

The importance of consulting children and young people about data literacy

Judy Robertson

University of Edinburgh, United Kingdom

E. Kay M. Tisdall

University of Edinburgh, United Kingdom



Peer-reviewed article

Citation: Robertson, J., & Tisdall, K. E. M. (2020). The importance of consulting children and young people about data literacy. *Journal of Media Literacy Education*, 12(3), 58-74. <https://doi.org/10.23860/JMLE-2020-12-3-6>

Corresponding Author:

Judy Robertson
judy.robertson@ed.ac.uk

Copyright: © 2020 Author(s). This is an open access, peer-reviewed article published by Bepress and distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. JMLE is the official journal of [NAMLE](#).

Received: February 7, 2020

Accepted: October 8, 2020

Published: December 14, 2020

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Competing Interests: The Author(s) declare(s) no conflict of interest.

[Editorial Board](#)

ABSTRACT

Given the importance of data skills to the economy and the skills shortage within data science, educational policy makers have identified the importance of including technical and analytical data skills in the school curriculum. An equally important aim is to educate children and young people to become *data citizens* who are aware of the current uses of data in society, able to use data to make decisions in their lives, and are actively engaged in critiquing the societal implications of future uses of data. The paper will explore the meanings of data citizenship, in light of the findings of a consultation with 96 children and young people (aged between 10 and 16 years old), from 11 schools in South East Scotland and the wider conceptual debates on citizenship and children and young people's rights to privacy, participation, and education.

Keywords: *data literacy, children, young people, citizenship, data privacy.*



Journal of Media Literacy Education

THE OFFICIAL PUBLICATION OF THE

NATIONAL ASSOCIATION FOR MEDIA LITERACY EDUCATION (NAMLE)

Online at www.jmle.org

INTRODUCTION

As data driven decision-making becomes more common within private, public, and commercial domains, children and young people¹ need to develop the technical and analytical skills – such as structuring data, conducting and interpreting statistical analyses, and creating data visualisations – to be able to utilise data to inform their own decisions. However, they should also have opportunities to engage with wider issues about how data is used in society. Organisations and education systems across the world are beginning to develop frameworks for data education (e.g., IDSSP Curriculum Team, 2019; Pittard, 2018; Tang & Sae-Lim, 2016). We argue that children and young people need to be involved in discussions about what they already know about data and what they would like to learn. Children and young people have participation rights under the United Nations Convention on the Rights of the Child (UNCRC) and evidence has accumulated on how their views and experiences can make highly productive contributions to educational developments (Lundy, 2007; Struthers, 2016). Secondly, as the use of data has proliferated rapidly in society and children and young people’s experiences relating to data usage may be different to those of adults, research is required to understand children and young people’s current knowledge, attitudes, and concerns about data.

The Data Education in Schools (DES) programme aims to educate all children and young people within South East Scotland about data (<https://dataschools.education/>). A goal of this government funded programme is to enable citizens to take an active part in shaping data driven innovation within the region. This project is primarily focussed on the formal educational context of primary and secondary schools. DES is developing an interdisciplinary curriculum framework for data education, drawing on reviews of pedagogical approaches to data literacy (Wolff et al., 2017), a mapping of existing curriculum outcomes to Wolff’s adaptation of the “Problem, Plan, Data, Analysis and Conclusions (PPDAC) problem solving cycle”, and consultations with the children and young people who will learn from this curriculum.

This article draws on a consultation event for the project with 96 young people aged between 10 and 16

years old. During the event, the children and young people took part in a set of group activities designed to gauge their initial knowledge of where data is used in everyday life to provoke discussion about data privacy issues and to give them the opportunity to suggest topics that they would like to pursue in the future. In analysing the data, the research team identified themes that were illuminated by considering concepts of children’s rights and data citizenship alongside data literacy. These concepts were not within the original research questions but are provocative ways to understand some of the central findings from the consultation. Below, the article first considers the academic grounding for these concepts. It then describes in more detail the consultation and its associated activities.

Citizenship and data literacy

As with many powerful concepts, citizenship is conceptualised in different ways across academic literature. A frequent starting point, particularly in the British literature, is the seminal definition provided by T. H. Marshall: “Citizenship is a status bestowed on those who are full members of the community. All who possess the status are equal with respect to the rights and duties with which the state is endowed” (Marshall, 1963, p. 87). This definition showcases the concept’s power: citizenship promises equality with respect to rights and duties, which are protected and enabled by the state. It is a desirable status with the goal to be included as a full community member rather than being excluded or marginalised. A particular legacy of Marshall’s concept is moving citizenship beyond a solely legal status associated with nation states to one of belonging to a community. These components contribute to the powerful claims of the citizenship concept and why many potentially marginalised groups seek to claim citizenship.

While T. H. Marshall’s definition and work on citizenship is seminal to the field, it has also been critiqued and subsequently developed. For example, substantial criticism has been made of liberal notions of citizenship, which require people to be autonomous, rational individuals able to assert and claim their rights in order to be recognised as citizens (Arneil, 2002). Further, the undue separation between public spheres of civic society and employment from the private spheres

¹ This article broadly addresses those up to the age of 18, following Article 1 of the UN Convention on the Rights of the Child. The phrase ‘children and young people’ is used because

older children in the UK prefer to be called ‘young people’ rather than ‘children’.

of family and care privileges certain types of citizenship participation in the public sphere and excludes others (Lister, 2007). Instead, a more relational notion of citizenship is widely advocated (Sandel, 1992; Sclater et al. 2009). This recognises the interdependency between individuals, collectivities, and communities and the ongoing struggle and performances of citizenship. There is an accumulation of literature that argues for a “difference-centred” model of citizenship (Moosa-Mitha, 2005). This turns around liberal interpretations, as echoed in Marshall’s definition, of citizens being equal in their rights and duties, to citizens being “differently equal” (Yuval-Davis, 1999). The issue is then not to transcend differences but that equality can be defined through difference (Moosa-Mitha, 2005). Thus, these strands of citizenship theory developments emphasise citizenship practices, the potential for a more inclusive concept, and a focus on citizens’ participation and agency to shape their own identities and communities.

