ambitions. One teacher stated that the girls with whom they have spoken still see computing as inaccessible, or express concern that they would excel more in a different subject. In this way, enjoyment of the course may not translate into pursuing computing beyond the AfG programme.

Technical aspects of the programme

Teachers felt that a focus on the technical aspects of computing would help improve girls' intention to study GCSE Computer Science. Teachers also commented on the lack of this focus on the AfG course, even though this was not deemed the aim of the programme, and seemingly demonstrated a mismatch between teacher expectations and programme objectives. The AfG staff member highlighted the importance of providing context around the IT aspects of the course, however teachers felt that the balance was too far on the non-IT side. One teacher explained that their pupils were keen to get moving on the app building, and to spend less time on the refinement process, meaning that some of these activities were cut short. The teacher acknowledged that in the real world, the refinement process of an app is undoubtedly important, however felt that the focus on planning and refining left too little time for the actual technical components linked to pupil pride in their computing work.

“It did feel like there's a lot of the idea generation screening and filtering in the scheme of work and at the expense of actually building something that the kids can be proud of” (S03).

Another teacher similarly recalled condensing the reflective and refining elements of the course in order to focus on the practical app building. This teacher said that their girls were not interested in reflecting on soft skills (one of the additional activities to the AfG course), but wanted to build the app, which the teacher acknowledged was more of their own interest as well. The AfG staff member had stated that the context around an app is extremely important, and it is possible that reducing the reflective time around broader computing skills and identity may influence whether girls perceive their informal learning as being linked to GCSE computer science studies. This constitutes an issue of fidelity, as girls may not have experienced the programme as intended and therefore the measured outcomes may be unreliable.

2. Direct intervention from the teacher

Girls were also subject to influences outside of the AfG course, in particular the important role played by teachers in turning an interest in computing into subject selection for girls in Year 8. Teachers spoke about their awareness of the gender gap in their GCSE classes, as well as the interventions staged by teachers while girls are in Year 8. Specifically, interviewed teachers outlined their role in reassuring girls with computing ability that they are good enough to take computer science at the GCSE level. The case example below conveys the experience of one teacher working to improve the gender balance in his own school.

In 2018, computing teacher James' department discussed the need to encourage more girls to take computer science at a GCSE level. James decided to put extra effort into having individual conversations with the girls in Year 8, in particular girls who demonstrated ability for computing but who lacked confidence to pursue computing from Year 9 onwards.

"I think we both - all of us, teachers in the department put a lot of effort in that particular year and I think it paid off."

James reflected that the girls who had direct teacher intervention had the highest conversion rate into GCSE students that he has ever taught. James went on to highlight the importance of one-on-one teacher discussions, as this aspect has been compromised since the pandemic began. James believes that the absence of these discussions has contributed to a lower number of girls moving into GCSEs for the next school year. James highlighted the need for teachers to consistently reassure female students with an interest in computing, for them to overcome their belief that they can not succeed at computing due to their gender.

"If you don't have the chance to do that, if you don't repeatedly tell them from Year 7 to Year 9 that, 'You're good at this, you should do this.' Then I don't think girls generally are willing to believe that they're good at computing".

3. Family support at home

According to teachers, families can play a critical role in converting an interest in computing to studying computing at school. Teachers reflected that girls they have taught who pursue computing have often had parents (mothers and fathers) who work in the computing industry, and that they have been exposed to computers and computing from an early age. One teacher reflected upon the role of families in encouraging girls throughout the AfG course, stating that the levels of support differed for different pupils. The teacher used the development of the user survey as an example where on one end, some pupils reported that

they had shared their survey with relatives who praised them for what they had achieved. On the other end, some pupils reported having negative reactions from family members who criticised the work and put the pupil down.

“Some kids have the support at home, that they can go home and they can show their family their app and get tons of feedback, and some kids go home and don't have that support”

(S03).

