Early Developmental Activities and Computing Proficiency

Quintin Cutts
University of Glasgow
Glasgow, United Kingdom
quintin.cutts@glasgow.ac.uk

Elizabeth Patitsas
University of Toronto
Toronto, Canada
patitsas@cs.toronto.edu

Elizabeth Cole
University of Glasgow
Glasgow, United Kingdom
e.cole.2@research.gla.ac.uk

Peter Donaldson
University of Glasgow
Glasgow, United Kingdom
peter.donaldson.2@glasgow.ac.uk

Bedour Alshaigy
Oxford Brookes University
Oxford, United Kingdom
bedour.alshaigy-2012@brookes.ac.uk

Mirela Gutica
British Columbia Institute of Technology
Vancouver, Canada
Mirela_Gutica@bcit.ca

Arto Hellas
University of Helsinki
Helsinki, Finland
arto.hellas@cs.helsinki.fi

Edurne Larraza-Mendiluze
University of the Basque Country
Gipuzkoa, Spain
edurne.larraza@ehu.eus

Robert McCartney
University of Connecticut
Mansfield, USA
robert@engr.uconn.edu

Charles Riedesel
University of Nebraska
Lincoln, USA
chuckr@unl.edu

ABSTRACT

As countries adopt computing education for all pupils from primary school upwards, there are challenging indicators: significant proportions of students who choose to study computing at universities fail the introductory courses, and the evidence for links between formal education outcomes and success in CS is limited. Yet, as we know, some students succeed without prior computing experience. Why is this?

Some argue for an innate ability, some for motivation, some for the discrepancies between the expectations of instructors and students, and some – simply – for how programming is being taught. All agree that becoming proficient in computing is not easy. Our research takes a novel view on the problem and argues that some of that success is influenced by early childhood experiences outside formal education.

In this study, we analyzed over 1300 responses to a multi-institutional and multi-national survey that we developed. The survey captures enjoyment of early developmental activities such as childhood toys, games and pastimes between the ages 0 — 8 as well as later life experiences with computing. We identify unifying features of the computing experiences in later life, and attempt to link these computing experiences to the childhood activities.

The analysis indicates that computing proficiency should be seen from multiple viewpoints, including both skill-level and confidence. It shows that particular early childhood experiences are linked to parts of computing proficiency, namely those related to confidence with problem solving using computing technology. These are essential building blocks for more complex use. We recognize issues in the experimental design that may prevent our data showing a link between early activities and more complex computing skills, and suggest adjustments. Ultimately, it is hoped that this line of research will feed in to early years and primary education, and thereby improve computing education for all.

CCS CONCEPTS

•**Social and professional topics → Computational thinking;** *Informal education; Computing literacy; K-12 education;*

KEYWORDS

Pre-requisite, pre-computational, thinking, proficiency, primary, early-years

ACM Reference format:
Quintin Cutts, Elizabeth Patitsas, Elizabeth Cole, Peter Donaldson, Bedour Alshaigy, Mirela Gutica, Arto Hellas, Edurne Larraza-Mendiluze, Robert McCartney, and Charles Riedesel. 2017. Early Developmental Activities and Computing Proficiency. In *Proceedings of ACM Innovation and Technology in Computer Science Education conference, Bologna, Italy, July 2017 (ITiCSE'17),* 18 pages.
DOI: 10.1145/nnnnnnn.nnnnnnn

ITiCSE'17, Bologna, Italy

DOI: 10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

As countries adopt CS education for all pupils from primary upwards, there are challenging indicators: traditionally, significant proportions choosing to study CS have failed introductory courses [10, 76]; recent studies suggest that not all pupils achieve intended outcomes using block-based languages [3, 28, 50]; and there is limited

evidence for links between formal education outcomes and success in CS [8, 61, 67]. It is a particularly important question as we move from CS as a largely optional subject to one that all pupils must take early in their school career. How can we ensure that all learners have a fair chance of success if we are unable to explain why some students fail while others succeed? For, as we know, some students do succeed in our courses without prior computing experience, while others, seemingly similar, fail. Why is this?

This study aims to shed light on this question and is based on the hypothesis that early childhood experiences *outside* formal education are a predictor of computational thinking skill/understanding demonstrated later in life. Given the uncertainty over the so-called *geek gene* [13, 23] or strong academic prerequisites for success in CS, this is intuitively appealing. We choose the broader measure of CT skills demonstrated later in life, rather than success in a CS course, to recognize those who, for example, successfully learn to program or are power-users of technology such as spreadsheets and databases without taking specific courses. As we argue in Section 6, the skills needed for complex problem solving with technology are also required in order to succeed in CS courses. The purpose of the study is therefore to look for evidence for or against this hypothesis.

To this end, we produced a survey that asks adults about

(1) their experiences, preferences, and activities as a child up to eight years old,
(2) their **current** level of success with computing and programming, and
(3) demographic information: their age, gender, and country where they were a child.

This survey, then, provides data about the participants at two points in their lives and our goal is to see if there is any relationship between these two points. Identifying a relationship between childhood activities and adult computing skills could be useful to educators in at least two ways: learners' earlier activities could be used to predict future computing success; alternatively, one could encourage children to engage in those activities that enable future computing success, ultimately improving outcomes in our classrooms.

1.1 Research questions

Our primary research questions are:

(1) Are there any childhood games and activities that are predictive of computing skills later on?
(2) If so, to what degree are particular childhood activities predictive of later computing skills?

The particular kinds of computing skills that we are concerned with lie on a spectrum from competent use of computing technology to sophisticated problem solving, and on to success in programming and software engineering.

2 RELATED WORK

Work relating to this study includes work about the relationships between play and development in general, and work that relates aptitude, experience, or behavioral attributes to success or failure of computing students.

2.1 Early developmental activities

This study requires a broad understanding of play categories in order to ensure that any survey properly captures respondents' early activities. Whitebread et al. [22] provide an overview of the anthropological, sociological, historical, psychological and educational research concerned with children's play. They acknowledge the challenges of play categorisation in an academic landscape where for every aspect of a child's development, there is a form of play, as proposed in [51]. However, their framework provides a sound foundation to base the selection of toys, games and activities for the survey, identifying five core types of play — physical play, play with objects, symbolic play, pretence/socio-dramatic play, and games with rules. The framework shows a link between play and cognitive development and takes account of Vygotsky's widely supported view that play is the first medium through which children explore the use of symbol systems [75].

Of Piaget's four childhood developmental stages [55], we wanted to explore mainly the preoperational stage (before the age of eight) when children engage in pretend and symbolic play with objects. Symbolic play was researched by Piaget [55]; however he does not separate symbolic play from playing with objects as Whitebread et al. do. We believe that symbolic play with objects is important for abstraction which is an important aspect of logical and computational thinking. Playing with objects is important as it "allows children to try out new combinations of actions, free of external constraint, and may help develop problem solving skills" [68].

Table 1 shows a breakdown of Whitebread's framework, highlighting some of the main developmental functions/foci including the main psychological benefit and typical developmental trajectories in physically and psychologically healthy children.

2.2 Innate capabilities

The claim that the ability to program or not is innate and measurable was proposed by Dehnadi and Bornat[23] as the "two humps" hypothesis. They claimed that grades in CS1 were bimodal, reflecting two populations of students, those who can learn computing and those who cannot. Although Bornat later retracted their claims [13], a number of researchers examined these claims. Caspersen et al. [17] tested that instrument with their students and found that the instrument was not a good predictor. Robins [60] investigated the "characteristic bimodal grade distribution" in CS1, and presented a model that explains the presence of extreme grades. The notion that CS grade distributions are commonly bimodal has been debunked by Ahadi and Lister [2] and Patitsas et al. [54]. Patitsas et al statistically tested nearly 800 different grades distributions and found bimodal grade distributions were extremely rare, and normal distributions to be most common.

A counter position to innate characteristics predicting success or failure is provided by Dweck's [25] notion of fixed versus growth mindsets–whether students see their abilities as something that can change over time based on their own efforts. Influenced by this, Lewis [45] asked faculty and senior students whether they agreed with the statement, "Nearly everyone is capable of succeeding in the computer science curriculum if they work at it"; 77% of the faculty members disagreed or strongly disagreed, evidence of a fixed mindset, while 59% of the senior students agreed or strongly agreed,

Table 1: Types of play and their connections to childhood development [22].

Type of play	Description	Development function / focus
Physical play	Active exercise. Jumping, climbing, dancing, skipping, bike riding and ball play.	Whole body and hand-eye coordination, strength and endurance.
	Rough-and-tumble with friends, siblings or parents/guardians. Chasing, grappling, kicking, wrestling and rolling on the ground. Distinguished from actual aggression through enjoyment by participant.	Developing emotional and social skills and understanding. Controlling aggression, independence, resourcefulness and self-regulation.
	Fine-motor practice. Sewing, colouring, cutting, junk modelling and manipulating action and construction toys.	Fine motor hand and finger co-ordination. Often solitary and due to absorbing nature develop children's concentration and perseverance skills.
Play with objects	Building, making and construction. Exploring and arranging objects and materials, sorting and classifying. Often within pretence or socio-dramatic context and a narrative is involved. The narrative (private speech) involves children commenting on their activity and setting themselves goals and challenges, monitoring their progress towards them.	Perseverance, positive attitude and self-regulation towards challenge. Investigative behaviour, thinking, reasoning, and problem-solving. The private speech increasing repertoire of cognitive and physical skills and strategies and generally self-regulate.
Symbolic play	Reading, writing, number, painting, drawing, collage, music. (Playing with language) making new rhymes, mark making, jokes, counting, sing dance making sounds with their own bodies and objects.	Expression, reflecting on experiences, ideas and emotions. Early literacy and numeracy skills. Develops childrenś ability to understand pictures, photographs, diagrams, scale models, plans, maps.
	Music	Social interaction, communication, emotion understanding, memory, self-regulation, communication and self-regulation and creativity. Cooperative behaviour.
Pretence/Sociodramaticplay	High-quality pretend play including private speech. Dolls, role play.	Cognitive, social and academic abilities. Narrative skills. Representational and self-regulatory. Children follow social rules governing the character, improvement of socio responsibility.
Games with rules	Includes making their own games with rules. Chasing, hide and seek, throwing, catching, board games, card games, electronic and computer games and sporting activities.	Helps young children make sense of the world. Social skills related to sharing, taking turns, understanding other's perspectives.

evidence of a growth mindset. This is consistent with McCartney et al. [48], who found that many CS instructors believe that innate capabilities are predictive of student success to some degree.

2.3 Other predictive factors

Wilson and Schrock [79] developed a model using (up to) 12 factors to predict midterm scores in CS1; their most effective predictive factors were comfort level, math experience, and how students attributed their success or failure on the midterm. They also distinguished the effects of previous computer experience. Similarly, Bergin and Reilly [12] developed a model to predict students' performance in introductory programming courses, and found that using only factors measurable at the start of the course (such as math scores and previous computing experience) resulted in poor prediction; including course-related factors measured during the course such as score on first test and student comfort level improved the prediction. Simon et al. [67] found that students who go on to be successful programmers (as measured by CS1 marks) "tend to have pre-existing strengths in a strategic / algorithmic style of articulation."