Children and childhood have not been central to theorisations of citizenship generally. If they are mentioned at all, it is usually as a juxtaposition to the adult citizen – children are the quintessential non-citizens or – at best – “citizens in waiting” (Hill & Tisdall, 1997; Lister, 2007). There has been a strong strand of literature that seeks to include children as citizens – from those suggesting that children can claim some citizenship rights even if not others and could be semi-citizens in that way (Cohen, 2005; Cox, 2018). Moosa-Mitha (2005) puts forward powerfully that the difference-centred model of citizenship can include children as citizens as childhood becomes one form of difference amongst others. Others emphasise the practice and lived experiences of citizenship with children part of their communities (Baraldi & Cockburn, 2018).

These active notions of citizenship may now have penetrated discussions of children’s digital citizenship. Emejulu & McGregor (2019) write a sharp critique of the current framing of digital citizenship in formal and informal educational contexts, which is primarily “the ability to effectively make sense of, navigate and exist in the digital world” (p. 132). They argue for a process, “by which individuals and groups committed to social justice critically analyse the social, political and economic consequences of digital technologies in everyday life and collectively deliberate to take action to build alternative and emancipatory technologies and technological practices” (Emejulu & McGregor, 2019, p. 140). This call for action mirrors developments

generally in considerations of digital citizenship, which focus on citizens’ agency and the potential for progressive social change (Hintz et al., 2018). There, thus, is a call to recognise the politics of digital engagement and not see digital citizenship narrowly as acquiring technical skills.

Data literacy, then, is important for data citizenship. The Data-Pop Alliance – a consortium of researchers investigating the impact of big data – regards the promotion of data literacy as means to “empower [...] citizens and communities as free agents”, noting that it “empowers citizens to keep governments accountable and transparent” as well as enabling “local populations to understand and solve local problems” (Data-Pop Alliance, 2015, p. 8). In this view, data literacy is broadly “the desire and ability to constructively engage in society through and about data” (Data-Pop Alliance, 2015, p. 32). Wolff et al. (2017), after a review of how data literacy has been conceptualised in the literature, elaborate on the abilities required for the constructive engagement with data:

Data literacy is the ability to ask and answer real-world questions from large and small data sets through an inquiry process, with consideration of ethical use of data. It is based on core practical and creative skills, with the ability to extend knowledge of specialist data handling skills according to goals. These include the abilities to select, clean, analyse, visualise, critique and interpret data, as well as to communicate stories from data and to use data as part of a design process. (p. 23)

In this view of data literacy, learners encounter the data problem-solving cycle in real world settings with ethical awareness as a part of each of the stages in the cycle (i.e., problem, plan, data, analysis, communication). For Gould (2017), data literacy is a more up-to-date extension of statistical literacy in which datasets may include the personal. The Open Data Initiative (2016) characterizes it as:

the data literate individual understands, explains, and documents the utility and limitations of data by becoming a critical consumer of data, controlling his/her personal data trail, finding meaning in data, and taking action based on data. The data-literate individual can identify, collect, evaluate, analyse, interpret, present, and protect data. (p. 2)

Recent research with children about their attitudes and concerns about protecting personal data indicates that children and young people care deeply about their privacy although they may be initially unaware of the extent to which personal data trails can be disclosive (Livingstone et al., 2018; Zhao et al., 2019). Stoilova et

al. (2019) review the existing evidence relating to children and young people's data privacy, identifying three privacy contexts in the digital data ecology: interpersonal, institutional and commercial. The interpersonal context refers to data that is created and accessed about an individual, and extended through social networking, whereas the institutional context refers to data that is maintained about an individual by government agencies such as health care or education. In the commercial context, data about individuals is harvested, bought and sold by third-party companies. Stoilova et al. (2019) also distinguish between data types in terms of *data which is given* by users knowingly, *data traces* which users unknowingly leave behind through interactions with technology (such as data tracked by cookies), and *inferred data* that is new insights which are derived from linkage and analysis of multiple data sources. The data types and contexts were used to inform the design of the activities in the consultation study.

A consultation study with children and young people

The consultation aimed to explore children and young people's existing understanding of how data may be used within a community to give them the opportunity to express creative ideas about how data could be used in positive ways within the community and to gain insight into their intuitions and expectations about personal data privacy in interpersonal, institutional, and commercial contexts. The research questions were:

- RQ1: What do children and young people know about data and where it may be found in familiar contexts?
- RQ2: Given an introduction to data technologies which are used in real life contexts, to what extent can children and young people use these technologies in design ideas for community contexts?
- RQ3: When presented with a fictional personal data scenario, what are children and young people's opinions about the collection of different sorts of personal data, depending on

the organisation which collects it and the purpose for which the data is collected?

- RQ4: If given the opportunity, would children and young people wish to learn more about data, and, if so, what topics are of interest to them?