3. Intention to study STEM subjects at GCSE

Key findings:

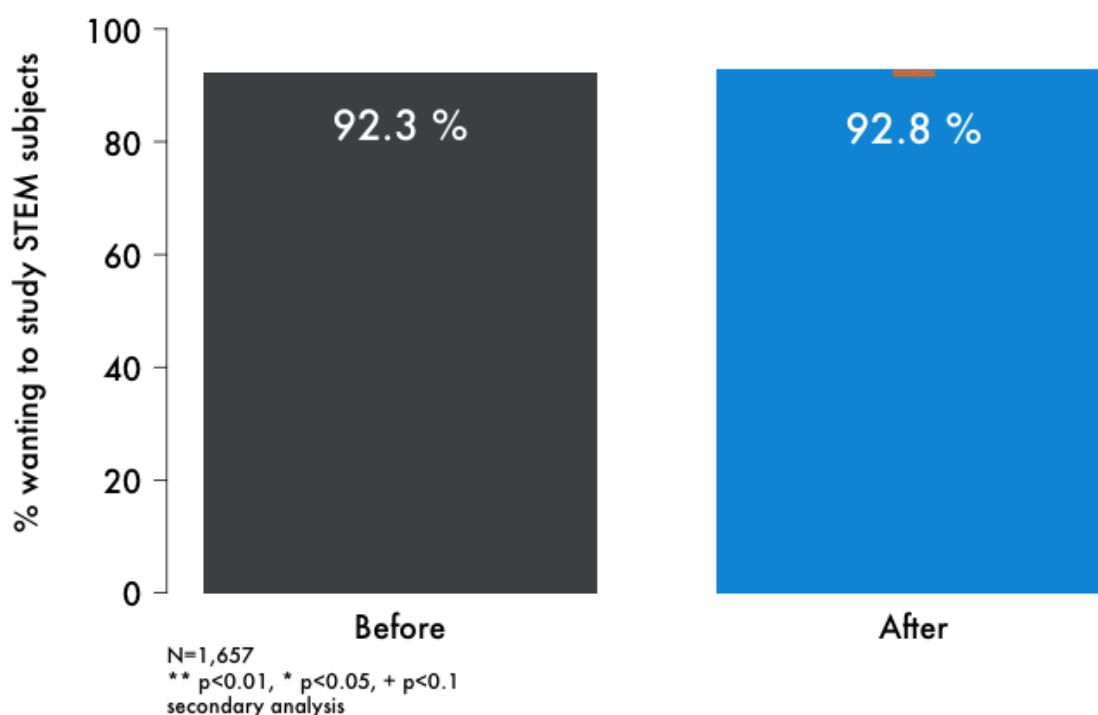
As measured by the survey:

- Girls' intent to study STEM subjects (except computing) stayed the same throughout the intervention

Girls' intention to study STEM subjects

The proportion of girls that intend to study a STEM subject except computer science at GCSE level did not significantly change. We detected a slight increase of 0.5 percentage points, but this effect was not significant. It's important to note that overall there is already a high level of intent to study STEM subjects at GCSE (92% at baseline). The proportion of boys and non-binary pupils wanting to take STEM subjects showed a slight but non significant decrease of 1.5 percentage points. However, these rates are also quite high overall (81.7%), though lower than the girls' baseline.

Figure 5. **Proportion of girls wanting to study any STEM subject at GCSE level (except computing), before and after the intervention.** The bar on the left shows the average likelihood that a girl would intend to study a STEM subject before the intervention began, and the bar on the right after it had ended. Orange lines indicate confidence intervals.



4. Programme delivery

Key findings:

According to interviews with teachers:

- While the workload was deemed heavy by teachers, overall it was manageable
- Teachers felt supported by RPF and suggested that a live Q&A prior to programme launch may be helpful for teachers in future to get their questions answered
- Teachers requested additional materials to help them assess how pupils are progressing through the course
- Demands on teacher time and limited time in the timetable meant that the course was rarely implemented as intended as teachers cut and condensed content

In addition to the outcome measures, the IPE provided evidence on aspects of delivery. This section briefly covers aspects of the intervention implementation that were successful and others that require improvement for future iterations of the course.

1. Workload for teachers

Overall, the teachers interviewed stated that the workload of the course, though heavy, was manageable. Teachers stated that the materials and training explained the purpose of the

programme, and in particular the roles, responsibilities and avenues of support from RPF. Teachers felt confident and supported to start delivering the programme, with the main concern being the amount of content and the teacher's capacity to move through it given the tight school timelines. Teachers spoke of minor technical difficulties that created some stress, including an outdated licence for the app design product, however these challenges were overcome when teachers reached out to AfG and RPF to amend such errors.

2. Support from RPF

Interviewed teachers found RPF to be responsive and supportive as they attempted to fit the course to the needs of the school, the timetable and the pupils. Teachers stated that contacts at RPF provided good advice and helped them redesign the course to meet their needs, however teachers also expressed a desire to have an opportunity to speak with AfG and RPF prior to programme launch. One teacher felt that running an open forum Q&A session while teachers read materials and considered lesson plans would have been more helpful than responding to emails on an individual basis.