Leeper and Silver [44] used factors from before University – SAT exam scores, number of units of Mathematics, English, Science, and Foreign Language taken in High School, and High School class rank – in a regression to predict grade received in CS1. They found that their predictive equation was significant, but only accounted for about 26% of the variation in CS1 grades. Rountree et al. [61] used class survey data to predict performance in CS1, and found that the strongest predictor of success was the expectation of receiving a high grade in the course.

A number of researchers have found behavioral attributes to be the best predictors. Watson et al. [77] found that students' programming behavior was a far more effective predictor than 38 traditional predictors of programming performance. Ventura [74] found that "measures of effort are the primary predictors of success" in a programming course. Höök and Eckerdal [33] investigated the relationship between final CS1 exam results and attributes of the course, including time students spent coding, time spent reading

and listening to theory, and whether they perceived the course as fun, difficult, or intimidating. The factor that had the highest impact on exam results was the amount of time students spent coding.

2.4 Summary

Drawing together the related work, innate capabilities may be recognised by staff, but are not upheld by subject-specific or general education research. There is some evidence of prior academic skills influencing success, but they do not account for a large proportion of the variation. Rather than pre-course predictors, many researchers have found that behaviour on the course is a better predictor. How childhood activities outside formal education influence CS skills later in life appears to be unexplored. However, there is a good agreement on an encompassing range of play categories on which to base our survey.

3 METHODOLOGY

3.1 The survey

To address RQ1, we chose to develop a survey that collects information about respondents' early activities, their later CS skills, and demographics. We use "early activities" as a shorthand for the toys, games and activities that respondents recalled from their early life. We set the bar for "early life" as age 8 or less because we are interested in applying any relevant findings to pre-school and early primary education. We recognize that recall of activities from that age may be considered to be problematic. However, we believe that participants' responses can be trusted based on our literature review as follows: Studies indicate that adults recall well childhood memory in terms of activities, location and who was present; however recollection of detail is generally poor [78]. Additionally, Howes, Siegel and Brown [35] found that especially positive or emotionally neutral memories are accurate. Hardth and Rutter [31] found from a comprehensive review study that "recall of experiences that rely heavily on judgment and interpretation have not been found to have satisfactory validity" (p.270). In our study we do not ask our participants to interpret or make judgments, we ask only for recall of activities which correspond to positive or neutral memories. Issues around accuracy of memories and reliability of memory recall are discussed in more detail at the end of the Discussion section.

In the overall structure of the survey, which is reproduced in full in the Appendix, there was a deliberate attempt not to bias respondents' thinking before they started the questionnaire. The second question asks them to recall their favorite three early activities, responding in free text, preceded only by text and a question about preliminary consent to continue with the survey. We defer asking for full consent until the end of the survey, explaining to the respondent that we are not going to explain the purpose of the survey until then, to avoid biasing their responses.

We deliberately aim to be vague on the definition of the CS skills we are looking for, as that is a topic of significant debate. Instead, we are looking for activities later in life that require broadly the skill set we are interested in. Where responses to the questions on programming success and ease (Q8 and Q9) clearly indicate that the respondent has written programs for themselves or for money, we infer that they have the relevant skills we are looking for. This leaves a number of respondents for whom the proficiency in skills we are looking for is unclear. Question 7 aims to assess whether such respondents have our target skill set. It is based around use of IT, on the assumption that all respondents would use IT to some degree. It is recognized from literature [18, 32, 40, 80] and our practice that nowadays programs are written by specialists in other domains who do not identify themselves as programmers and may not have ever been in a programming class. Dorn et al. [24] (p.30) report of the projects completed by end-users that "the projects were typical of what one might expect from a novice programmer". So, the assumption is that the complexity of the IT use is related to our target skill set. The elements of Q7 relate to commonly accepted themes around use of technology:

- whether a user can set up / get started with IT systems or tend to seek help from others;
- whether a user makes use of the advanced features of typical software;
- Whether a user helps others to fix their IT problems, or are recognized by friends/colleagues as the person who can debug IT issues; and
- whether a user writes scripts, macros or programs associated with their IT use.

3.2 Data collection

The surveys were distributed via Survey Monkey. The participants in our study voluntarily responded to the survey. We did not prior disclose the content or the goal of the survey and no monetary compensation was offered for completion. We recruited our participants in various ways: invited friends, family, members on non-profit organizations (e.g., church) and close work colleagues to participate and share the survey within their networks; invited undergraduate and postgraduate students from our institutions by sending mass e-mails; invited members of student clubs by sending the link to the club chairs; posted the questionnaire on our Facebook, LinkedIn and Twitter accounts; and posted the questionnaire on our institutions' websites, internal shared space, and also national and international networks.

Our survey did not ask anything about the educational background of participants. However, the relatively even distribution of different job titles, ages, types of childhood experiences and confidence and skill in solving complex problems using computing technology suggest respondents were drawn from a wide range of different backgrounds. More detail on the likely backgrounds of participants is provided in section 4.1.

3.3 Data cleanup

Survey Monkey was used as the primary data collection tool by all of the researchers. They either used a specific collector link to the English language survey or were given access to a copy of the survey to translate and then localise.

After the collection period ended there were a total 2002 survey responses recorded. These were then filtered to remove partially completed responses where participants had exited the survey prematurely giving 1340 responses and an average response rate of 65.89%. The final step, before exporting the data, was to filter just the responses where consent had been given after the full purpose

Table 2: Summary of data sources

Collector	Number	Description
Collector 1	273	(1) personal Twitter, Facebook accounts, (2) work Twitter, Facebook and internal shared space of national education organization, (3) link was sent for sharing to a friend teaching abroad.
Collector 2	198	(1) national and international work contacts, (2) European project platform, (3) teacher communities and on-line course participants, (4) CoderDojo community, (5) social networks Twitter and Facebook, (6) past students
Collector 3	120	(1) closed CS Faculty Facebook group , (2) open Facebook group for ICT in Education, (3) students in two courses , (4) research assistants and their friends and family.
Collector 4	120	(1) 25% friends and family; (2) 75% faculty, postdocs and students in local academic department
Collector 5	101	(1) Faculty members at local institution, (2) professionals from industry contacted directly or via LinkedIn, (3) students from one of the local courses, (4) graduate students from local institution from non-CS disciplines, (5) family friends and Facebook friends groups.
Collector 6	95	(1) friends and family, (2) members of the elementary school parent association, (3) members of a national organization including professionals and academics in many disciplines, (4) faculty professors, researchers (mainly from the areas of computing and linguistics) and students.
Collector 7	93	(1) undergraduate students from local academic department, (2) academic staff from the local academic department, (3) personal contacts, friends and family, on Facebook and Twitter, (4) professional contacts from different disciplines via LinkedIn.
Collector 8	83	friends and colleagues and their connections: (1) industry professionals (including IT and other industries), (2) engineering undergraduate, Masters and Ph.D. students across the country.
Collector 9	73	friends, distant family, national teacher association members and their connections (1) Industry professionals (includuing IT, Legal, Health and Retail industries) (2) Secondary and Primary School teachers
Collector 10	71	labour union at local institution from all disciplines who are (1) graduate students and postdocs, (2) instructors/TAs
Collector 11	59	(1) friends and family members with a link, suggested that they pass it onto their acquaintances, and (2) daily email posting that goes to all faculty and staff at University of Connecticut.
Collector 12	42	Family, local Kiwanis club members, Church, Honors class (second year algorithms and data structures) at local institution
Total	1329	

of the survey had been explained. This left 1332 responses that were then exported and combined into the master data set.

All responses to restricted choice items needed to be exported as text labels rather than numeric values due to a Survey Monkey bug that meant some questions had different text response values mapping to the same numeric value. Text response values for childhood activities and adult experiences in different languages were translated to their English language equivalents. In addition, some of the "response required" settings were not copied over to two of the foreign language versions of the survey which meant that 38 of the complete responses contained one or more missing values. Three of these entries had so many values missing that they were excluded from the dataset leaving 1329 complete responses. Any unexpectedly missing value was marked with a -1 to distinguish it from an optional response that had been left blank and items that had been marked as non-applicable which were coded as 0.

Values for country varied by method of data entry, with fixed choice responses containing a two letter country code in addition to the name of the country and free text entry containing just the name. All two letter prefixes were removed and the name of the country was also capitalized.

The final stage in the cleanup process involved mapping fixed response text values to an ordinal value or categorical value that could be subsequently analyzed.

For the demographic information:

- age range was assigned an ordinal value starting at 1 for the youngest group up to 7 for the oldest group
- gender was assigned a categorical value between 1 and 4.
- jobs were assigned a categorical value indicating whether Computing skills would be required with (1) unknown, (2) not required and (3) required.

For the other response options:

- level of enjoyment was assigned an ordinal value from (1) none or low up to (5) very high.
- level of agreement was assigned an ordinal value from (1) strongly disagree through to (5) strongly agree.
- frequency of activity was assigned an ordinal value from (1) not at all to (5) a lot.
- programming proficiency was assigned an ordinal value from (1) never tried to program at all through to (5) you have successfully written programs mainly requested/ paid for by others.

Table 3: Gender

Gender	Count	%
Female	818	61.6%
Male	492	37.0%
Other	8	0.6 %
Did not disclose	11	0.8%

- relative difficulty in learning to program was assigned an ordinal value from (1) no experience through to (5) they found learning to program much easier than most others.

4 ANALYSIS AND RESULTS

A strict statistical analysis of survey data was conducted as well as a more exploratory analysis. The former is described in this section, the latter in the following section.

4.1 Data summary

The participants in this study, after data cleaning, were 1329 individuals who represent different age groups, genders and have different levels of computer experience. Our participants spent their first eight years of age in 67 different countries over all continents. Table 5 presents the top 10 countries. As we see in Table 4, we have representation for each age-group with similar percentages for young and middle-age participants. The largest group is represented by 45 - 54 year olds (22.3%) and the smallest is 75 and above (the survey was responded by only 5 participants). More women than men responded the survey (ratio of 1.66/1); eight participants reported their gender as "other", and 11 did not disclose their gender (see Table 3). Table 6 represents the level of computer experience. As we can see, we have a balanced number of participants in each category. The majority of the participants who have been in a programming class (75.8%) responded that it was much easier or easier for them than for other individuals in the class (see Table 7).

As explained in Section 3.2, we sampled our data from a variety of sources including students and individuals working in different fields including computing and non-computing professions. Our survey results are a convenience sample due to the use of our own networks and are therefore less likely to represent the population at large. As our networks include students and generally more professionals, we report that the majority of participants are from medium to upper level economic classes. We believe that there is no bias in our selection of participants as we compare individuals with CS background to individuals with other professions.

There is evidence to suggest that sociological differences in access to early cultural capital exist [42]; however our study doesn't address this issue.

4.2 Quantitative Analysis

To answer our research questions, we wanted to perform an ordinal logistic regression using the childhood activity questions as predictors, and some measure(s) of adult computing as the outcome variable(s). We selected this stricter form of regression as the included response values have a relative ordering but the exact difference between the levels of response is unknown.