This was an exploratory study seeking to establish key elements for a subsequent curriculum programme to be co-designed with children and young people. Local schools were invited to bring groups of children and young people to a three-hour long event called "Data Town" at the University of Edinburgh. We invited participants in the age range 10-16 years old because examination of the curriculum expectations and outcomes within the local school system indicated that children and young people in this age range would be beginning to study related topics in school within literacy, social studies, technologies, and mathematics. An invitation to sign up was issued on Twitter. 96 children and young people from 11 schools – four secondary (42 people) and seven primary (54 people) attended.

While the schools were selected because of their diversity by socio-economic contexts², the school management was responsible for selecting which children and young people would attend the events; the study does not purport to be a representative sample but rather a purposive sample of children and young people who confirmed their interest to attend. Table 1 shows the breakdown of participant numbers, school type, and SIMD status of the school location. This article draws on findings from schools in daily transport distance of Edinburgh so were a range of urban schools but not rural nor remote ones.

Equal numbers of boys and girls attended although we did not ask children and young people to self-identify their genders. Other demographic and background characteristics were not collected for the children and young people as the collective and group-based activities meant that individuals' answers were not tracked through the day's activities. The event was themed around a fictional "Data Town" and all learners were allocated into mixed-age groups.

² Two measures are commonly used in Scotland to identify deprivation and poverty of pupils in schools: a relative measure of deprivation in small areas (<https://www.gov.scot/collections/scottish-index-of-multiple-deprivation-2020/#:~:text=If%20an%20area%20is%20identified,to%20services%2C%20crime%20and%20housing>) and percentages of

pupils receiving Free School Meals (<https://www.gov.scot/policies/maternal-and-child-health/free-school-meals/>). Across these measures, the schools come from a mix of deprivation deciles and percentages of Free School Meals suggesting, but not guaranteeing, that children and young people involved in the consultation had a mix of socio-economic backgrounds.

Table 1. *School information*

School type	Number of learners	Scottish index multiple deprivation decile
Primary	10	10
Primary	2	2
Primary	10	3
Primary	12	9
Primary	10	1
Primary	10	10
Secondary	10	7
Secondary	10	5
Secondary	10	8
Secondary	12	10

The session was led by four experienced teachers on the “Data Education in Schools” team. Participants were welcomed to the event and given an explanation of the event purpose. The three main activities were:

- *Activity 1 – drawing.* (RQ1). The purpose of Activity 1 was to explore in an open-ended way what the children and young people already know about data and how it may impact people’s lives by asking them to draw where data may be found in a town.
- *Activity 2 – themed design and discussion tasks.* The activity themes were chosen as a focus for the participants to learn more about a data topic and think more deeply about the implications for society. The themes included data for health (RQ3) and data in the community (RQ2).
- *Activity 3 – short personal data activities.* (RQ4). There was a series of short, fun activities spread throughout the morning, which were designed to illustrate that it is possible to inadvertently give out personal data online.

Further details of the activities can be found in Appendix.

Ethics. The study was approved by the Moray House School of Education & Sports’ Ethics Committee at the University of Edinburgh. All participants and their parents gave informed consent to attend the event and for the research team to gather and store drawings, discussion notes, and survey data. Parents also had the option to grant photographic and video consent; children and young people with parental consent for this had the option of filming their views. All the research team and adult helpers had the appropriate criminal record checks for working with children as required by Scottish law.

They were instructed about the local authority child protection procedures in the event that the children and young people raised concerns about inappropriate behaviour that they had encountered online.

Data gathering and analysis. Feedback forms containing a series of open-ended prompts were completed by each participant at the end of the event. The forms were transcribed and stored in electronic form on University of Edinburgh secure servers. The prompts were designed to enable the participants to express what mattered to them about the event, areas for future learning and aspects of the events that were unexpected or raised questions for them (see **Errore. L'origine riferimento non è stata trovata.**). Participants could choose whether to complete none, all or just some of the prompts. The forms were analyzed according to prompts, categorized, frequency counted per category, and summarized.

Table 2. *Feedback prompts*

The first thing I want to say is...
I would like to learn more about...
To help me learn more about data, I would need...
I was surprised that...
I liked it when...
I am not sure if...
I would like to ask...

The large sheets of paper from each table containing the children’s drawings of Data Town (Activity 1) were photographed in sections, one for each artist. The drawings were annotated by the artist and sometimes by the facilitator if requested. For analysis, the items written in the annotations were listed and categorised by category of place within the town (e.g. street, house,

shop). Researchers' notes of conversations at each group were also used to interpret items in the drawings.

For the community app design activity, participants recorded their design ideas on worksheets containing prompts about the app (e.g., what does it do, what problem will it solve, what phone features will it use) and prompts relating to data input, data processing, and data output. There was also a blank mobile phone template for the designers to sketch their interface design. The worksheets from each sub group were categorised by the prompts and summarised. For the Fitbuddy task, participants wrote what they liked and disliked about each advertisement on sticky notes that were later transcribed and categorised. Thus, the activities resulted in a large amount of qualitative data, which was largely analysed thematically using NVivo (Braun & Clarke, 2006) building up from categories as described above. A large sample of children and young people were involved, and it was possible to consider patterns across the data (e.g., did a large number of children and young people indicate an answer or a very small number). While not without debate, we followed typical practice in the social sciences and related fields in presenting enumerated responses in order to evidence these patterns (Braun & Clarke, 2006; Ritchie et al., 2003). As stated above, we discuss the findings in light of concepts of data citizenship, data rights, and data literacy as these concepts illuminated the findings and led to provocative conclusions for data education. These concepts were not embedded in the original research questions; rather, they emerged as ways to understand the study's findings that illuminate the analysis and its implications as well as disrupt certain assumptions.