“Having that chance to discuss the course probably before I'd started it in a live question and answer session might have been a very useful thing” (S03)

3. Pupil assessment

One teacher spoke about the lack of pupil assessment criteria to go along with the course. The teacher felt that they struggled to assess how pupils were progressing through the course, and stated that as it was a whole term of work, some guidance on assessment would be welcome in future iterations of the course. The teacher's particular concerns related to the collaborative aspect of the course, worrying that some students would rely on stronger or more confident students to get through the work. Instead, the teacher requested some plenary activities or suggested questions they can ask all students at various points to check they are engaged and learning.

4. Delivery of content

Teachers felt that the amount of content to deliver did not fit into their timetable, and that pupils were rarely provided an opportunity to undertake the full AfG programme. Interviewed teachers all described a process of reducing the content prior to the programme starting. One teacher attempted to squeeze the content into five double lessons instead of the prescribed 12 one-hour lessons, stating that this approach seemed to work for them, while other teachers felt forced to cut out course components altogether.

“My biggest issue with the individual lessons was that I think there was way too much stuff in each lesson to get through. Basically, you had to make a decision about what you were going to cut out or skim over, because there's just no way of doing absolutely everything, I found”
(BS03).

Examples of adaptations included teachers consolidating the various idea filtering activities into a single session instead of three. Other classes attempted to complete all activities meaning the teacher lost their ability to track pupil progress through the curriculum, instead moving from activity to activity without assessing pupils in between. One teacher spoke of rushing through the research component, and acknowledged that they did not push students to complete this component thoroughly.

“When they gave their presentations, they were supposed to investigate careers in tech, and I don't think they really did that at all. I did ask them to, but I don't think they did. They just said, 'We don't know any careers in tech,' and stopped” (BS03).

Teachers expressed tensions between their desire to consolidate formal learning and engage in informal learning. Primarily, teachers were concerned over the differences between pupils, where some used creative informal learning time to improve their understanding of computing, and others spent their time without learning a lot. One teacher stated that the limited curriculum time meant that they had to balance creative, informal time with ensuring that everyone is able to continue to move through core curriculum objectives.

“So it's a difficult thing to do, to build that into the curriculum, build enough time for that creative, informal learning, while still making sure everyone makes good progress through the curriculum objectives” (S03).

Furthermore, teachers did not feel that the informal to formal learning link was as explicit as required. One teacher who used the updated materials (which had originally been intended to be tested as part of an RCT) believed themselves to be part of a control group as they could not detect the connecting aspects in the course materials and assumed they were not in there. Given that interviewed girls in part attributed their more positive attitudes toward computing to the creative elements of the course, the balance between informal and formal

objectives is paramount in keeping girls interested in computing but also giving them confidence to take on the challenges of computing as a formal GCSE subject.

Conclusions and lessons learnt

Due to the lack of a control group, we cannot be certain that our intervention has caused any of the effects we observed. It is highly likely that other factors could have caused similar effects on pupils in these 12 weeks that the course was running for, such as coming out of lockdown or other environmental factors. Furthermore, while the aim of the pre-post design was to increase power, this only succeeded to a limited degree as we had to use models that allow for the treatment effect to differ by school (contrary to if this had been an RCT) and that in turn increased the minimum detectable effect size. Finally, the pandemic had an effect on recruitment of schools and attrition during the pilot and the trial, with many students only taking the first survey due to being in self-isolation or lockdown when the second survey was administered.

Interpretation of findings

1. Benefits for girls

The intervention seemed to be of a small benefit to girls' attitudes about computing, however due to design limitations, we are not able to attribute this to the additional informal learning activities. The significant, but small, improvement in attitudes on the pre and post surveys was supported by qualitative data from teachers and pupils who highlighted the course's focus on role modelling computing careers through industry experts and by showing how computing can be relevant in the girls' everyday lives. In this way, the qualitative data suggests that the key mechanisms of change detected in this evaluation relate to *relevance* to the real world and a sense of *belonging* to the computing industry, two core aspects of the programme theory of change which are being explored in parallel interventions and evaluations under the GBIC programme. Interestingly, we do not observe any change in survey scores on the belonging and usefulness subscales of the SCSAS. This may be due to the issues related to teachers reporting on outcomes on behalf of girls, or else the qualitative data has detected enjoyment from the girls, but this enjoyment has not translated into a perceived sense of belonging or relevance.