Table 4: Age – this sums up 100.1%

Age group	Count	%
18 - 24	236	17.8%
25 - 34	290	21.8%
35 - 44	283	21.3%
45 - 54	296	22.3%
55 - 64	172	12.9%
65 - 74	46	3.5%
75 -	5	0.4%
Did not disclose	1	0.1%

Table 5: Country – top 10

Country	Count	%
Scotland	271	20.4%
Italy	170	12.8%
England	140	10.5%
United States	137	10.3%
Finland	117	8.8%
Canada	98	7.4%
India	88	6.6%
Basque Country	86	6.5%
Netherlands	36	2.7%
Romania	22	1.7%
(Other)	165	12.3%

Table 6: Level of programming

Level	Count	%
Never tried to program	429	32.3%
Tried, but did not feel successful	268	20.2%
Successfully written programs, for self	290	21.8%
Successfully written programs, for others	342	25.7%

Table 7: Programming easiness when compared to others; not all reported (maybe people in online courses where you cannot compare?) – n = 752

Easiness	Count	%
Much easier than others	192	25.5%
Easier than others	378	50.3%
Harder than others	158	21.0%
Much harder than others	24	3.2%

4.3 Cluster analysis of adult computing questions

We asked 12 questions about adult computing use however we needed to ascertain whether they measured different aspects of computing proficiency or the same underlying characteristic(s). To try to reduce the dimensionality of this set of questions, we

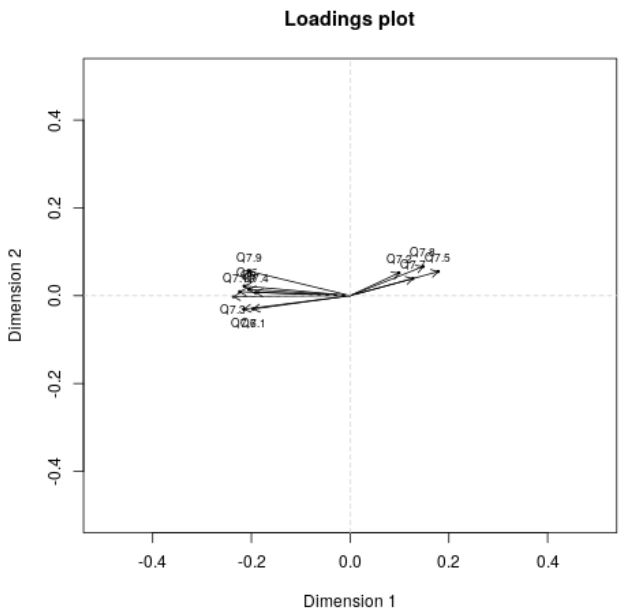

Figure 1: NPCA of Qs 7-9

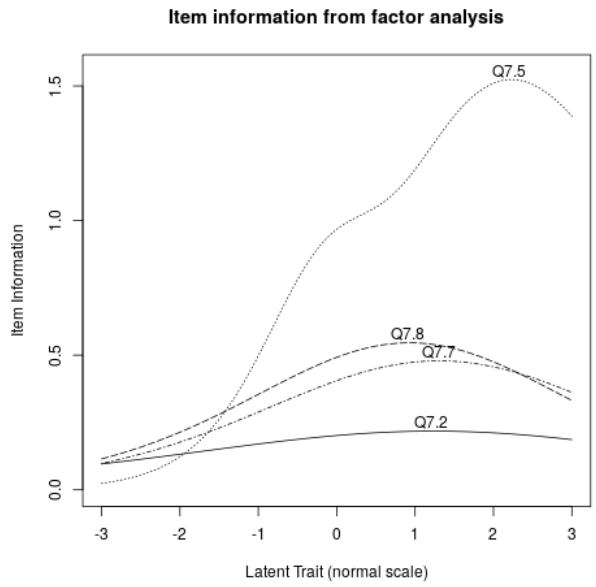

Figure 2: IRTFA1

performed a non-linear principal component analysis (NPCA). This was done with the `homals` package in R version 3.4.

As seen in Figure 1, two clusters were present in the set of twelve adult computing questions. We ran the NPCA with higher dimensions (3 through 6) and did not find the higher-dimension models had less loss than the 2-dimensional NPCA.

The two clusters were:

(1) Cluster 1: Q7.2, 7.7, 7.8, 7.5 — questions generally related to confidence in understanding and using computing technology
(2) Cluster 2: Q7.4, 8, 7.6, 7.1, 7.10, 9, 7.3, 7.9 — questions generally related to complexity of computational problems they can solve

4.4 Identifying proxies for the two clusters

At this point we faced a design decision: whether to use the principal components (or a varimax rotation thereof) as our measures of adult computing, or to use one question from each cluster as a proxy for the cluster. We chose the latter to improve reproducibility: this way future studies can use the two proxy questions rather than all twelve questions.

We used Item Response Theory Factor Analysis (IRT-FA) to identify which questions were most discriminating in each cluster. This was done with the `irt.fa` function in R's `psych` library.

Based on the IRT-FA results we chose Q7.4 and Q7.8 to be our proxies for the two clusters. They were chosen because they were the most centered in figures 2 and 3 with an amplitude closest to 0.5. These questions were:

7.4 "I develop spreadsheets, databases or other systems to analyse data"

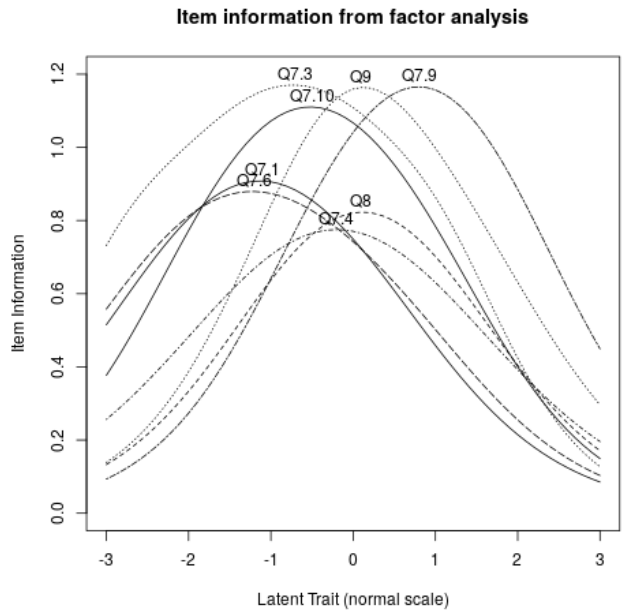

Figure 3: IRTFA2

7.8 "I get other people to set up the computing technology in my house, e.g. internet, home entertainment, PC/laptop/tablet, printer"

4.5 Predictive modelling

With our two proxy measurements of adult computing behaviour identified, we then performed two ordinal logistic regressions using R's `polr` function. Both regressions used all 43 questions about childhood activities as predictive variables. Using Bonferroni correction, we used an alpha value of 0.0005 to reduce the likelihood of false positives in our models. This does however mean that our model is very conservative and has a greater chance of missing some associations between childhood activity and adult computing proficiency.

To test goodness of fit for the models, we used the Hosmer-Lemeshow goodness of fit test (`logitgof` in R).

The goodness of fit for our model for cluster 1 was statistically significant ($p = 0.0000003$). But the goodness of fit for the model of cluster 2 just failed to meet the threshold for statistical significance ($p = 0.06$). As a result we failed to reject our null hypothesis that there was no association between childhood activities and cluster 2.

The cluster 1 model had a McFadden pseudo R^2 of 0.2. For cluster 1, we found two statistically significant predictors. These were questions 3.15 and 3.19:

3.15 Reading on my own
3.19 Playing with construction toys without moving parts (e.g. lego, duplo)

The two predictors had an log(odds ratio) of -2.11 and -2.06 respectively. Per [19] these are both considered to be medium effect sizes. The full regression table is available in section 8.

5 SPECULATIVE EXPLORATION

Independently of the analysis described above, and inspired by magazine questionnaires, we attempted to combine the "current use" answers from Question 7 into a single number that reflects the overall computer use over a range of tasks (affinity for computational thinking). The questions are given in Table 8. To do so, we converted the scales of each question to (-2, -1, 0, 1, 2) corresponding to Not at all, A little bit, Somewhat, Quite a bit, and A lot for the positively-stated questions. The statistic, which we denote AHMCTA, for Ad Hoc Measure of Computational Thinking Affinity, is a linear combination of these 10 values, with weights of 1 for each of the positive-viewpoint questions and -1 for each of the negative-viewpoint questions. For each individual, this statistic is a number over the range [-20,20], with higher numbers corresponding to greater sophistication or comfort in use of computing technology. As there were very few missing values in Question 7, AHMCTA had only three missing values (corresponding to individuals who did not answer all of the parts of Question 7). The analyses done with AHMCTA are not meant to be statistically rigorous in terms of meeting assumptions, rather they are explorations that might suggest areas of further examination.

The overall distribution of AHMCTA is shown in Figure 4. It is relatively smooth, and skewed somewhat to the right.

Does the AHMCTA statistic plausibly reflect the level of technology use or comfort of a user? Arguably so: those respondents who give the "maximum" answer on all of these questions report a high amount of technology use and comfort across a large range of activities; similarly those at the other extreme are less sophisticated (or comfortable) technology users. On the margin, if a person has a lower value for one of these factors than another, they are lower on this overall scale.

Table 8: Parts of Question 7

	question	viewpoint
1	I understand advanced features of a word processor such as styles, accurate positioning of images, and automatic section numbering.	positive
2	I avoid using computing technology wherever possible, preferring e.g. pencil and paper	negative
3	I fix technology issues that I experience	positive
4	I develop spreadsheets, databases or other systems to analyze data	positive
5	When the computer doesn't do what I expect I immediately ask someone for help.	negative
6	I can pick up new apps and computing technology without really thinking - I work out how to use it just by trying things out	positive
7	I use computing technology by writing down or memorising the steps I have to follow - if something goes wrong, I'm a bit lost	negative
8	I get other people to set up the computing technology in my house, e.g. internet, home entertainment, PC/laptop/tablet, printer	negative
9	I set up rules or macros in programs to automate common actions	positive
10	People at work or in my home life come to me with their technology problems	positive

Inherent in this measure (and its use as a dependent variable in prediction) are some assumptions that we have not tested:

(1) Assigning the qualitative scores to numbers assumes the differences between adjacent categories are equal: the difference between "Not at all" and "A little bit" is the same as the difference between "A little bit" and "Somewhat", and so forth.
(2) We weight the questions equally; if some questions measure the same thing, then that thing is given more weight.

Neither of these problems is insurmountable, they suggest that a more sophisticated combination function might lead to a better statistic.

Given this measure, we examined how well it could be predicted by answers from other questions, specifically the answers from Questions 3 and 5, each of which is answered on a 5 point scale measuring the amount of enjoyment or level of agreement respectively.

5.1 Question 3 regressions

As a first cut, we attempted to build a linear model predicting AHMCTA from Question 3 answers. Unlike Question 7, "N/A" is a possible answer for people having no experience. The number of N/A answers in Question 3 ranges from 12 to 728 out of 1329 responses; the median value is 77, the mean is 134.86. Six of the questions have over 10% N/A: Q3.21 (55%), Q3.10 (39%), Q3.13 (14%),

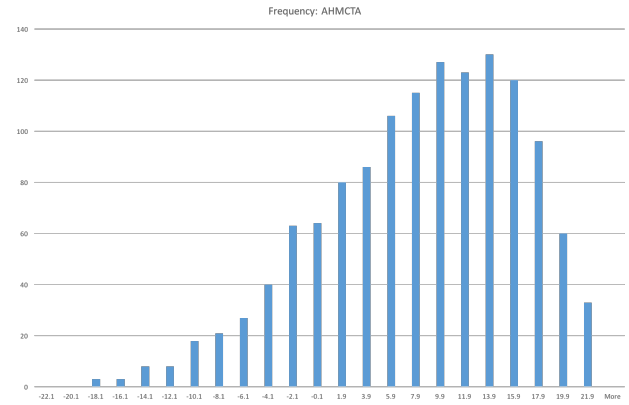

Figure 4: Histogram of AHMCTA over all the survey responses.