FINDINGS

Overall reflections

Data from the feedback forms indicated the participants' general reflections about learning about data, both from the activities in the Data Town session and in the future. Not all of the participants chose to complete the evaluation form. The prompt about "the first thing I would to say is ..." produced largely positive answers about the participants' enjoyment of the activity and thank you messages to the organizers. The prompt "I liked it when..." received 78 responses which fell into four main categories relating to: 1) the drawing activities (28); 2) designing an app (18); 3) meeting the robots, Pepper and Sim Man (12); and 4) the short personal data activities (9).

There were 78 responses to the prompt "I would like to learn more about...". Beyond the common answer of "data", which is unsurprising given the title of the event, four main categories for responses emerged: 1) robotics (11 responses); 2) sensors (9 responses); 3) hacking (8 responses); and 4) data sharing and privacy (6 responses). When asked what they would need to help them to learn about data (72 responses), the young people identified resources including particularly: "experts" or "specialists" (8 references), personal research (6 references), and school-based learning (5 references). There were 75 answers to the prompt "I was surprised that...". Some of the answers (11) expressed surprise about the educational approach taken during the workshop (such as drawing activities or mixed age groups), and others were about the technology they encountered (13 responses). However, the most frequent categories which emerged related to the workshop's themes with 21 responses referring to "surprises about data privacy" and 19 responses about the "uses of data in the world". For example, the children and young people were surprised that data could be sold or shared between companies. Some of the children and young people were struck by the Data Town fictitious example of hacking and sale of personal information. On a more positive note, other participants noted that they were surprised that data could be used to help with the climate crisis. Many of the comments about data in the wider world revealed that, prior to the workshop, the participants did not appreciate how pervasive data is in modern society. Responses to the "I'm not sure if..." and "I would like to ask ..." prompts suggested that the activities had been successful in provoking interest and curiosity about data. One participant wondered if "hacking is always a bad thing", while another wanted to know "how data could be used for nefarious purposes". Other children and young people wondered whether "data is always safe" and whether "data is a good thing".

In summary, the feedback forms indicated that the participants enjoyed the event, were keen to learn more about data (particularly robots, sensors, hacking, and data privacy), and were surprised about how prevalent data is in society. They had questions regarding the collection and sale of personal data.

Activity 1: Where is data?

The drawings showed data in homes, gardens, a public park, a swimming pool, in shops, a hospital, and the street. There were 103 annotations indicating data

sources across the large drawings created at each group table. Of these, 34 of the items do not necessarily require data. Frequently, these items use electricity or are related to sustainability technology. For example, household appliances or electrical items such as a kettle or a toaster could potentially use data but commonly do not. It is impossible to say whether imaginative items like “transport portal”, “floating beds”, or a “teleporting ambulance” require data!

Within the home, the drawings show a range of security devices such as smart doorbells and locks, voice activated doors and windows, security sensors, and a “border to stop hackers”. Wi-Fi is commonly depicted. A range of convenience devices were drawn in the homes: Alexa personal assistants, smart fridges, a washing machine that weighs clothes, taps with sensors, and monitored toilets. At the hospital, patient records are collected and a drone was depicted that can collect things that patients need from their homes. Data collected in shops includes address data for deliveries,

purchases by customers, sales by staff, temperature data to stop ice-creams melting in the ice-cream parlour, and “new trends of fashion discovered”. In the garden, a “robot gives you food that it has grown and plants new food”, and the plants are self-watering. Robots also maintain the park, the swimming pool keeps a record of swimmers’ abilities, and self-driving police cars patrol the street.

In summary, the drawings suggest that the children and young people are aware of a number of everyday devices which currently, or potentially, use data. They were most familiar with uses of data in the home, but they also identified uses for data in public places or commercial premises. The drawing activity encouraged the children and young people to express creative ideas of where technology could be used in the near future. For some children and young people, there may be a confusion between electrical devices and those which use data (e.g., washing machines use electricity but not data).



Figure 1. An example drawing of a house and a park in Data Town

Activity 2: Data for Health

The participants were generally positive about features of the first FitBuddy app such as encouraging messages and challenges. Some participants approved of how the app would incentivize exercise and healthy behaviors. Others appreciated the vouchers and rewards as a means to save money although three comments pointed out that reward vouchers to redeem at a movie theater/cinema may be counterproductive because “cinemas aren’t sporty”.

The most commonly identified negative aspect of the advertisement was related to selling of data to third parties (e.g., “I don’t want my info sold. It shares most stuff”). Some participants mistrusted the vendor or wanted more information before making a decision to install it. One participant wrote, “the location part is a bad idea because if it got stolen or hacked then they would know where you go and they could abduct you”. Others did not like the idea of a company “stalking” their location or selling location information about them. Other negative aspects included the similarity of Fitbuddy to existing products and that it was “bribery” to incentivize exercise with monetary rewards.

Many of the comments weighed up both the positives and negatives together. Participants identified the trade-offs a consumer would make when installing the product: “I like that you can get rewards but I don’t like that it can track you” or “Dislikes: you give away your information. Likes: you get a voucher for it”.

When asked to consider FitBuddy 2, participants were generally positive about the additional features (“tracks more important things than the last app”, “It is good now that it tracks more stuff like heart rate, monitor sleep, eating”). Many participants identified sleep tracking, heart rate monitoring and nutrition as beneficial. Again, participants weighed up the tradeoffs of the advertisement: “It is still unfair with data being sold to external partners, but it is also good by helping you understand your health”. Once again many privacy and safety concerns were raised – “stop selling our data!”. One participant commented that “You’d feel violated if they sold this info about you”, and another stated that “Your periods should be personal”. Some participants would prefer to opt out of the collection of some data: “You should have the choice to turn off location and period trackers”. Doubts were also raised about the accuracy of the data.