Girls also focused on the collaborative aspects of the app development, enjoying the learning and teaching process with their peers. Furthermore, girls highlighted the creative process and the ability to work on their own ideas as key features that changed their attitudes; however, teachers felt that this focus was at the expense of the technical components of app development. Neither teachers nor girls attributed any improvements in attitudes to the links between informal learning and GCSE education, but rather that the intervention seemed to help girls enjoy computing and to see it as interesting and useful.

2. Detractors for boys

Where girls' attitudes improved, we observed a decline in boy's attitudes following the course. The AfG course particularly focused on gender and gender balance, with teachers reporting that they were able to secure prominent female experts in the computing industry. Furthermore, as stated above, the course focused on creativity and collaboration, on top of the technical components, and it is possible that the focus on females in computing and on the non-technical aspects left boys feeling worse about computing than when they started. Certainly this is something for AfG to consider in future iterations of the course, as it would appear while it was hoped that gender balance would be stable for boys while improving the attitudes of girls (see the ToC), tipping the focus toward female empowerment may lower boys' interests in the course.

3. Converting interest into intent

Where girls' attitudes may have improved, the findings from the impact evaluation suggest that this interest was not converted into intent to study either computing or other STEM subjects. However, rates of intent to study other STEM subjects was already very high (92%), so there is possibly limited scope for improvement on this outcome. Unfortunately the qualitative findings were limited in their ability to explain the girls' intent and decision making, as the fieldwork took place during the course and not after it had concluded (as initially intended). The evidence collected from teachers suggested that the course may not have adequately prepared girls for the technical aspects of GCSE computer science, and evidence from one teacher suggested that the links between formal and informal learning may not have been explicit enough to be noticed by pupils. This is addressed in the section below discussing lessons for the programme. Furthermore, limited evidence from female pupils and teachers suggested that the intervention had not shifted their confidence with computing or their image of computing as something inaccessible to them. The main enablers in closing the gap were support at home from the family, role modelling from family members and direct intervention from teachers in the form of encouragement over a prolonged period of time. Future interventions in this space may focus on these mechanisms in addition to informal learning to help close the gap between attitudes and intention to study computing.

Lessons for the programme

1. Balancing creative, collaborative and technical components

The qualitative findings indicated a difference between teacher and pupil priorities relating to the AfG course. Interviewed pupils highlighted the importance of collaboration and creativity in changing their attitudes toward computing, enjoying the freedom to do what they would like to do in class, as well as to learn from their peers. Teachers, while acknowledging the importance of relevance and role modelling, felt that the course lacked a focus on the actual app development, spending too much time on idea refinement and presentation. While these aspects may be helpful in improving subject attitudes, they may not result in conversion to GCSE computer science uptake, precisely what is suggested by the impact evaluation findings. On the other hand, in line with the ToC, attitudes may need to change before subject choice decisions are likely to change, and therefore engagement and enjoyment of the course remain crucial mechanisms of change, albeit proximal and not distal. To this end, the course requires a balance between engaging creative and collaborative elements and more challenging technical aspects that teachers believe create a sense of pride, achievement and confidence in girls to continue pursuing computing.

2. Making links to formal computing more explicit

Girls liked many aspects of the course, and there were significant changes in their attitudes toward computing over the 12 weeks. However, qualitative data indicated that the components that were most appreciated were relevance, role modelling, collaboration and creativity. It seems that these aspects may have influenced girls' attitudes, however we did not observe any change in intention to study computing. This may have been as a result of girls not appreciating or understanding the links between informal and formal computing education. Teachers felt that while the course catered for different levels of computing confidence, allowing girls to be creative and to work on their presentation skills, it may not adequately prepared girls for the technical requirements for computer science at GCSE level. One teacher in the treatment group did not even realise they were in the treatment group, as they had not detected the links to formal education in the course. Instead, we suggest that either throughout the course at the end of each lesson, or at the end of the course, teachers are provided with content that explicitly links what the girls have learned to GCSE computer science to help convert interest and positive attitudes into intention and confidence to take up computing following the course.

