Empowering Students to be Data Literate: The Design and Implementation of a Learning Environment to Foster Critical Data Literacy

Andria Agesilaou, Eleni A. Kyza
aa.agesilaou@cut.ac.cy, Eleni.Kyza@cut.ac.cy
Media, Cognition and Learning Research Group
Department of Communication and Internet Studies
Cyprus University of Technology

Abstract: Business giants collect and analyze personal data on a massive scale, including those of children’s, to predict behavior and inform algorithmic decision-making. Such efforts reflect a form of “biopower”, allowing control over other people, whose data are exploited. To safeguard children’s data rights, scholars call for the need to empower children through critical data education. In our work we seek to understand how children can be empowered to reflect on the use of their own self-tracking devices and gain a deeper understanding of the political economy. This paper describes the design-based research process of designing an educational intervention that employed the use of activity trackers by elementary school students. This effort helped us peel back the layers of children’s understanding of their own digital traces and supported their engagement in critical discussions. The paper concludes with lessons learned from the implementations on developing learning materials to foster critical data literacy.

Keywords: critical data literacy, personal digital data, design-based research, data privacy, self-tracking

Introduction
In an age when almost all dimensions of our lives are constantly digitized and datafied, concerns have been raised about the detailed personal information that is continuously generated through ubiquitous technologies. The term “datafication” is often used in the literature to describe how people’s daily activities, practices and sensitive personal data are turned into digital information (Van Dijck, 2014). “Dataveillance” as a result of constant datafication, is the tracking and use of data for the purposes of monitoring people’s activity (Van Dijck, 2014), and often happens through opaque algorithmic techniques. Such techniques collect and store people’s personal information in large datasets, giving rise to Big Data accessed and exploited by third parties. Lupton (2020) explains two uses of personal data by third parties: (a) Collection and use of personal data to improve or market products or sell them to other companies for other types of exploitation, and (b) User profiling to inform the process of algorithmic decision-making with the support of artificial intelligence and deep learning models.

Scholars argue that this algorithmic decision-making has started to “unbalance” civil society, since it is used to sort and categorize individuals based on their data profiles (Pangrazio & Selwyn, 2020). Such categorization may have negative impact on individuals, including children, and may be unjustly used to exclude citizens from future opportunities. This data profiling and algorithmic decision-making is characterized in the literature as a form of soft “biopower” which affects individuals through a series of dividing practices manifested within relations of power by governments, business giants and other authorities (Lupton 2020; Pangrazio & Selwyn, 2020). Foucault discussed the issue of biopower long before the digital age as a means to control people through their personal information. Nowadays, these control practices can lead to the stratification and categorization of citizens based on data profiling (Pangrazio & Selwyn, 2020) because individuals lack the necessary knowledge and skills to identify where, how and why their personal data are being collected, stored, monetized and manipulated by others (Livingstone, StoiIova & Nandagiri, 2019; Pangrazio & Selwyn, 2020). This form of biopower affects children in a significant manner. As children and youth are constantly connected to digital devices, they become producers of a vast amount of personal data. Collecting, processing and assessing information on appearance, growth, development, health, social interactions, moods, behaviors, educational achievements and other features, without the consent of the children and, on several occasions, of their guardians, undermines their rights (Lupton & Williamson, 2017) and jeopardizes their future.

Scholars suggest that researchers and educational designers should focus on how to empower students and youth through critical data education and the development of critical data literacies (Pangrazio & Selwyn, 2020; Stornaiuolo, 2020). Children and youth’s education about critical data literacy is necessary for developing
the data skills to navigate in a datafied culture and see themselves as people with data rights (Stornaiuolo, 2020). Nonetheless, the kind of pedagogical support needed to empower students for recognizing, understanding and acting on their data practices should be further explored (Stornaiuolo, 2020). This study continues work reported by Agesilau and Kyza (2021) on upper elementary school students’ awareness of digital personal data, feelings of ownership of digital data and their understanding of privacy risks; informed by this work, this paper reports on how these findings informed the iterative design of a scaffolded inquiry environment designed to engage students in critically discussing about personal digital data in a formal educational setting.