The third advertisement for Healthy Heart attracted the most positive comments. Some participants explicitly noted that this version of the advertisement

compared favorably to the previous ones in terms of data privacy (e.g., “this is a lot better data wise”, “privacy is much better”). Participants liked the option for the user to control who would see their data and the purpose for which the data would be used (e.g., “Data is not being sold. It is only being used for research”). Some people enjoyed the possibility of contributing to research which could benefit others. The potential for personalized health care was identified – partly, as one participant noted, because it would stop you lying to your doctor. However, two participants expressed doubts about data reliability (e.g., “What happens if the health data is wrong because then the doctor might give you wrong medication?”). Some participants were still concerned about who would be able to access the data or the loss of privacy if their phones were stolen.

In summary, the participants displayed a range of attitudes to what they consider acceptable with respect to collecting personal health data. They identified positive and negative aspects of personal data collection, and their views on data sharing depended both on the nature of the data and the purpose of sharing. Some children and young people were uncomfortable about the collection of location and intimate health data including periods, mood, and heart rate. For some people, the idea of a company selling information about them was unacceptable. The participants generally reacted favorably to user privacy control features and the idea of sharing data to help with research.

Activity 3: Data in the community

The designs contained a wide range of ideas for facilitating connections between community members: apps to match people with other people who have specified hobbies, skills, or personalities. There were two designs which were not related to community – a single app that aggregates the features of all other apps installed on the phone and a sketching app. Three of the ideas are notable for suggesting very specific or original application areas: “Warns people of a natural disaster or cause that is happening or about to happen”, “contact nearest person with a cat or dog to help the person”, and the mouse catcher app that locates mice in a house by listening for mouse heartbeats.

Only two answers justified the phone features that are needed: “Notifications, SMS for messaging, microphone for accessibility, GPS to find people, camera for profile picture, check they’re a great helper before it stores data”. Often answers tended to list many phone features, some of which seem unlikely to be

required given the stated purpose of the application. For example, for an app which is “kind of like a big group chat where everyone says their queries and people respond”, the decision to use the map, camera, cloud storage and contacts features would deserve careful consideration. The use of those features could collect (and share) unnecessary data, suggesting children and young people could benefit from critically considering phone features and the potential privacy implications, especially when agreeing to prompts from commercial apps. It is possible that the written answers do not reflect the depth of verbal discussion about privacy and safety; researchers’ notes comment that groups discussed the importance of sending pictures first before exchanging details for skill transfer, or considered whether apps should check if someone has a criminal record before allowing them to register for a service.

The answers about data input for the app mostly refer to personal user data, with the exception of the earthquake app, which contains data about previous disasters, and the mouse finder app which knows what a mouse heartbeat sounds like. Some of the personal data is not required for the purpose of the app, and some data is incomplete for the purpose. For example, it is not clear whether a marketplace app would need user data of birth, and an app which helps people to find others with building skills would need to know more from the user than “how many years they have been an expert”. In general, the participants did not give much detail of what processing would be required in computational terms. Answers such as: “share it, process, store it”, “save progress”, “update every so often”, “look for what you want” illustrate the beginnings of an understanding of what data processing might be required. Other answers reported more details of *what* tasks would be performed, but not *how*: “Turn it into a profile with pictures of them doing helpful things. App robot finds people saying similar things nearby. Survey for improvement”. There were three answers that indicated specific computational tasks such as calculating an average star rating, matching key words, and sorting data into categories.

The prompt to write about the output of the app seems to have confused almost all of the participants, who interpreted it as “outcome”. They wrote about the impact the app would have on the lives of users. For example, “To help lonely people stay happy and positive” rather than the output of the data processing that the event organisers had sought to elicit.

In summary, while the participants generated design ideas for community connection apps, their answers about data input, processing, and output showed a lack

of knowledge about how such apps might work and what input would be necessary or desirable. They would also benefit from additional education about which phone features could collect what sort of data and whether it is appropriate for a given purpose.

DISCUSSION

The Data Town study explored children and young people’s awareness, knowledge, and attitudes to data.

It is encouraging that the children and young people in this study were curious about data and wanted to learn more about it. They were particularly surprised and concerned about the issue of commercial data sharing and selling but also wanted to pursue technical topics such as robotics or Internet of Things sensors. This suggests that data literacy would potentially be well-received if it were embedded within the school curriculum. It also illustrates that children and young people do not wish to restrict themselves to narrowly acquiring technical skills but are eager to learn about the politics of digital engagement. Children and young people showed their interest and abilities to engage in the critical learning advocated by Emejulu & McGregor (2019).

During the drawing activities, some people were initially unsure as to what the term “data” referred to, but with facilitation from the student teachers, identified a large range of devices which could collect data in domestic, public, and commercial spaces. The findings suggest that some of the children and young people are abreast of current and emerging data technologies. However, it would be unwise to assume that this is true of all children and young people, and therefore, children and young people should have the opportunities to learn how data technologies are used in everyday life. Because this is rapidly changing and non-specialist teachers may themselves not be familiar with such technology, the research community could assist in producing accessible guides to innovations in data technology.