3. Determining core elements and cutting down content

The interviewed teachers struggled to include the entire course in their timetables over the 12 weeks. Teachers either condensed lessons or cut out whole lessons, seemingly prioritising the technical app-development aspects over the creative, reflective or refinement elements. As both technical and reflective activities have been identified as core components of the course, it may be helpful to review content and refine the course to be more aligned with the realities of school timetables without relying on teachers to cut content. Teachers felt uncomfortable making decisions about how to adapt the course, and felt RPF were helpful in providing guidance for teachers about adaptations. It would be helpful to produce a list of core components which are essential to complete and optional additions if teachers have time in their class schedule. This would help teachers prioritise the aspects of the programme that make links between informal and formal computing more explicit.

Lessons for evaluation

1. A randomised controlled trial would be useful to understand causal impact of the programme, with careful consideration regarding school recruitment

The current evaluation findings present some interesting trends in which we observe outcomes increasing for girls and decreasing for boys and non-binary pupils. Due to the evaluation design, we cannot say whether this is solely due to the programme or some combination of other factors. In order to better measure these trends, running a randomised controlled trial (RCT) with a larger sample size, as originally intended, would be ideal. This would obtain a better estimate of the intervention effect and reduce uncertainty around whether other factors were at play. The AfG team may want to consider whether they would prefer to understand the impact of the enhanced version of the programme compared with AfG without informal linking activities, as originally planned, or the enhanced programme

against business-as-usual. The COVID-19 pandemic posed significant challenges for school recruitment during this evaluation. Given the uncertainty about ongoing COVID-related disruptions, any future evaluation may want to carefully consider school recruitment plans and targets.

2. More comprehensive qualitative research

The COVID-19 pandemic limited our ability to investigate the mechanisms of change across a range of pupils. While we were able to get a range of views from teacher interviews, we were not able to achieve our sample of students or observe the lessons in real time through field visits. Given we observed decreases in boys' attitudes and intent to study computing, future iterations of this course should include more comprehensive qualitative research with boys, particularly focusing on their perspective on the course and whether the focus on creativity and collaboration put them off studying computer science at a GCSE level.

Reference list

Childs, K. (2021). Factors that impact gender balance in computing. In Understanding computing education (Vol 1). Proceedings of the Raspberry Pi Foundation Research Seminar series

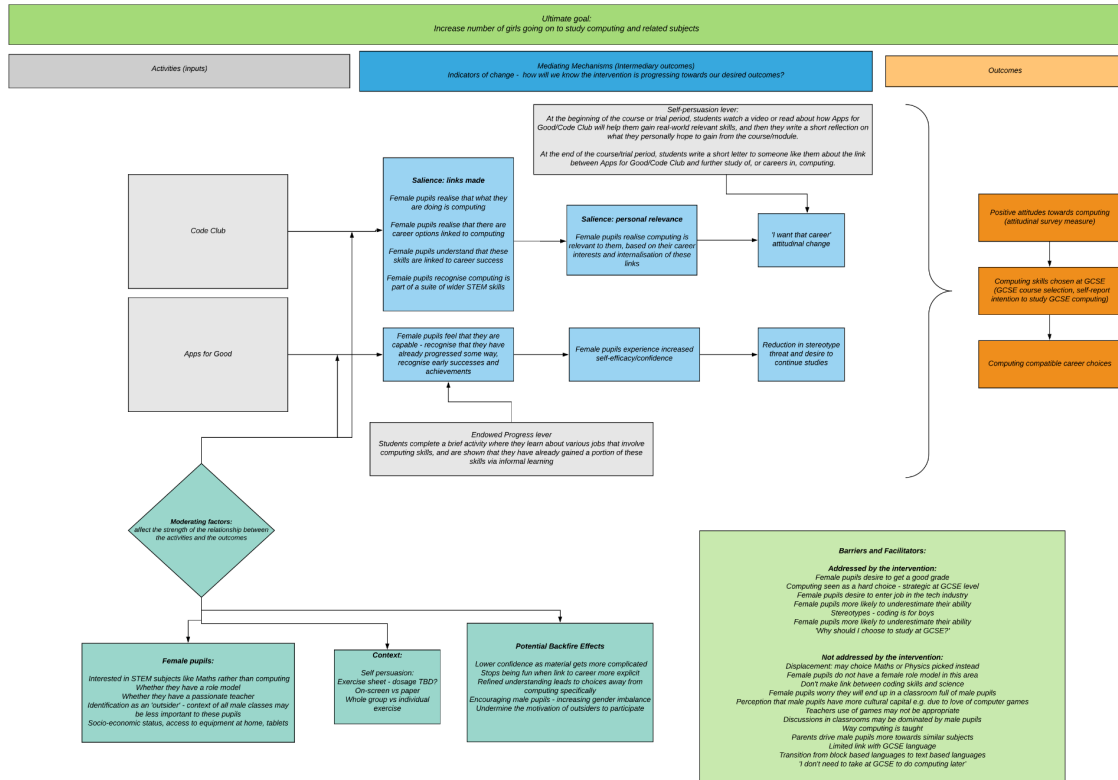
Leadbetter, Hazeldean & Quinlan. (2018). Code Clubs annual surveys 2017, Raspberry Pi Foundation Research, No 4