Q3.20 (13%), Q3.18 (12%), and Q3.17 (10%). We chose potential independent variables for the model by taking the questions that correlated best with AHMCTA (ignoring the questions with N/A rates over 10%). We built a linear predictor for AHMCTA using multiple linear regression using these variables, then eliminated the least significant variable in the model one at a time until each variable remaining had a p-value < 0.005.

The result of this sequence of operations on Question 3 was a predictor using Questions Q3.14, Q3.15, Q3.16, and Q3.19; Q3.15 (Reading on my own) and Q3.19 (Playing with construction toys without moving parts) had positive coefficients, while Q3.14 (Fantasy play - puppets, Barbie dolls, Action men) and Q3.16 (Being read to by others) had negative coefficients. This regression had an R^2 value of 0.0837, and was significant with $p < 2.2e - 16$. While the predictive value is not particularly high, the variables that contribute positively, Q3.15 and Q3.19, are the same ones found as significant with the ordinal logistic regression discussed earlier.

5.2 Grouping question 3 activities

An alternative to using the individual questions in Question 3 is grouping them into variables based on what sort of developmental activities they correspond to. We use the types of play presented in Table 1 to group the Question 3 activities: Physical play (3.11, 3.20, 3.21), Play with objects (3.4, 3.7, 3.8, 3.13, 3.19), Symbolic play (3.9, 3.15, 3.16, 3.18), and Games with rules (3.2, 3.3, 3.5, 3.6, 3.10, 3.12, 3.17). Rather than sum the individual questions in each group, we used the mean value; with the large number of N/A values, the number of observations where a group had no N/A answers was small).

Using these groups as independent variables, we went through the same sequence of regressions as with the individual questions above, starting with all four variables. The resulting model had two predictive factors, Physical play with a positive coefficient, and Symbolic play with a negative coefficient. It had an R^2 value of 0.0465, and was significant with $p < 4.4e - 14$. While significant, the predictive value is quite low.

5.3 Question 5 regressions

As we did with Question 3, we attempted to build a predictive model of AHMCTA from the Question 5 answers. With one exception, there were relatively few N/A answers (Q5.6, "People around me were into programming", had 442 N/A answers out of 1329), so all but that one could be considered, but the correlations between Questions and AHMCTA were relatively low. Building successive models as above, we ended with a model based on a single question (Q5.9, Trying new things was encouraged) with a positive coefficient, with an R^2 value of 0.0095, significant with $p = .0004$.

5.4 Summarizing Explorations

Using AHMCTA as a measure did not lead to particularly strong predictive models. The model from the Question 3 regressions included the two questions found to be significant in the logistic regressions, plus two more, suggesting that those two questions might be worth further investigation. Grouping the Question 3 activities by type of play did not improve things, but assigning the questions to the play types was inexact, and using the average may have muddied the data. The Question 5 regression is significant, but with a very low R^2; these particular statements do not seem to be very predictive of AHMCTA.

6 DISCUSSION

In the strict statistical analysis of our outcome measure of computing skills, we have identified two clear traits, which we deem to be related to *confidence* in the use of computing technology, and the *complexity* of that use. In looking for connections between these two traits and the participants' childhood activities, we find: no statistically significant relationship with the complexity trait (it was marginally outside the significant range); and a highly-sigificant relationship of medium effect-size with the confidence trait, based on two childhood activity items – reading on one's own, and playing with construction toys with no moving parts (bricks, simple lego, etc.) The exploratory analysis further confirms this relationship.

These results, and the study generally, raise many questions for discussion:

- Why do we deem the two traits to relate to confidence and complexity with technology use?
- How do reading on one's own and construction toys with no moving parts relate to confidence with using computing technology?
- Why has the study found no relationship between the complexity of problems solved using computing technology and childhood activities?
- Do age or gender of respondents have an effect?
- Can recall of childhood experiences be trusted, and other potential threats to validity?

6.1 The Two Traits

The statistical analysis has identified two clear traits in the outcome measure of computing skills. What do we think these traits represent, and why? One trait consists of the following questionnaire items:

- 7.2 I avoid using computing technology wherever possible, preferring e.g. pencil and paper
- 7.5 When the computer doesn't do what I expect I immediately ask someone for help.
- 7.7 I use computing technology by writing down or memorising the steps I have to follow - if something goes wrong, I'm a bit lost
- 7.8 I get other people to set up the computing technology in my house, e.g. internet, home entertainment, PC/laptop/tablet, printer.

These questions mostly consider the level of engagement with computing technology, not the actual activities performed on the technology. For example, "I avoid using computing technology...", "I immediately ask someone for help", "I get other people to...". They seem to concern an internal or visceral response to computing technology which we interpret as an overall measure of *confidence* in using it. Question 7.7 consists of two parts. The first, "I use computing technology by writing down or memorising the steps I have to follow", relates to actually working with the technology and gives an indication of the depth of understanding of the technology, but the second part, "if something goes wrong, I'm a bit lost", is back to more of a confidence statement.

The second trait consists of the question items (see the Appendix for the full versions of questions 8 and 9):

- 7.1 I understand advanced features of a word processor such as styles, accurate positioning of images, and automatic section numbering.
- 7.3 I fix technology issues that I experience
- 7.4 I develop spreadsheets, databases or other systems to analyze data
- 7.6 I can pick up new apps and computing technology without really thinking - I work out how to use it just by trying things out
- 7.9 I set up rules or macros in programs to automate common actions
- 7.10 People at work or in my home life come to me with their technology problems
- 8 What of the following best describes your level of programming?
- 9 If you've ever been in a programming class, how easy did you find it compared to most other individuals in the class learning at the same time?

Many of these are closely connected to what the respondent is *able to do* with the computing technology: "I fix...", "I develop spreadsheets... to analyse data", "I work out how to use it", "I set up rules or macros", and descriptions of programming ability and apparent level of ease of learning to program. Question 7.10, about how folk come to the respondent to get their technology issues fixed, speaks to an underlying ability to both understand and work with technology. Question 7.6 is the opposite of the first part of Question 7.7, discussed earlier, getting at the depth of understanding of computing technology generally: responding positively to 7.6 indicates that an underlying model of how the technology works is being developed by the respondent as they play around with the technology; whereas the activity described in 7.7 indicates a very shallow understanding of technology, with use only possible via tightly following a set of precise instructions provided by someone else.

We see this trait, therefore, as a measure of the ability level of the respondent in using computing technology, or the *complexity* of their use.

We note in passing that all the confidence trait items are expressed negatively, while all the complexity items are positive or neutral. However, given the nature of the principal component analysis that determined the two traits, it is not clear that this unnoticed distinction between the questions could be having any effect.

6.2 Relating Reading and Construction Toys to Confidence

In order to understand more about how reading on one's own and playing with construction toys might contribute to confidence with technology, we will explore some of the pertinent characteristics of using computing technology.

Software applications tend to be low on explicit instructions, requiring the user to infer from what can be seen on the screen what it is that they need to do to operate the technology. With careful design, such as in a well-made smartphone app operating in a narrow context, the software guides the user effectively on the steps they should follow. But more complex software tends to be more of a general purpose tool, with a complex underlying model that is well-hidden. For example, a modern word processor has all the power once commanded by a professional type-setter, with a hugely complex model of a document embedded within it. This model cannot be displayed on screen in full, and so a successful user develops an increasingly complete internalised model over time as he or she uses the software, where each use is a kind of learning experiment resulting in a little more knowledge. Investigation skills are strongly in play during this process, and also when the software inevitably doesn't do what was expected. Furthermore, the very nature of a general purpose tool demands that the user work out how to complete their tasks using the tool at hand - a problem solving activity. Such a focus on problem solving is reported in Attard, Mountain and Romano in their study of computing use [5]. Computing technology changes all the time, of course, and so being able to adapt to change, to be able to transfer knowledge and skills from one context to another, is important. In the face of regular change, investigation, and problem solving, then a level of resilience is also required. Hence, confident use of computing technology is likely to require:

- confidence and resilience in the face of uncertainty and change
- ability to develop and hold mental models associated with the technology
- investigatory and problem solving skills

We have found no direct studies that link reading ability to confident use of computing technology, and so we develop an argument based on the literature. Terminologically, reading *instruction* is distinguished from reading *for pleasure* [21], with significant benefits accruing from the latter, including particularly the development of willpower and self-direction. We judge our activity of *Reading on your own* to be the same as reading for pleasure. A longitudinal

study [71] followed 6,000 children from similar socio economic backgrounds to establish the cognitive benefits from reading for pleasure. Findings show that those who read books often at an early age gain higher test scores at the age of 16. While researchers were not surprised with the impact on children's literacy attainment, the link to maths scores was not expected. Their justification for this is that a strong reading ability enables children to absorb and understand new information, which must be structured in some kind of internal model, e.g. of people, relationships, places, events, and which therefore impacts positively on their attainment in all subjects. This could relate to resilience and the ability to develop mental models when using computing technology as noted above.

An alternative view on the reading attribute is that, in order to enjoy reading on one's own, a high level of *executive function* (EF) is required [66]. EF is a loosely-defined term capturing a number of higher-order cognitive skills necessary for independent goal-directed behaviour. Such skills include holding and manipulating information in working memory, planning and sequencing tasks, ascertaining the "big picture" from a complicated set of details, thinking "outside the box", and self-control to, for example, meet unanticipated challenges. Hence reading on one's own may be an indicator of a deeper set of skills that are related to confidence with computing technology, since skills associated with executive function relate well to holding mental models and problem solving identified above.

As for the other factor, playing with construction toys with no moving parts, we refer back to Whitebread's findings [22] given in Table 1, which outlines a wide range of benefits for children playing with objects, incorporating construction toys with no moving parts, in the early years. Examples are perseverance, positive attitude and self-regulation towards challenge, investigative behaviour, thinking, reasoning and problem-solving. These all relate to the items identified above, associated with confident use of technology. Clavio and Fajardo [20] explain why playing with blocks develops problem solving skills. They label blocks and building sets as *divergent* materials. This means that they lead to multiple uses and the open ended result encourages open ended thinking. They suggest that problem solving requires memory, reasoning and metacognition and that playing with blocks can help develop these.

Studies by Jirout and Newcombe [36] and Oostermeijer, Boonen and Jolles [53] and Richardson, Hunt and Richardson [59] link playing with blocks (construction toys with no moving parts) and the development of spatial awareness. The relevance of this links to Sorby's work on the importance of particular kinds of spatial skills, such as mental rotation, on achievement in a variety of different STEM disciplines [69]. The analysis of factors involved in confidently using computing technology above points to the need for identifying hidden models in software, related to spatial awareness, and the work of Sorby has clearly identified the value of spatial awareness skills for science and technology subjects.