The consultation results indicate the children and young people had design ideas for socially useful applications of data driven smartphone software. This is promising in terms of the Data-Pop Alliance’s (2015) perspective that data literacy should enable local populations to understand and solve local problems. The children and young people in this study would benefit from technical knowledge about how such ideas might be put in practice. Thoughtful ethical decision-making should be embedded throughout the design,

implementation and testing: “it requires making a series of small decisions that are often fraught, forcing reflection at each step” (Barocas & Boyd, 2017, p. 23). For example, when engaging in design challenges, learners should learn how to carefully consider the minimal necessary set of personal data that is required and whether the use of smartphone features such as location or camera has privacy repercussions. Ideally, learners would deepen their understanding of data processing algorithms, so that they can articulate to what extent and how their designs could be put into practice without treating the “computer” as a black box.

Children and young people would benefit from knowing about public datasets that could be used in socially useful applications as their answers about the data sources that could be useful in their app designs were focused on personal data that would be collected anew for their application. Their interest in how data could be used for health research shows potential for data activism where a suitable concern for data privacy can be combined with data sharing for the common good.

Children and young people’s discussions of the fictional personal health app indicate that they formed nuanced opinions about personal data collection, depending on the sort of data collected and how it would be shared. The participants were not apathetic, disengaged, or reckless in their views about data privacy and were aware of the trade-offs of exchanging data for services. On this basis, we recommend that data privacy education need not be limited to admonishing children and young people to safeguard their data; it could also include a reflective exploration of the children and young people’s values with respect to privacy and a realistic evaluation of the potential harms of sharing personal information.

The contexts for data privacy in activities ranged from personal, to institutional to commercial, with fluid boundaries between the contexts in the Fitbuddy scenario. In terms of the data types referred to by Stoilova et al. (2019), the participants appeared to be most familiar with personal *data given* and to some extent *data traces* and were surprised by the possibility of *inferred data*. On the basis of their consultation with young people, Livingstone et al. (2019) commented that: “Children focus on data they know they give, much more than data that is taken or inferred – and they think all of it is ‘none of their business’” (p. 3). In our study, at least some of the children and young people did consider it to be their business in the sense that they were concerned about the idea that their data could be tracked

or inferred in commercial contexts and actively objected to it (“Stop selling our data!”).

Limitations and future work

The findings of this study are an initial exploration of the knowledge and views of children and young people aged 10 to 16 on a broad set of data related topics. The present study was not intended to gather in-depth data to document the progression of technical knowledge or the development of data privacy views as children grow older, but these would be valuable areas to explore in the future.

There was a tension in designing the consultation session between educating the participants who may have known very little about data to begin with and biasing their answers towards data relating to particular themes. For example, we chose to include some demonstrations of robots as examples of how data could be used. As a result, many of the answers about what the children and young people would like to learn more about related to robotics when other areas of data which we did not present do not occur in participants’ feedback. On the other hand, it is a logical possibility that the participants could have had low interest in robotics and stated that they did not want to learn anything further about it.

Within the “Data Education in Schools” project, we have begun to act on the findings from the consultation. We have developed and tested a set of learning materials about personal health data as well as data privacy issues more generally, and the development of set of Internet of Things materials is underway. We are establishing a Young Person’s Advisory Group for the programme, so we might engage in further dialogue with young people on this topic.

In the future, it would be beneficial to explore potential gender inequalities in children and young people’s ambitions and interests with respect to data education. As the technology industry in general is highly male-dominated (Ashcraft et al., 2016), it would be unfortunate if this were perpetuated in the emerging field of data science. It is of particular importance that women and girls participate in discussions about data in society because of the gender data gap. Criado Perez (2019), a high profile author and data activist, demonstrates that there is a pervasive lack of data collected about women across many domains, stemming from the unstated assumption that men are default humans, and that the gender data gap has profoundly negative effects on women’s lives. It would be

beneficial to further research girls' awareness and aspirations with regards to data-related careers and also explore their views on how the ways in which data about women is collected and analysed can effect social change.

CONCLUSION

With rapidly changing technologies, children and young people's official digital literacy in schools has largely been framed by two adult concerns. The first is adults' fears for children and young people's safety and wellbeing, about how children and young people will use social media and growing concerns about where their data will go and the implications thereof. The second is a desire to train the data citizens of the future – to ensure that children and young people have the skills to compete in the workforce, to be leaders in the digital economy, and to navigate global society successfully. Another frame is to consider them as data citizens, benefitting from opportunities to practice and develop critical reflections on their and others' digital use, to apply and develop themes for social justice and data activism. This frame was not one that motivated the consultation study's original research questions but provides a provocative way to consider the study's findings.

Applying the frame of data citizenship would ensure that the discussions within schools incorporated this critical learning within their official lessons. It would help recognise that the division between the public sphere of the classroom or official school places and the private spheres of children's play time or home no longer hold, particularly in the digital age, so that children and young people's data citizenship permeates their everyday lives and must be considered holistically rather than officially in separate domains. This is even more evident with the reliance on digital connections as a response to the emergency measures for COVID-19 and the accompanying risk of digital exclusion (Tisdall et al., 2020).

A focus on data citizenship highlights how some children and young people are excluded from these digital communities, while others may be included. They can be excluded by lack of access to knowledge, they can be excluded by age restrictions (that are often unsuccessful in their application (OfCom, 2019), and they may lack opportunities to know of the options available. A citizenship of difference, that can consider the differences, intersectionalities, and inequalities embedded in children and young people's everyday

lives can be usefully applied to data citizenship, to recognise the multiplicities of children and young people's engagements and identities, and to ensure that school learning recognises children and young people's data citizenship now and not just in the future.