**Literature review**

**Empowerment through critical data literacy**

Philip, Schuler-Brown and Way (2013) highlight the importance of “empowering students with knowledge” (p.106), to help students understand the role of their personal data contributions to the big data industry and exert control over the commercial exploitation of their data by third parties. Livingstone, Stoilova and Nandagiri (2020) discuss the difficulty of empowering children via personal data literacy because of the complexities involved in data protection and privacy regulations. They argue that learning how to navigate the contemporary digital landscape in terms of privacy is necessary but not sufficient to protect and empower children (Livingstone, Stoilova & Nandagiri, 2020). What is needed for children to ensure their data rights in a datafied society is a deeper critical understanding of the digital environment behind the digital interface, including the social relations and interactions that are embedded within it and which drive the emergence of the current business models (Livingstone, Stoilova & Nandagiri, 2020).

The idea of helping students acquire informed understanding of the massive dataveillance of everyday activities and be able to act based on this understanding has its roots in Paulo Freire, who talked about deliberate education in which teachers are not the authority, and students have the opportunity to question their reality and assume an active role in changing their society (Freire, 1970). This pedagogy supports students’ opportunity to make decisions and act as agents of change in a datafied society, through the adoption of protective and resilient behavior towards the exploitation of their digital traces. It is of paramount importance to explore the extent to which students understand their personal data practices and the implications these have in the big data era. Such understanding will inform current efforts on how to better guide children through education for consciously being involved in practices that lead to the generation of personal and sensitive digital data. Such approaches have been reported as missing from contemporary education (Pangrazio & Sefton-Green, 2020).

Pangrazio and Sefton-Green (2020) discuss how data literacy approaches reported in the literature have been implemented in educational settings. They first note that research on data literacy pedagogy is limited. A conclusion derived from their study is that current data literacy efforts are unable to capture the range of skills and dispositions needed to tackle datafication and that the term needs more theorization to address a more critical and practical approach (Pangrazio & Sefton-Green, 2020). Pangrazio and Selwyn (2020) suggest that a critical data education should reflect the following objectives: (1) “Developing understanding of datafication processes, especially around what becomes a ‘data object’, how it is processed and used; (2) Critically reflecting on the importance of metrics in social media use, and the significance of numerisation in society; (3) Developing awareness and understanding of the political economy of digital platforms, including the capacity to track, profile and predict; (4) Identifying how digital platforms (including the interface and design) position users and shape their participation; (5) Identifying and using a range of strategies and tactics to protect their personal data; and (6) Gaining a sense of the longer-term implications of personal data practices” (Pangrazio & Selwyn, 2020, p.6).

**Children’s awareness of data privacy and third-party exploitation**

Children are immersed in a variety of social networking activities that entail the disclosure of personal information. According to Livingstone, Stoilova and Nandagiri (2019) the full complexity of children’s privacy online can be captured by three main types of relationships: between an individual and (i) other individuals or groups (‘interpersonal privacy’); (ii) a public or third sector (not-for-profit) organization (‘institutional privacy’); or (iii) a commercial (for-profit) organization (‘commercial privacy’). In addition, Livingstone et al. distinguished between the types of data being shared online: (a) *Data given* by the individuals knowingly during their online media consumption, (b) *Data traces*, which are data left online, most of the times unknowingly, and are captured from data-tracking devices or plug-ins such as geolocation, cookies, etc., and (c) *Inferred data* which are created from the combination of the two former categories, using algorithms leading to the profiling of the data subjects. The inferred data fall into the category of commercial privacy. According to Livingstone, Stoilova and Nandagiri (2019) much less research has been focused on inferred data.
Research on children’s online practices shows that their online media consumption has dramatically increased, even though children are reported to lack the necessary understanding and skills to securely navigate the digital environment. According to a survey conducted with 5,436 Canadian students in grades 4-11 (Steeves, 2014), 32% of students in grades 4-6 have a Facebook account and 16% have a Twitter account. Steeves also examined students’ privacy attitudes towards strangers and marketers. A large percentage of students (39%) stated that companies are not interested in what they do online, and 68% believed that if a website has a privacy policy, it means they will not share students’ personal information with others. Livingstone, Stoilova and Nandagiri (2019) reviewed 105 empirical research studies from countries around the world, to assess students’ skills in regulating their online privacy and the level of their understanding of the commercial use of their personal data. Children, ages 11-16, are reported to have difficulties in understanding the concept of personal data, have gaps in their knowledge and awareness of issues surrounding their online privacy and display a general inability to see how their data might be valuable to anyone (Livingstone et al., 2019).