6.3 Why no relationship with Complexity?

Our original hypothesis suggested that childhood activities outside formal education are a predictor of computational thinking skills demonstrated later in life. Using the complexity trait in our outcome measure, we have found no statistically significant predictor, with the model developed having a significance value of only 0.06, and the exploratory analysis found nothing of real significance here either.

The highly significant model derived for the confidence trait suggests that asking survey respondents about early childhood experiences and relating these to later behaviours is a valid method. We discuss this in more detail later in this section. So the general methodology is workable. Intuitively, however, we were surprised that no connection emerged. It may of course be the case that the experiment is correct in every respect and there is no connection; having considered our survey design carefully, however, and as a prompt for further analysis and research, we raise the following two issues with the design.

6.3.1 Unexpected responses on skills. Even though we trialled the questionnaire prior to distributing it widely, we found conflicting responses in respondents' answers that would reduce the significant of the link we were looking for.

Building questionnaires for a global audience is not easy. In our case, the questions that asked about favorite childhood toys, games and activities and their key features, as well as how supportive the environment in which the respondents grew up in, were easy enough. Participants were provided an extensive list of options, and they could also type in additional data points that they thought were missing from the questionnaire. When asking about the current use of computers and applications, questions were constructed to capture as wide a relevant experience as possible from the participants: there are questions about studying programming and programming proficiency, as well as questions to determine individuals who are proficient with the advanced use of technologies of interest.

The questions were built by the researchers with an extensive domain knowledge on both teaching computing and computing in general. It is possible that an individual who knows the domain understands the questions as intended, but it is also possible that some participants do not understand the questions as intended. For example, some participants may have not understood the question on developing spreadsheets, databases and systems to analyze data in the way the researchers intended – some may, for example, consider digitizing information into a spreadsheet as developing a spreadsheet.

Furthermore, we noticed a potential confound between items in question 7 that asked about advanced use of common office-based software such as spreadsheets and databases, and question 8 which asked about programming proficiency. While we wanted to capture both aspects individually, they are linked in the context of this study: in the broadest sense, computer programmers are individuals who develop applications written in programming languages - these could be general purpose programming systems, but also for example, spreadsheets, databases and web systems. As identified in [40], there is a large category of professionals who write spreadsheets, databases and websites to support their job or hobbies. Indeed, we note in our data that these professionals do not necessarily consider themselves to be programmers: in our analysis we have observed that a large proportion (44%) of participants who indicated that they write macros, databases, spreadsheets or other

systems to analyze data (see Q 7.4 and 7.9) did not report themselves as programmers (Q8), in that they reported that they never took a programming course or they failed it, rather than indicating that they were successfully programming. We believe that these participants are in fact programmers, but we note that we should have treated them as related/complementary variables, rather than independent variables, as we have done. The confound in the data will weaken/confuse the complexity trait with a knock-on effect on its relationship with childhood activities. In order to address this aspect we aim in future work to include a definition of what a "programmer" means, or alternatively to treat the answers to these two questions in a different manner.

6.3.2 Lateness of the CT skills measure. The confound noted in the previous paragraph sheds light on a larger issue with the study as a whole.

The study has been driven by the widely-acknowledged anecdotal observation that some students succeed in introductory CS courses at university level and some don't, even when both groups of students have no prior experience in CS. Given that the evidence is not strong for pre-requisite school subjects or innate ability contributing to success, we conjectured that success may come from childhood activities outside formal education. We operationalized this as a search for common childhood activities evident in respondents who were demonstrating computational thinking skills later in life. We used the measure "demonstrating CT skills later in life" as a relaxation of a stricter measure "passed/failed an introductory CS course" (upon which the motivating anecdotal observations is based) so that we could poll individuals from a wider range of backgrounds (e.g. non-university, non-CS), and because a self-report is easier to manage than needing to connect responses to course grades.

However, the study design captures an unintended viewpoint on learning computing skills: that childhood experiences are the dominating influence in being able to pick up these skills. It is one thing to conjecture that an 18-year-old university student's success on a CS course might be strongly attributable to key activities undertaken just 10 years ago, where other academic activities have already been largely discounted; but it is quite another to claim that these early activities are the key predictors of success with technology for a 30, or 50, or 60 year old. Considering Carol Dweck's growth mindset [25], those who wish to learn about computing technology, have the necessary resources, and the belief and ability to persevere, will be able to teach themselves; as part of their jobs, they may also attend a range of training courses directly targeting technology skills. Just because they may have failed our very particular kind of computing course at university does not mean that they are unable, with appropriate motivation and opportunity, to pick up the skills later in life.

A final observation about the questions on current skills, and why there may be no link between them and childhood activities. Whereas every respondent will have sufficient experience of computing technology to answer the four confidence questions accurately, some of the questions in the skill-based trait are not necessarily measuring a skill level. For example, a respondent may answer that they don't build spreadsheets and databases to analyse data. How is this response to be interpreted? Is it that they choose not to, or they are unable to, build such artefacts? They may have the necessary skill, but it's not something they ever need to do.

To understand more clearly the observations of students' varying success in our courses, and whether these are related to early childhood activities, we will need to run a more tightly controlled study.

6.4 "Why didn't you look at ...?'

The goal of our work has been to take the first steps towards understanding whether and how early developmental activities affects computing proficiency. This exploration was conducted through an analysis of responses to a questionnaire that asked for memories of joyful childhood activities and linked them with computing related activities and beliefs in current life. The responses to our study ($n > 1300$) represent a cross-section of members in various societies and countries. To paint the overall picture and to open up avenues for future research, we chose to focus on the analysis of the data as a whole, without looking for sub-populations in the data.

We do not know the extent of selection bias in our data, but one could argue that the majority of the responses may come from so-called upper, upper-middle and middle classes. Similarly, as we focused on the analysis on the responses as a whole, it is possible that the results could have been different had we studied the responses of each data provider individually.

We ran a sense check on the data for age and gender, as the most obvious sub-populations where responses might differ. Averages for each question were calculated for each age and gender category. Averages were then compared, for each question, to see if any particular age group or gender responded clearly differently to the others. We did the same for the number of respondents who indicated for each question whether they had experienced the corresponding activity, given that there was a *not-applicable* response category. They did not appear to differ significantly. With respect to age, the only childhood questions that showed some difference between the age groups were the questions related to the numbers involved in, and their enjoyment or use of, design and tech toys and activities involving programming. There were steadily fewer respondents in, and they were steadily less positive about, these categories as age increases. This can be explained by older people in general having either limited or no access to such toys and activities; where there was only limited access, this may not have been enough to have formed a significant impression. When looking at gender, there is also some difference in enjoying design and tech toys and toys with a programming element, in favour of males. The questions that were in the predictive model (Q3.15, Q3.19) were rather evenly distributed across the groups.

The country of the responder could possibly influence weights in the data and thus influence the overall predictive model, although we didn't explore this further. For example, running the analysis on data only from India may produce different results as per the UNESCO Institute of Statistics report from 2015 approximately 72.1% of all people aged 15 and above can read and write, whilst the same number for e.g. Romania is 98.8%. Differences in gross domestic product (GDP) per capita, which can be observed between some of the countries from which data was entered to our survey, could also influence some of the results.

These observations all suggest ways in which the questionnaire could be improved upon in future studies.

6.5 Threats to validity

Clear concerns in a study of this kind relate to the self-reporting of early memories and to what kind of memories are stored from early childhood.

6.5.1 Self-reporting and memory recall reliability. The following key issues were elicited from the literature:

- Telescoping of memories [6], where the precise date for events is misplaced forwards or backwards in time. Respondents to our survey did comment that they were unsure whether the recall was from age 8 or earlier, or in truth from some time afterwards. This is not a major issue for us as we are concerned in this initial study only with the broad period of childhood up to around 8 - a year or two later doesn't matter.
- Fadnes [43] identifies a number of aspects. Recall period is important as a longer period reduces recall but makes it more representative. This is fine for this study as we are looking for a representative picture over time rather than an activity enjoyed once only.
- Some items are selectively recalled more easily than others. The survey had free response items subject to this issue, but also many rating questions on fixed items. Furthermore, Howes [35], focusing on early memories, indicates that positive memories, such as we are looking for, are just as well remembered as negative or traumatic experiences, counter to prevailing beliefs.
- Suggestibility and social desirability. This is where the questionnaire leads the respondent, or the latter tries to "look good" in the eyes of the experimenter. We avoid biasing the respondent in this way by minimally explaining the context or purpose of the questionnaire in advance.
- Bias. Unintentional bias introduced via the questions is always a concern. Where the questionnaire was to be delivered in other countries or cultures compared to its originating country, Scotland, the local experimenter ensured that it was adjusted for the local context. An example was ensuring that named games or activities would be commonly known to local respondents. Age and gender bias have been discussed in the previous section.

6.5.2 Accuracy of childhood memories. This concerns the formation and reliability of childhood memories. Work by Bauer and Larkina [7] suggests there are different memory mechanisms up to age 3 1/2, then up to 9, up to 18, and finally to adult. Most of the memories in the earliest phase are lost. Wells [78] examines the earliest memories in which one could have confidence. Both positive and negative memories were found in the age range 3 1/2 up to 10, closely related to our target range; the memories were rarely highly detailed, focusing on activities, locations and the people present, but this is enough for our needs. Some respondents in our study were concerned at the validity of their memories. While Wells notes that there will always be a degree of inference and construction, a study by Howes [35] on very early memories showed that the majority of memories proved to be accurate, irrespective of whether they were positive or negative experiences being recalled. Asking about the most enjoyable early activities is therefore reasonable. Furthermore, participants in [14] noted particularly vivid memories of times when they were playing alone, which is relevant to some of the activities in the questionnaire.

7 CONCLUSIONS AND FUTURE WORK

Considering our initial research questions, we have met RQ1: Yes, there are childhood activities that predict aspects of computer use later in life. We have found two childhood activities that are linked with a measure of confidence in using computing technology: these are reading on one's own and playing with construction toys with no moving parts. And we have an answer to RQ2 as well, in that the two childhood activities identified have a medium effect size.

We have demonstrated the principle that recollection of childhood experiences can be correlated with current abilities with a high level of statistical significance.

As for the particular findings, confidence with computing technology is an essential first step on the journey to becoming a high-end user of technology, and the activities identified as being related to that confidence can be considered for greater emphasis in early years education. It is interesting to speculate whether playing with construction toys might develop executive function which in turn supports the ability to read on one's own - and that the increased executive function ultimately leads to a range of traits supporting confident use of technology.

We did not find a link from childhood activities to more advanced use of computing technology. But, as we have shown in the discussion, there are a number of methodological reasons that explain why this may be, principally to do with issues with the way the complexity questions were framed and then analysed. Further work could either re-run the questionnaire in a similar way, across the wider population as was done here, but with complexity questions that are unambiguous; or it could focus a related questionnaire specifically on the target age groups we're interested in, using in-class performances as the outcome measure.

8 ACKNOWLEDGEMENTS

We thank Francesco Maiorana and Shitanshu Mishra for their sterling efforts in collecting several hundred responses to our survey, and Shitanshu additionally for assisting our analyses in Bologna.

REFERENCES

[1] 2014. Re: Motivating Students. Was: Survey Results. (14 October 2014). SIGCSE mailing list, Item 013729. http://sigcse.org/sigcse/membership/mailing-lists. (Last accessed 26 August 2016).