Royal Society. (2017) After the reboot: computing education in UK schools. <https://royalsociety.org/~media/policy/projects/computing-education/computing-education-report.pdf>

Appendices

Appendix 1: Logic model

Informal Learning Logic Model_V4



Appendix 2: Survey measures

Page 1

Hello! It's time to do the survey.
 Please read each question carefully and take your time to answer.
 Please don't worry about people you know seeing your answers - that won't happen.

1.1	Please type your first name			
1.2	Please type your last name			
1.3	How would you describe your gender?	Female	Male	Non-binary
1.4	Please select the day you were born/month you were born/year you were born Remember to double check the year.			
1.5	Please pick the name of your school from the list below			
1.6	How many of these lessons have you been to? (only in endline survey)			
1.7	Do you want to study any of these subjects for GCSE?			
	Computing	Yes	No	Don't know
	Physics	Yes	No	Don't know
	Biology	Yes	No	Don't know
	Chemistry	Yes	No	Don't know
	Combined Science	Yes	No	Don't know
	Design & Technology	Yes	No	Don't know
	Engineering	Yes	No	Don't know

Page 2

Thanks! Now it's time for the rest of the questions.

2.1	I have self-confidence when it comes to computing	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly Agree
2.2	I am confident that I can solve problems by using computing	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly Agree
2.3	I can learn computing without a teacher to explain it	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly Agree
2.4	I am sure I could do advanced work in computing	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly Agree
2.5	I think I will do well in computing	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly Agree
2.6	I would take additional computing classes if I were given the opportunity	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly Agree
2.7	I hope that my future career will require the use of computing	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly Agree
2.8	I like to use computing to solve problems	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly Agree
2.9	The challenge of solving problems using computing appeals to me	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly Agree
2.10	I like writing computer programs	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly Agree

Page 3

Well done! Keep going - you are already half way through.

2.11	I feel comfortable in computing	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly Agree
2.12	I feel I belong in computing	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly Agree
2.13	I feel accepted by my classmates in computing	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly Agree
2.14	I know someone like me who uses computing	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly Agree
2.15	I know a lot of pupils like me who are interested in computing	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly Agree
2.16	Knowledge of computing will help me earn a living	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly Agree
2.17	Learning to use computing skills will help me achieve my career goals	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly Agree
2.18	Computing skills help me understand things in everyday life	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly Agree
2.19	I'll need to be good at computing for my future work	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly Agree
2.20	Computing is a worthwhile and necessary subject	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly Agree

Page 4

Almost done!

2.21	A friend or someone in my year group has encouraged me to study computing	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly Agree
2.22	Someone I know has given me the desire to study computing	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly Agree
2.23	Someone I know has praised my work in computing	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly Agree
2.24	Someone I know has discussed the subject of computing with me	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly Agree
2.25	Someone in my family has encouraged me to study computing	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly Agree

Page 5

You have completed this survey!

Thank you for taking the time to answer this survey.

Appendix 3: Changes to SCSAS

The original SCSAS was designed for pupils aged 13+ in the US.⁴ For the current trial, the language of the SCSAS was adapted to a UK sample and underwent cognitive testing to ensure pupils understood all questions.

Confidence

Original

- I have self-confidence when it comes to computer science.
- I am confident that I can solve problems by using computing.
- I can learn computer science without a teacher to explain it
- I am sure I could do advanced work in computer science.
- I think I will do well in computer science.

Adapted

- I have self-confidence when it comes to **computing**.
- I am confident that I can solve problems by using computing.
- I can learn **computing** without a teacher to explain it
- I am sure I could do advanced work in **computing**.
- I think I will do well in **computing**.