A rights-based approach, as required by the citizenship concept, means a balance is needed between children's rights to protection and their rights to participate. It is a familiar debate in the children's rights field and one that has challenged many policy areas, including education, to rethink their perceptions of children and childhood. Practices have often needed to change to ensure that children's rights to have their views given due regard in matters that affect them, to give and be provided with information, and to freedom of assembly and expression, are recognised. This balance is needed in children and young people's data citizenship particularly given the changing digital technologies and, thus, their implications for surveillance, privacy, and opportunities. Children and young people will benefit from critical learning where they can develop their own skills and knowledge to use in their everyday lives now as well as in the future.

ACKNOWLEDGEMENTS

The authors would like to thank the children, young people, teachers and student teachers who contributed to the consultation. We would also like to thank Kate Farrell, Tommy Lawson, Jo Spiller and Holly Linklater for their work in designing and leading the activities.

REFERENCES

- Arneil, B. (2002). Becoming versus Being. In D. Archad & C. M. Macleod (Eds.), *The Moral and Political Status of Children* (pp. 70-96). Oxford UPs.
- Ashcraft, C., Mclain, B., & Eger, E. (2016). Women in tech. National Centre for Women in Information Technology [PDF]. https://www.ncwit.org/sites/default/files/resources/ncwit_women-in-it_2016-full-report_final-web06012016.pdf
- Baraldi, C., & Cockburn, T. (2018). Introduction. In C. Baraldi & T. Cockburn (Eds.), *Theorising Childhood Studies* (pp. 1-27). Palgrave Macmillan.
- Barocas, S., & Boyd, D. (2017). Engaging the ethics of data science in practice. *Communications of the ACM*, 60(11), 23-25. <https://doi.org/10.1145/3144172>

- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101. <https://doi.org/10.1191/1478088706qp063oa>
- Cohen, E. (2005). Neither Seen Nor Heard. *Citizenship Studies*, 9(2), 221-240. <https://doi.org/DOI:10.1080/13621020500069687>
- Cox, P. (2018). Children and Relational Citizenship. In M. Shaw & S. Bailey (Eds.), *Justice for Children and Families* (pp. 42-49). <https://doi.org/doi:10.1017/9781108619554.007>
- Criado Perez, C. (2019). *Invisible Women*. Chatto & Windus.
- Data-Pop Alliance. (2015). Beyond Data Literacy [PDF]. Data-Pop Alliance. <http://datapopalliance.org/wp-content/uploads/2015/11/Beyond-Data-Literacy-2015.pdf>
- Emejulu, A., & McGregor, C. (2019). Towards a radical digital citizenship in digital education. *Critical Studies in Education*, 60(1), 131-147. <https://doi.org/10.1080/17508487.2016.1234494>
- Gould, R. (2017). Data literacy is statistical literacy. *Statistics Education Research Journal*, 16(1), 22-25.
- Hill, M., & Tisdall, K. (1997). *Children and Society*. Routledge.
- Hintz, A., Dencik, L., & Wahl-Jorgensen, K. (2018). *Digital citizenship in a datafied society*. <http://ebookcentral.proquest.com/lib/ed/detail.action?docID=5613296>
- IDSSP Curriculum Team. (2019). Curriculum Frameworks for Introductory Data Science [PDF]. http://idssp.org/files/IDSSP_Frameworks_1.0.pdf
- Lister, R. (2007). Unpacking children's citizenship. In A. Invernizzi & J. Williams (Eds.), *Children and citizenship* (pp. 9-19). <https://doi.org/doi:10.4135/9781446214756.n1>
- Livingstone, S., Stoilova, M., & Nandagiri, R. (2018). Conceptualising privacy online: what do, and what should, children understand? *Parenting for a Digital Future*, 3, 1-4.
- Lundy, L. (2007). Voice is not enough. *British Educational Research Journal*, 33(6), 927-942. <https://doi.org/10.1080/01411920701657033>
- Marshall, T. H. (1963). *Citizenship and Social Class in Sociology at the Crossroads and Other Essays*. Heinemann.
- Moosa-Mitha, M. (2005). A Difference-Centred Alternative to Theorization of Children's Citizenship Rights. *Citizenship Studies*, 9(4), 369-388. <https://doi.org/10.1080/13621020500211354>
- Oceans of Data Institute. (2016). Building Global Interest in Data Literacy. <http://oceansofdata.org/our-work/building-global-interest-data-literacy-dialogue-workshop-report>
- OfCom. (2019). Children and Parents [PDF]. *Ofcom*, 220. https://www.ofcom.org.uk/__data/assets/pdf_file/0023/190616/children-media-use-attitudes-2019-report.pdf
- Pittard, V. (2018). *The Integration of Data Science in the Primary and Secondary Curriculum Final Report* [PDF]. <https://royalsociety.org/~media/policy/Publications/2018/18-07-18-The-Integration-of-Data-Science-in-the-primary-and-secondary-curriculum.pdf>
- Ritchie, J., Lewis, J., & Elam, G. (2003). Designing and selecting samples. In Jane Ritchie & Jane Lewis (Eds.), *Qualitative research practice* (pp.77-108). Sage.
- Sandel, M. (1992). *Liberalism and the Limits of Justice*. Cambridge University Press.
- Sclater, S. D., Ebtehaj, F., Jackson, E., & Richards, M. (2009). *Regulating Autonomy: Sex, Reproduction and the Family*. Hart.
- Stoilova, M., Livingstone, S., & Nandagiri, R. (2019). Children's data and privacy online Growing up in a digital age [PDF]. <http://www.lse.ac.uk/media-and-communications/assets/documents/research/projects/childrens-privacy-online/Evidence-review.pdf>
- Struthers, A. E. C. (2016). Breaking down boundaries: voice and participation in English primary education. *International Journal of Children's Rights*, 24(2), 44-468. <https://doi.org/10.1163/15718182-02402011>
- Tang, R., & Sae-Lim, W. (2016). Data science programs in U.S. higher education. *Education for Information*, 32(3), 269-290. <https://doi.org/10.3233/EFI-160977>
- Tisdall, E. K. M., Powell, M. A., Reid, K., Kong, G., Davidson, J., Doyle, M., Harris, J., & Sutton, N. (2020). Independent Children's Rights Impact Assessment on the response to COVID-19 in Scotland [PDF]. <https://cypcs.org.uk/wp-content/uploads/2020/07/independent-cria.pdf>
- Wolff, A., Gooch, D., Cavero Montaner, J., Rashid, U., & Kortuem, G. (2017). Creating an understanding of data literacy for a data-driven society. *Journal of Community Informatics*, 12(3), 9-26. www.ci-journal.net/index.php/ciej/article/view/1286