[2] Alireza Ahadi and Raymond Lister. 2013. Geek Genes, Prior Knowledge, Stumbling Points and Learning Edge Momentum: Parts of the One Elephant?. In *Proceedings of the Ninth Annual International ACM Conference on International Computing Education Research (ICER '13)*. ACM, New York, NY, USA, 123–128. DOI : http://dx.doi.org/10.1145/2493394.2493416

[3] Efthimia Aivaloglou and Felienne Hermans. 2016. How Kids Code and How We Know: An Exploratory Study on the Scratch Repository. In *Proceedings of the 2016 ACM Conference on International Computing Education Research (ICER '16)*. ACM, New York, NY, USA, 53–61. DOI : http://dx.doi.org/10.1145/2960310.2960325

[4] anonymous. Recalling Memory in Survey Response. (????). http://www.activecampaign.com/blog/recalling-memory-in-survey-response/

[5] Conrad Attard, Gail Mountain, and Daniela Maria Romano. 2016. Problem solving, confidence and frustration when carrying out familiar tasks on non-familiar mobile devices. *Computers in Human Behavior* 61 (2016), 300–312.

[6] Oztas Ayhan. Memory recall errors in retrospective surveys: A reverse record check study (PDF Download Available). (????). https://www.researchgate.net/publication/226044747_Memory_recall_errors_in_retrospective_surveys_A_reverse_record_check_study DOI: http://dx.doi.org/10.1007/s11135-005-2643-7.

[7] Patricia J. Bauer and Marina Larkina. 2014. The onset of childhood amnesia in childhood: A prospective investigation of the course and determinants of forgetting of early-life events. *Memory* 22, 8 (Nov. 2014), 907–924. DOI:http://dx.doi.org/10.1080/09658211.2013.854806

[8] Jens Bennedsen and Michael Caspersen. 2005. An investigation of potential success factors for an introductory model-driven programming course. In *ICER-05*. 155–164.

[9] Jens Bennedsen and Michael E. Caspersen. 2006. Abstraction ability as an indicator of success for learning object-oriented programming? *SIGCSE Bulletin* 38, 2 (2006), 39–43.

[10] Jens Bennedsen and Michael E. Caspersen. 2007. Failure Rates in Introductory Programming. *SIGCSE Bull.* 39, 2 (June 2007), 32–36. DOI:http://dx.doi.org/10.1145/1272848.1272879

[11] Jens Bennedsen and Michael E. Caspersen. 2008. Abstraction Ability As an Indicator of Success for Learning Computing Science?. In *ICER '08*. 15–26.

[12] Susan Bergin and Ronan Reilly. 2005. Programming: Factors That Influence Success. In *SIGCSE '05*. ACM, 411–415.

[13] Richard Bornat. 2014. Camels and humps: a retraction. (July 2014). Retrieved from http://www.eis.mdx.ac.uk/staffpages/r_bornat/papers/camel_hump_retraction.pdf (Last accessed August 17, 2016).

[14] Urie Bronfenbrenner. 1979. *The Ecology of Human Development: Experiments by Nature and Design*. Harvard University Press.

[15] Jerome Bruner. 1996. Folk Pedagogy. In *The Culture of Education*. Harvard University Press, Chapter 2, 44–65.

[16] Faye Marsha G Camahalan. 2014. Recalling childhood memories: An instructional approach using grounded theories. *Journal of Instructional Pedagogies Volume* (2014).

[17] Michael E. Caspersen, Kasper Dalgaard Larsen, and Jens Bennedsen. 2007. Mental Models and Programming Aptitude. In *ITiCSE '07*. ACM, 206–210.

[18] Polina Charters, Michael J. Lee, Andrew J. Ko, and Dastyni Loksa. 2014. Challenging Stereotypes and Changing Attitudes: The Effect of a Brief Programming Encounter on Adults' Attitudes Toward Programming. In *Proceedings of the 45th ACM Technical Symposium on Computer Science Education (SIGCSE '14)*. ACM, New York, NY, USA, 653–658. DOI:http://dx.doi.org/10.1145/2538862.2538938

[19] Henian Chen, Patricia Cohen, and Sophie Chen. 2010. How big is a big odds ratio? Interpreting the magnitudes of odds ratios in epidemiological studies. *Communications in Statistics, Simulation and Computation®* 39, 4 (2010), 860–864.

[20] Josefa Carina V Clavio and Amelia C Fajardo. 2010. Toys as Instructional Tools in Developing Problem-Solving Skills in Children. *Education Quarterly* 66, 1 (2010).

[21] Kathleen E Cox and John T Guthrie. 2001. Motivational and cognitive contributions to students' amount of reading. *Contemporary Educational Psychology* 26, 1 (2001), 116–131.

[22] Martina Kuvalja David Whitebread, Marisol Basilio and Mohini Verma. 2012. *The importance of play: A report on the value of children's play with a series of policy recommendations*. Technical Report. University of Cambridge.

[23] Saeed Dehnadi and Richard Bornat. 2006. The camel has two humps (working title). (2006). Retrieved from http://www.eis.mdx.ac.uk/research/PhDArea/saeed/paper1.pdf (Last accessed August 17, 2016).

[24] Brian Dorn, Allison Elliott Tew, and Mark Guzdial. 2007. Introductory Computing Construct Use in an End-User Programming Community. IEEE, 27–32. DOI:http://dx.doi.org/10.1109/VLHCC.2007.35

[25] Carol S. Dweck. 2008. Can Personality Be Changed? The Role of Beliefs in Personality and Change. *Current Directions in Psychological Science* 17, 6 (2008), 391–394.

[26] BG Embree and PC Whitehead. 1993. Validity and reliability of self-reported drinking behavior: dealing with the problem of response bias. In *Journal on the Study of Alcohol*, Vol. 54. 334–344.

[27] Howard Gardner. 1993. *Frames of Mind: The Theory of Multiple Intelligences*. Basic Books.

[28] Shuchi Grover and Satabdi Basu. 2017. Measuring Student Learning in Introductory Block-Based Programming: Examining Misconceptions of Loops, Variables, and Boolean Logic. In *Proceedings of the 2017 ACM SIGCSE Technical Symposium on Computer Science Education (SIGCSE '17)*. ACM, New York, NY, USA, 267–272. DOI:http://dx.doi.org/10.1145/3017680.3017723

[29] Mark Guzdial. 2009. Are some people suited for CS1, and others not? Computing Education Blog. (November 29 2009). https://computinged.wordpress.com/2009/11/29/are-some-people-suited-for-cs1-and-others-not/ (last accessed August 2, 2016).

[30] Norbert Győrbíró, Henry Larkin, and Michael Cohen. 2010. Long-term memory retention and recall of collected personal memories. In *ACM SIGGRAPH 2010 Posters*. ACM, 90.

[31] Jochen Hardt and Michael Rutter. 2004. Validity of adult retrospective reports of adverse childhood experiences: review of the evidence. *Journal of Child Psychology and Psychiatry* 45, 2 (Feb. 2004), 260–273. DOI:http://dx.doi.org/10.1111/j.1469-7610.2004.00218.x

[32] Felienne Hermans, Martin Pinzger, and Arie van Deursen. 2009. Domain-Specific Languages in Practice: A User Study on the Success Factors. In *Model Driven Engineering Languages and Systems (Lecture Notes in Computer Science)*. Springer, Berlin, Heidelberg, 423–437. DOI:http://dx.doi.org/10.1007/978-3-642-04425-0_33

[33] Lars Josef Höök and Anna Eckerdal. 2015. On the Bimodality in an Introductory Programming Course: An Analysis of Student Performance Factors. In *LATICE '15*. 79–86.

[34] Matthew Horton and Janet C. Read. 2012. Parents and Children Having and Using Technology: What Should We Ask?. In *Proceedings of the 11th International Conference on Interaction Design and Children (IDC '12)*. ACM, New York, NY, USA, 252–255. DOI:http://dx.doi.org/10.1145/2307096.2307138

[35] M. Howes, M. Siegel, and F. Brown. 1993. Early childhood memories: Accuracy and affect. *Cognition* 47, 2 (May 1993), 95–119. DOI:http://dx.doi.org/10.1016/0010-0277(93)90001-C

[36] Jamie J. Jirout and Nora S. Newcombe. 2015. Building Blocks for Developing Spatial Skills. *Psychological Science* 26, 3 (2015), 302–310. DOI:http://dx.doi.org/10.1177/0956797614563338 PMID: 25626442.

[37] Ros Johnson. 2014. Adult Recall of Childhood Memories: The Research. (May 2014). http://www.mindingtherapy.com/childhood-memories/

[38] Geetha Kanaparan, Rowena Cullen, and David Mason. 2013. Self-Efficacy and Engagement as Predictors of Student Programming Performance. In *PACIS 2013 Proc.* http://aisel.aisnet.org/pacis2013/282.

[39] Päivi Kinnunen, Robert McCartney, Laurie Murphy, and Lynda Thomas. 2007. Through the eyes of instructors: a phenomenographic investigation of student success. In *ICER '07*. ACM, 61–72.

[40] Andrew J. Ko, Robin Abraham, Laura Beckwith, Alan Blackwell, Margaret Burnett, Martin Erwig, Chris Scaffidi, Joseph Lawrance, Henry Lieberman, Brad Myers, Mary Beth Rosson, Gregg Rothermel, Mary Shaw, and Susan Wiedenbeck. 2011. The State of the Art in End-user Software Engineering. *ACM Comput. Surv.* 43, 3 (April 2011), 21:1–21:44. DOI:http://dx.doi.org/10.1145/1922649.1922658

[41] Jeff Kramer. 2007. Is abstraction the key to computing? *Commun. ACM* 50, 4 (2007), 36–42.

[42] Annette Lareau. 1987. Social class differences in family-school relationships: The importance of cultural capital. *Sociology of education* (1987), 73–85.

[43] Adam Taube Lars T. Fadnes and Thorkild Tylleskr. 2008. How to identify information bias due to self-reporting in epidemiological research. *The Internet Journal of Epidemiology* 7, 2 (Dec. 2008). http://ispub.com/IJE/7/2/9194

[44] R. R. Leeper and J. L. Silver. 1982. Predicting Success in a First Programming Course. In *Proceedings of the Thirteenth SIGCSE Technical Symposium on Computer Science Education (SIGCSE '82)*. ACM, New York, NY, USA, 147–150. DOI:http://dx.doi.org/10.1145/800066.801357

[45] Clayton Lewis. 2007. Attitudes and Beliefs About Computer Science Among Students and Faculty. *SIGCSE Bull.* 39, 2 (June 2007), 37–41.

[46] Melissa E. Libertus, Lisa Feigenson, and Justin Halberda. 2011. Preschool acuity of the approximate number system correlates with school math ability. *Developmental Science* 14, 6 (2011), 1292–1300.

[47] G Mazzoni, A Scoboria, and L Harvey. 2010. Nonbelieved Memories. In *Psychological Science (9)*, Vol. 21. 1334–1340.