Interest

Original

- I would take additional computer science courses if I were given the opportunity.
- I hope that my future career will require the use of computer science.
- I like to use computer science to solve problems.
- The challenge of solving problems using computer science appeals to me.
- I like writing computer programs.

Adapted

- I would take additional **computing classes** if I were given the opportunity.
- I hope that my future career will require the use of **computing**.
- I like to use **computing** to solve problems.
- The challenge of solving problems using **computing** appeals to me.
- I like writing computer programs.

Belongingness

Original

- I feel comfortable in computer science.
- I feel I belong in computer science.
- I feel accepted by my peers in computer science.
- I know someone like me who uses computer science in their work.

⁴ Haynie and Packman (2017). Available at: <https://csedresearch.org/tool/?id=156>.

- I know a lot of students like me who are interested in computer science.

Adapted

- I feel comfortable in **computing**.
- I feel I belong in **computing**.
- I feel accepted by my **classmates** in **computing**.
- I know someone like me who uses **computing in their work**.
- I know a lot of **pupils** like me who are interested in **computing**.

Usefulness

Original

- Knowledge of computer science will help me earn a living.
- Learning to use computing skills will help me achieve my career goals.
- Computing skills used to understand computer science material can be helpful to me in understanding things in everyday life.
- I'll need a mastery of computer science for my future work.
- Computer science is a worthwhile and necessary subject.

Adapted

- Knowledge of **computing** will help me earn a living.
- Learning to use **computing** will help me achieve my career goals.
- Computing skills **help me understand things in everyday life**
- **I'll need to be good at computing for my future work**
- **Computing** is a worthwhile and necessary subject.

Encouragement

Original

- A friend or peer has encouraged me to study computer science.
- Someone I know has given me the desire to study computer science.
- Someone I know has praised my work in computer science.
- Someone I know has discussed with me the computer science field.
- Someone in my family has encouraged me to study computer science.

Adapted

- A friend or **someone in my year group** has encouraged me to study **computing**.
- Someone I know has given me the desire to study **computing**.
- Someone I know has praised my work in **computing**.
- Someone I know has discussed **the subject of computing** with me.
- Someone in my family has encouraged me to study **computing**.

Appendix 4: Model specification

We used multilevel models to assess the Intention-To-Treat (ITT) effect of our treatment on the outcomes. Owing to the clustered nature of the data, we used random effects to account for different baselines and treatment effects between schools and pupils. We added fixed effects for the trial run (2020 or 2021) and timing of the observation (before or after the intervention). As recommended by previous research⁵, we included the maximal random effects structure *justified by the design*. We used the Satterthwaite approximations for degrees of freedom for determining significance (a discussion of why this method is preferred can be found in Luke (2017)⁶).

Primary outcome: SCSAS scores

The primary outcome is mechanically bounded above and approximately normally distributed (see Figure 6 below), and therefore we used a multilevel linear regression model to assess the Intention-To-Treat (ITT) effect of our treatment on this outcome. Owing to the clustered nature of the data, we used random effects to account for different baselines and treatment effects between schools. As recommended by previous research⁷, we included the maximal random effects structure *justified by the design*.

$$Y_{ji} = \beta_0 + \mu_{0ji} + \beta_1 t_{ji} + \mu_{1ji} t_{ji} + \beta_2 y_{ji} + \varepsilon_{ji}$$

$$\varepsilon_{ji} \sim N(0, \sigma^2)$$

Where:

- Y_{ji} is the SCSAS survey score for pupil i in school j
- β_0 is the intercept
- μ_{0ji} is the random intercept for pupil i in school j
- t_{ji} is the treatment indicator for pupil i in school j

⁵ Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language*, 68(3), 255-278.