Yuval-Davis, N. (1999). Ethnicity, gender relations and multiculturalism. In R. D. Torres, L. F. Mirón, & J. X. Inda (Eds.), *Identity and Citizenship* (pp. 112-126). Blackwell.

Zhao, J., Wang, G., Dally, C., Slovak, P., Childs, J., Van Klee, M., & Shadbolt, N. (2019). *'I make up a silly name': Understanding Children's Perception of Privacy Risks Online*.
<https://doi.org/10.1145/3290605.3300336>

APPENDIX

Session Design

Purpose

The event was themed around a fictional Data Town, with all learners allocated into mixed-age group ‘streets’ containing around 9 learners. Participants were welcomed to the event and given an explanation of the event purpose. They were then asked to complete a short survey about their use of technology and learning about data in school. Three main activities followed:

1. *Drawing.* The participants were asked to draw where data is within a town using large sheets of paper and pens. If they were stuck for ideas, they could look at twelve prompt cards which contained icons representing: a shopping trolley, a police car, a road, a town hall, a park, a hospital, a bike, a shop, a water tap, an aeroplane and city buildings. The purpose of activity 1 was to explore in an open ended way what the young people already know about data and how it may impact people’s lives.
2. *Short personal data activities.* There was a series of short, fun activities spread throughout the morning which were designed to illustrate that it is possible to inadvertently give out personal data online. The young people had the option to complete paper versions of online quizzes which give away personal information (including birth day and month, initials of forename and surname and favourite colour) in exchange for a “unicorn name” or ‘superhero name” (see Figure 1).



Figure 1. Example personal data activity

If they elected to do these activities they got a sticker for their sticker sheet. Participants also indicated whether they would install the FitBuddy app by adding a sticker to their sheet, and received stickers to indicate they had completed the sustainability and citizenship design exercises. Throughout the morning, the workshop leader interrupted activities to show news flashes from Data Town TV about the consequences of their decisions (see the example in Figure 2). It was revealed that the Mayor of Data Town had sold information about favourite colours to a marketing company, that FitBuddy had sold location data to a company who used it to decide a location for a new gym, and that University researchers had found a cure for a disease based on the Healthy Heart dataset. At one point, the news reported that Data Town had suffered a data breach and that anyone who had collected particular stickers for sharing personal information had their personal information hacked, and would receive a sticker to indicate this. The Mayor role was played by a member of the research team who attempted to justify his decisions about selling data or failing to protect data when questioned by participants or other members of the research team.



Figure 2. Example news flash

3. *Themed design and discussion tasks.* Each street took part in one of three activities with the following themes. The themes were chosen as a focus for the participants to learn more about a data topic, and think more deeply about the implications for society. Health and wellbeing, citizenship and sustainability are all cross-cutting themes within the curriculum in Scotland.

a. *Data for Health.* The participants were shown a series of four adverts for a fictional product called FitBuddy (see Figure 3). They were asked to discuss the adverts and to write on sticky notes what they liked and disliked about each. The first advert described a product which gathers users' step counts and location data. The user could gain tangible rewards for their steps including vouchers to spend in sports shops, gyms or cinemas. Small text on the advert stated that location and fitness data would be sold to external partners to provide the user with personalised offers. There is a similar commercial product available which is used by young people in the region (as we found in a previous consultation workshop). The second version of the advert was for a newer version of the product which collected more intimate data including heart rate, nutrition, sleep, mood and menstrual cycle. Again, the advert stated that the data may be sold to other companies. This scenario was intended to raise discussion about where individuals might draw boundaries about privacy. Finally, the third advert was for a product called Healthy Heart. It explicitly stated that the user can control their data. Although the app collects heart rate data for the benefit of the user, it will never be sold to external partners for a profit, and the data will only be shared with medical experts if the user chooses. The purpose of this scenario was to explore if the participants' views on data sharing were dependant on the purpose for sharing.

There were breaks for food, drinks and outdoor play between activities. During the event, the participant had the option to film videos in which they interviewed a partner about topics which emerged during the day. The purpose of this was to gather views in a way which was led by the issues which were important to the young people which the research team might not anticipate in advance.

The research team included twenty-five new teachers who were in the final stages of their postgraduate teaching degree at the University. These teachers, collectively referred to as 'Alexas' during the study, facilitated the group discussions, led activities, assisted the young people, observed and wrote research notes.