[48] Robert McCartney, Jonas Boustedt, Anna Eckerdal, Kate Sanders, and Carol Zander. 2017. Folk Pedagogy and the Geek Gene: Geekiness Quotient *(SIGCSE '17)*. ACM, New York, NY, USA, 405–410. DOI:http://dx.doi.org/10.1145/3017680.3017746

[49] Susan H. McLeod. 1995. Pygmalion or Golem? Teacher Affect and Efficacy. *College Composition and Commun.* 46, 3 (1995), 369–386. StableURL:http://www.jstor.org/stable/358711, Last accessed: 26-08-2016.

[50] Orni Meerbaum-Salant, Michal Armoni, and Mordechai Ben-Ari. 2011. Habits of Programming in Scratch. In *Proceedings of the 16th Annual Joint Conference on Innovation and Technology in Computer Science Education (ITiCSE '11)*. ACM, New York, NY, USA, 168–172. DOI:http://dx.doi.org/10.1145/1999747.1999796

[51] Janet R Moyles. 1989. *Just playing?: The role and status of play in early childhood education*. Open University.

[52] Andreas Mller, Matthias Kranz, Barbara Schmid, Luis Roalter, and Stefan Diewald. 2013. Investigating Self-reporting Behavior in Long-term Studies. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, New York, NY, USA, 2931–2940. DOI:http://dx.doi.org/10.1145/2470654.2481406

[53] Meike Oostermeijer, Anton JH Boonen, and Jelle Jolles. 2014. The relation between childrenfis constructive play activities, spatial ability, and mathematical word problem-solving performance: a mediation analysis in sixth-grade students. *Frontiers in psychology* 5 (2014).

[54] Elizabeth Patitsas, Jesse Berlin, Michelle Craig, and Steve Easterbrook. 2016. Evidence that Computer Science Grades are not Bimodal. In *ICER '16*. 8. (To appear).

[55] Jean Piaget. 1966. The Origins of Intelligence in Children. *Psychology in the Schools* 3, 2 (1966), 189–189. DOI:http://dx.doi.org/10.1002/1520-6807(196604)3:2⟨189::AID-PITS2310030223⟩3.0.CO;2-W

[56] Jean Piaget. 2013. *Play, dreams and imitation in childhood*. Vol. 25. Routledge.

[57] PM Podsakoff, SB MacKenzie, and JY Lee. 2003. Common method biases in behavioral research: A critical review of the literature and recommended remedies. In *Journal of Applied Psychology*, Vol. 88. 879–903.

[58] K. Raphael. 1987. Recall bias: a proposal for assessment and control. In *International Journal of Epidemiology*, Vol. 7. 167–170.

[59] Miles Richardson, Thomas E Hunt, and Cassandra Richardson. 2014. Children's Construction Task Performance and Spatial Ability: Controlling Task Complexity and Predicting Mathematics Performance. *Perceptual and motor skills* 119, 3 (2014), 741–757.

[60] Anthony Robins. 2010. Learning edge momentum: a new account of outcomes in CS1. *Computer Science Education* 20, 1 (2010), 37–71.

[61] Nathan Rountree, Janet Rountree, and Anthony Robins. 2002. Predictors of Success and Failure in a CS1 Course. *SIGCSE Bull.* 34, 4 (Dec. 2002), 121–124. DOI:http://dx.doi.org/10.1145/820127.820182

[62] Nathan Rountree, Janet Rountree, Anthony Robins, and Robert Hannah. 2004. Interacting factors that predict success and failure in a CS1 course. *SIGCSE Bull.* 36, 4 (2004), 101–104.

[63] S Salvucci, E Walter, V Conley, S Fink, and M Saba. 1997. Measurement error studies at the National Center for Education Statistics. U.S. Department of Education, Washington D.C.

[64] DL Schacter. 1999. The seven sins of memory: Insights from psychology and cognitive neuroscience. In *American Psychology*, Vol. 54. 182–203.

[65] N. Schwarz. 1999. How the questions shape the answers. In *American Psychologist*, Vol. 54. 93.

[66] Heather Whitney Sesma, E Mark Mahone, Terry Levine, Sarah H Eason, and Laurie E Cutting. 2009. The contribution of executive skills to reading comprehension. *Child Neuropsychology* 15, 3 (2009), 232–246.

[67] Simon, Sally Fincher, Anthony Robins, Bob Baker, Ilona Box, Quintin Cutts, Michael de Raadt, Patricia Haden, John Hamer, Margaret Hamilton, Raymond Lister, Marian Petre, Ken Sutton, Denise Tolhurst, and Jodi Tutty. 2006. Predictors of success in a first programming course. In *ACE2006 (CRPIT)*, Vol. 52. Hobart, Australia, 189–196.

[68] Peter K Smith and Anthony Pellegrini. 2008. Learning through play. *Encyclopedia on early childhood development* (2008), 1–6.

[69] Sheryl A. Sorby. 2009. Educational Research in Developing 3-D Spatial Skills for Engineering Students. *International Journal of Science Education* 31, 3 (2009), 459–480. DOI:http://dx.doi.org/10.1080/09500690802595839

[70] Peter Sprague and Celia Schahczenski. 2002. Abstraction the key to CS1. *J. Comput. Small Coll.* 17, 3 (2002), 211–218.

[71] Alice Sullivan and Matt Brown. 2013. Social inequalities in cognitive scores at age 16: The role of reading. *CLS Working Papers* 2013, 13/10 (2013).

[72] Thea F Van de Mortel and others. 2008. Faking it: social desirability response bias in self-report research. *Australian Journal of Advanced Nursing, The* 25, 4 (2008), 40.

[73] Marleen van Ooijen, Ulla I Ivens, Christoffer Johansen, and Torsten Skov. 1997. Comparison of a self-administered questionnaire and a telephone interview of 146 Danish waste collectors. *American journal of industrial medicine* 31, 5 (1997), 653–658.

[74] Jr. Ventura, P.R. 2005. Identifying predictors of success for an objects-first CS1. *Computer Science Education* 15, 3 (2005), 223–243.

[75] Lev S. Vygotsky. 1967. Play and Its Role in the Mental Development of the Child. *Soviet Psychology* 5, 3 (1967), 6–18. DOI:http://dx.doi.org/10.2753/RPO1061-040505036

[76] Christopher Watson and Frederick W.B. Li. 2014. Failure Rates in Introductory Programming Revisited. In *Proceedings of the 2014 Conference on Innovation & Technology in Computer Science Education (ITiCSE '14)*. ACM, New York, NY, USA, 39–44. DOI:http://dx.doi.org/10.1145/2591708.2591749

[77] Christopher Watson, Frederick W.B. Li, and Jamie L. Godwin. 2014. No Tests Required: Comparing Traditional and Dynamic Predictors of Programming Success. In *SIGCSE '14*. 469–474.

[78] Christine Wells, Catriona M Morrison, and Martin A Conway. 2014. Adult recollections of childhood memories: What details can be recalled? (2014).

[79] Brenda Cantwell Wilson and Sharon Shrock. 2001. Contributing to Success in an Introductory Computer Science Course: A Study of Twelve Factors. In *SIGCSE '01*. 184–188. DOI:http://dx.doi.org/10.1145/364447.364581

[80] Greg Wilson, D. A. Aruliah, C. Titus Brown, Neil P. Chue Hong, Matt Davis, Richard T. Guy, Steven H. D. Haddock, Kathryn D. Huff, Ian M. Mitchell, Mark D. Plumbley, Ben Waugh, Ethan P. White, and Paul Wilson. 2014. Best Practices for Scientific Computing. *PLOS Biology* 12, 1 (Jan. 2014), e1001745. DOI:http://dx.doi.org/10.1371/journal.pbio.1001745

[81] Chong-ho Yu. 2010. Reliability of self-report data. *Retrieved July* 11 (2010).

APPENDIX: THE SURVEY

Toys, games, and play

Thank you for opening up this questionnaire.

The questionnaire asks about the kinds of toys, games and activities that you preferred when you were young (age 8 or under) and the home environment at that time, as well as the kind of study or work in which you are now involved and some basic demographic information such as age and gender. There are 13 questions and we expect the questionnaire will typically take around 10-15 minutes to complete. We are collecting no information that would identify you, and so your responses will be completely anonymous.

We will explain why we are asking these questions in detail at the end of the questionnaire. We are not doing so at this point, because in pilot runs of the questionnaire, we have found that participants' answers are biased if we explain the study beforehand.

If, for any reason, you are not happy with participating in the study once we have explained it fully, you will be able to withdraw at once.

The questionnaire should only be completed by participants who are 18 years of age or over.

This study adheres to the British Psychological Society's ethical guidelines, and has been approved by the College of Science and Engineering's ethics committee at The University of Glasgow (ref: 300160139). The questionnaire is being deployed in many countries. You are free to discuss your participation in this study with the lead researcher, Professor Quintin Cutts at the University of Glasgow (contactable on +44 141 330 5619 / Quintin.Cutts@glasgow.ac.uk).

The questionnaire runs ok on Desktop Safari, Internet Explorer, Firefox and Chrome. We have noted occasional glitches with Safari on a tablet.

1. By selecting the "Yes" option below, I acknowledge that I have read and understood the explanation of the study above, that I am willing to complete the questionnaire, and that I have the opportunity to opt out once the study is explained in detail at the end of the questionnaire. Selecting "No" will exit you from the questionnaire.

◯ Yes
◯ No

Toys, games, and activities you enjoyed as a child

2. From memory, what are the your top three childhood games, toys or activities that you played with before the age of 8?

[ ]

3. Please indicate as a young child (age 0-8) your enjoyment levels for the following types of toys/game/activities. If you had no experience, tick "N/A". (Possible answers are N/A, None or low, Some, Moderate, High, and Very High)

- Dramatic and role play
- Puzzles - jigsaws, sliding grid, rubik's cube
- Memory games - Kim's game, match pairs game, Chinese whispers
- Sand play
- Guessing games - 20 questions, Guess who?, Who am I?, Cluedo, Botticelli
- Card and strategy games - Draughts/checkers, noughts & crosses/tictac- tie, Happy Families
- Water play - pouring, measuring

- Dough and clay play
- Painting, drawing, collage
- Design and tech toys (video games, mini bots)
- Outdoor play - climbing trees, den-building, making a snowman
- Organised outdoor play - sport
- Push/pull toys
- Fantasy play - puppets, Barbie dolls, Action men
- Reading on my own
- Being read to by others
- Cooking/baking
- Musical instruments
- Construction toys **without** moving parts: wooden bricks, simple Duplo/Lego blocks
- Construction toys **with** moving parts: Lego Technic, K-nex, Meccano, train sets, marble runs
- Toys with a programmable element
- Other (please specify)

4. What are the key features of the toys/games you most enjoyed as a young child? Click as many as you wish.

- ☐ Role play
- ☐ Fixing/improving things
- ☐ Strategy
- ☐ Building things from primitive components
- ☐ Problem solving
- ☐ Finding out how things worked
- ☐ Mechanisms that needed to be understood
- ☐ Programming
- ☐ Creativity
- ☐ Solitary play
- ☐ Cooperative play

☐ Other (please specify)

5. Considering your experiences at home between the ages of 0-8, how strongly do you agree/disagree with the following statements? (Possible answers are N/A, Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree)

- I was encouraged to explore
- Winning / getting the right answer was the main thing
- We tried to fix things in our house
- It was ok to fail
- Adults and siblings around me were into making things
- Others around me were into programming
- If things broke, they were thrown away
- It wasn't a competitive environment
- Trying new things out was encouraged
- I got told off or laughed at if I didn't succeed at tasks
- All of my time was timetabled into arranged activities

6. Tell us about aspects of the people and environment around you between the ages 0-8 that you think were the most influential on your development (e.g. "My parents read to me most nights", "Siblings were fixing cars and I watched/helped them", "My teacher encouraged me to practice maths outside of class", etc.), and in what way.