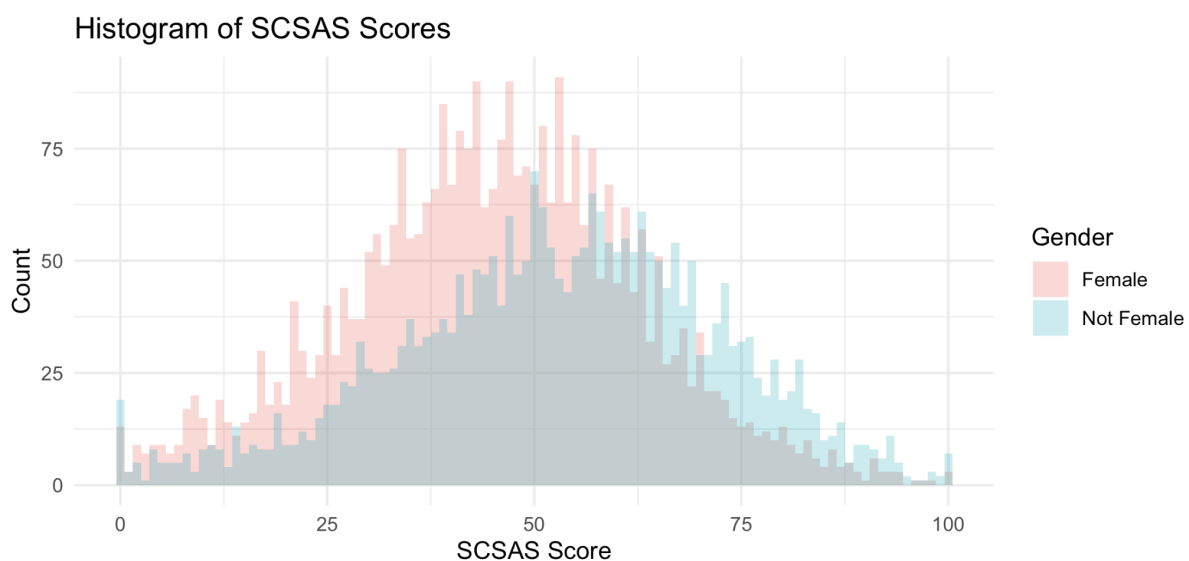
⁶ Luke, S. G. (2017). Evaluating significance in linear mixed-effects models in R. *Behavior research methods*, 49(4), 1494-1502.

⁷ Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language*, 68(3), 255-278.

- $t = 0$ for pre-treatment
- $t = 1$ for post-treatment
- β_1 is the treatment coefficient
- $\mu_{1ji}t_{ji}$ is the random treatment coefficient for pupil i in school j
- β_2y_{ji} is a dummy variable indicating the school year (either 2020 or 2021)
- ε_{ji} is the error term

We used the Satterthwaite approximations for degrees of freedom for determining significance (a discussion of why this method is preferred can be found in Luke (2017)⁸). If the model showed convergence warnings or signs of a singular fit, we iteratively reduced the random effects structure until no problematic random-effects estimates were observed.

Figure 6. Distribution of SCSAS scores in our sample, for female and non-female students.



Secondary outcome 1: stated intention to study computer science at GCSE level

As the secondary outcome is binary and we have a repeated measures design, we used a multilevel logistic regression model to assess the Intention-To-Treat (ITT) effect of our treatment on this outcome. Similar to the primary outcome variable, we used random effects to account for different baselines and treatment effects between schools.

⁸ Luke, S. G. (2017). Evaluating significance in linear mixed-effects models in R. *Behavior research methods*, 49(4), 1494-1502.

$$\log \left(\frac{\Pr(y_{ji} = 1)}{1 - \Pr(y_{ji} = 1)} \right) = \text{logit}(\Pr(y_{ji} = 1)) = \beta_0 + \mu_{0ji} + \beta_1 t_{ji} + \mu_{1ji} t_{ji} + \beta_2 y_{ji}$$

$$Y_{ji} \sim \text{Bern}(\phi_{ji})$$

Where:

- y_{ji} is a binary indicator for pupil i in school j reflecting whether the pupil intends to study Computer Science at GCSE
- β_0 is the intercept
- μ_{0ji} is the random intercept for pupil i in school j
- t_{ji} is the treatment indicator for pupil i in school j
 - $t = 0$ for pre-treatment
 - $t = 1$ for post-treatment
- β_1 is the treatment coefficient
- $\mu_{1ji} t_{ji}$ is the random treatment coefficient for pupil i in school j
- $\beta_2 y_{ji}$ is a dummy variable indicating the school year (either 2021 or 2020)

Again, we used the Satterthwaite approximations for degrees of freedom for determining significance. If the model showed convergence warnings or signs of a singular fit, we iteratively reduced the random effects structure until no problematic random-effects estimates were observed.

Secondary outcome 2: stated intention to study STEM

The second secondary outcome measured whether the pupil intends to study any STEM subject except computing. This included physics, biology, chemistry, combined science, design & technology or engineering.

As the outcome is binary, the same mixed model specification as for the first secondary outcome was used, substituting the outcome variable accordingly.