Your experience and success with computing technology...

7. **Moving to you now**, rate the extent to which the following statements describe your use of computers / apps in your **current** personal life, studies, or work. (Possible answers are Not At All, A Little Bit, Somewhat, Quite A Bit, A Lot.)

- I understand advanced features of a word processor such as styles, accurate positioning of images, and automatic section numbering.
- I avoid using computing technology wherever possible, preferring e.g. pencil and paper
- I fix technology issues that I experience
- I develop spreadsheets, databases or other systems to analyze data
- When the computer doesn't do what I expect I immediately ask someone for help.
- I can pick up new apps and computing technology without really thinking - I work out how to use it just by trying things out
- I use computing technology by writing down or memorising the steps I have to follow - if something goes wrong, I'm a bit lost
- I get other people to set up the computing technology in my house, e.g. internet, home entertainment, PC/laptop/tablet, printer
- I set up rules or macros in programs to automate common actions
- People at work or in my home life come to me with their technology problems

8. What of the following best describes your level of programming?

- ◯ You have never tried to program at all
- ◯ You have tried (on-line or face-to-face course, or just on your own) but didnfit feel you succeeded
- ◯ You have successfully written programs mainly for yourself (complete the box below)
- ◯ You have successfully written programs mainly requested / paid for by others (complete the box below)

If you've ticked one of the last two options, tell us about the most complicated program you've written, roughly how many lines of code, and the programming language used, in a few sentences.

9. If you've ever been in a programming class, how easy did you find it compared to most other individuals in the class learning at the same time? Please also indicate if you have no programming experience.

- ◯ I have no experience of programming
- ◯ Much harder than others
- ◯ Harder than others
- ◯ Easier than others
- ◯ Much easier than most others

10. What is your profession, line of work, or current subject of study?

11. What is your age?

- ◯ 18 to 24
- ◯ 25 to 34
- ◯ 35 to 44
- ◯ 45 to 54
- ◯ 55 to 64
- ◯ 65 to 74
- ◯ 75 or older

12. What is your gender?

- ◯ Female
- ◯ Male
- ◯ Other
- ◯ Prefer not to say

13. In which country did you mainly live before the age of 8?

Completing the questionnaire...

As you may have guessed from the questions in the second half of the questionnaire, the purpose of the study is to determine whether there is a link between early childhood activities **outside** school and the kind of study or work that people undertake later in life. In particular, we wish to determine whether specific childhood activities are linked to study or work that involves **computational thinking skills**.

In previous studies, only weak links have been discovered between prior academic success and success in computer science courses. This begs the question - what is it, then, that determines success in such courses? Some believe that success is based on some innate ability. The hypothesis underpinning this study is that success is not based on activities undertaken in typical school subjects, but instead, on a particular subset of those games / toys / play that are undertaken outside school.

Such links between success in computing courses and earlier activities, whether academic or external to school, are of interest because failure rates in computing courses in secondary and tertiary education are typically relatively high. We are entering a phase in most countries in the world where computing science teaching is being introduced into primary schools. If we can find activities that link with computing success and include them in early years and primary curricula, pupils stand a much greater chance of success in computing courses.

We will be analysing all questionnaire data during July and should have a preliminary report prepared by the end of August. If you wish to find out what we have discovered from the data, please email Quintin.Cutts@glasgow.ac.uk at the end of August, and we will send you a summary of our findings.

Having read this explanation, if you are happy for your data to be included in the study, please click Yes in the question below. If you navigate away from this page now, or if you select No below, your data will not be included in the study.

14. By clicking on the Yes option below, you acknowledge that you have read and understood the detailed explanation of the study and are content for your answers to be included in the study.

- ◯ Yes
- ◯ No

ORDINAL LOGISTIC REGRESSION

```
Call:
polr(formula = trait1 ~ x1 + x2 + x3 + x4 + x5 + x6 + x7 + x8 +
   x9 + x10 + x11 + x12 + x13 + x14 + x15 + x16 + x17 + x18 +
   x19 + x20 + x21 + x22 + x23 + x24 + x25 + x26 + x27 + x28 +
   x29 + x30 + x31 + x32 + x33 + x34 + x35 + x36 + x37 + x38 +
    x39 + x40 + x41 + x42 + x43)

Coefficients:
         Value Std. Error  t value
x12   0.669265     0.4606  1.45303
x13   0.674891     0.4694  1.43765
x14  -0.716153     0.5230 -1.36942
x15  -0.493537     0.6180 -0.79866
x22   0.074756     0.7060  0.10588
x23   0.652011     0.6904  0.94441
x24   1.115690     0.7621  1.46403
x25   1.767590     0.8192  2.15778
x32   0.089416     0.5095  0.17548
x33  -0.337606     0.5120 -0.65933
x34  -0.030697     0.5154 -0.05956
x35   0.029044     0.6827  0.04254
x42  -0.565122     0.5025 -1.12460
x43  -1.163844     0.5341 -2.17906
x44  -1.017452     0.5401 -1.88399
x45  -1.430190     0.6430 -2.22431
x52   1.229095     0.5119  2.40120
x53   0.457470     0.5378  0.85068
x54   0.398345     0.5331  0.74729
x55   0.276974     0.6394  0.43318
x62   0.611358     0.6721  0.90964
x63   1.165678     0.6869  1.69703
x64   0.374964     0.7012  0.53473
x65   0.798443     0.7608  1.04947
x72  -0.329828     0.4913 -0.67131
x73   0.418925     0.5053  0.82904
x74   0.799295     0.5401  1.47983
x75   0.717172     0.6161  1.16402
x82   1.074757     0.5050  2.12828
x83   1.344531     0.5046  2.66469
x84   1.338358     0.5316  2.51760
x85   0.401433     0.6611  0.60722
x92   0.159100     0.6854  0.23213
x93   0.484140     0.6760  0.71619
x94   0.993451     0.6846  1.45118
x95   1.507320     0.7200  2.09354
x102  0.098454     0.3994  0.24653
x103 -0.168103     0.4095 -0.41050
x104 -0.277462     0.4362 -0.63603
x105 -0.424991     0.5047 -0.84214
x112  0.490179     0.8383  0.58474
x113  1.133793     0.7803  1.45295
x114  0.159945     0.7500  0.21326
x115  0.423146     0.8014  0.52802
x122 -0.135128     0.4973 -0.27172
x123  0.328196     0.4892  0.67095
```

```
x124  0.430761     0.4897  0.87968
x125 -0.038906     0.5219 -0.07454
x132 -0.072262     0.4041 -0.17880
x133  0.620046     0.4245  1.46078
x134 -0.346415     0.4772 -0.72593
x135  0.052695     0.6711  0.07852
x142  0.338812     0.6060  0.55906
x143  0.559948     0.5929  0.94446
x144  0.294798     0.6136  0.48042
x145  0.394113     0.6000  0.65685
x152 -1.639437     0.6282 -2.60965
x153 -1.980330     0.5861 -3.37864
x154 -2.142483     0.5988 -3.57802
x155 -2.218576     0.5916 -3.74983
x162 -0.072319     0.4808 -0.15040
x163  0.371934     0.4911  0.75735
x164  0.665767     0.4863  1.36906
x165  0.365123     0.5205  0.70142
x172  1.055964     0.4221  2.50144
x173  0.664262     0.4082  1.62720
x174  0.589009     0.4633  1.27145
x175  0.751413     0.5540  1.35630
x182  0.171513     0.3857  0.44470
x183  0.316829     0.4380  0.72333
x184  0.340082     0.4440  0.76590
x185  0.202184     0.5301  0.38144
x192 -1.710779     0.6008 -2.84736
x193 -1.060646     0.6045 -1.75444
x194 -2.062083     0.6158 -3.34886
x195 -2.138306     0.6762 -3.16233
x202 -0.114500     0.4962 -0.23075
x203 -0.010438     0.5079 -0.02055
x204 -0.347661     0.5478 -0.63466
x205 -0.623807     0.5976 -1.04392
x212  0.124309     0.3755  0.33105
x213 -0.669127     0.4529 -1.47741
x214 -1.401514     0.6205 -2.25853
x215 -0.844750     0.7214 -1.17101
x221  0.694109     0.3137  2.21244
x231 -0.540992     0.3329 -1.62489
x241  0.558145     0.3056  1.82642
x251  0.115718     0.3145  0.36799
x261 -0.870556     0.3121 -2.78928
x271  0.654522     0.3238  2.02138
x281 -0.167841     0.3889 -0.43153
x291 -0.304933     0.5820 -0.52394
x301 -0.009231     0.3305 -0.02793
x311  0.083868     0.3015  0.27814
x321 -0.005880     0.2801 -0.02099
x332  2.700250     1.1076  2.43785
x333  1.166932     1.0345  1.12805
x334  1.474696     1.0381  1.42059
x335  2.452239     1.0693  2.29325
x342  0.036736     0.5758  0.06379
x343  0.153639     0.5878  0.26139
x344  0.968734     0.6157  1.57341
x345 -0.214511     0.9417 -0.22780
x352 -0.096535     1.2486 -0.07732
x353  0.596093     1.1987  0.49727
x354  0.575320     1.2137  0.47400
x355 -0.291855     1.2279 -0.23769
x362  0.632056     1.0590  0.59685
x363 -0.630046     1.0560 -0.59664
x364  0.558524     1.0464  0.53375
x365 -0.095032     1.1038 -0.08610
x372  0.971417     0.8345  1.16400
x373  0.401230     0.8451  0.47476
x374  0.689463     0.8035  0.85805
x375  0.422238     0.8490  0.49732
x382 -0.294313     0.3748 -0.78517
x383  0.134208     0.4421  0.30354
x384 -0.133277     0.4979 -0.26769
x385 -0.922507     0.8224 -1.12172
x392 -0.756586     0.4119 -1.83684
x393 -0.949112     0.4837 -1.96208
x394 -1.230848     0.5092 -2.41733
x395 -1.627651     1.3316 -1.22235
x402  0.021574     0.6832  0.03158
x403  0.530932     0.6570  0.80806
x404  0.854566     0.6589  1.29696
x405  1.027326     0.7562  1.35852
x412 -2.030681     1.3125 -1.54719
x413 -1.340584     1.1714 -1.14446
x414 -1.019986     1.1491 -0.88768
x415 -1.112677     1.1940 -0.93192
x422  0.218576     0.3203  0.68235
x423  0.819326     0.4550  1.80061
x424  0.200545     0.5712  0.35111
x425  1.746042     1.0246  1.70406
x432  0.240537     0.3188  0.75450
x433  0.200447     0.4226  0.47434
x434  0.986034     0.5363  1.83867
x435 -2.593365     1.6036 -1.61718

Intercepts:
    Value   Std. Error t value
1|2  3.2798  2.3926     1.3708
2|3  4.7959  2.3980     1.9999
3|4  6.0015  2.4032     2.4973
4|5  7.3435  2.4141     3.0419

Residual Deviance: 841.122
AIC: 1127.122
(927 observations deleted due to missingness)
```