Chi, Jeng, Acker and Bowler (2018) argue that in order to understand teens and their interactions with data, a holistic approach should be followed concentrating not just on behaviors, but also on the interplay of behavior with cognition and affect. They investigated how youth (aged 11-18) feel, think and behave in regard to their personal data online employing the Affective-Behavioral-Cognitive (ABC) model of attitudes. They found that young people were confident about their digital skills when using technology and were favorable towards having total control over their personal data. At the same time, they showed fear or anger when talking about who could access their online data, and the loss of empowerment and control over their digital data. These results, though, refer to students’ perceptions, not their actual skills. The study participants showed a lack of awareness of what data were released to other entities, and how this might affect them.

Ey and Glenn Cupit (2011) interviewed children five to eight years old to examine their understanding of online risks, as well as their management strategies. In general, children identified several internet risks, but were unable to connect these risks with potentially dangerous situations. Specifically, when children were asked about placing their personal information on the internet, 54 out of the 57 children reported that it is not safe to give your name, address or photo on the internet. Similarly, Selwyn and Pangrazio (2018) ran a series of workshops with young people, ages 13-17, to explore their current understandings of personal data and data mining from third parties, and also to help them develop resistance practices. The findings of the study showed that the workshop activities were successful in making young people more conscious about their data trails, but less successful in cultivating the participants’ need to act differently and employ resistance practices, such as using software to identify who might be tracking their personal data.

Such studies highlight the lack of young people’s awareness of their personal digital data and the implications of their data actions online, but also expose the inability of young users to adopt protective and resilient practices towards the exploitation of their digital traces.

Research aim
This study is part of a design-based research project investigating upper elementary students’ awareness of data privacy issues and psychological ownership of digital data. We are interested in understanding how students can be empowered through scaffolded inquiry experiences to reflect on their own personal digital data and gain a deeper understanding of the digital infrastructure and the political economy. Specifically, we focus on how the objectives of critical data education can be approached pedagogically, especially since the theory in this area is still evolving and research-based curricula are limited. Therefore, the main goal of this study was to analyze our design efforts to explore which pedagogical features have potential to scaffold students’ critical discussions about personal digital data in formal educational settings. In this paper, we report on the design iterations and findings, and conclude with a set of empirically driven design guidelines for developing learning environments that can engage students in critical reflections on the uses of their personal digital data.

Methods
Theoretical foundations of the learning environment
The design of the learning environment was informed by Freire’s critical pedagogy ideas and critical data education, as discussed in Pangrazio and Selwyn (2020): the goal of the designed materials was to scaffold students in understanding their personal digital data contributions and empowering them in making decisions based on their own values. The learning environment is also informed by critical theories of data literacy. The pedagogical approach for developing the curriculum is inquiry-based learning in an effort to engage students in authentic learning inquiries about their own personal digital data. Inquiry-based learning usually consists of five
phases: Orientation, Conceptualization, Investigation, Conclusion and Discussion (Pedaste et al., 2015). The inquiry-based learning approach has been chosen because it encourages Freire’s view of education, since the learner in inquiry-based learning is expected to assume responsibility of their own knowledge-building process, and students’ active participation is highly emphasized.

Design process
Learning materials aiming to foster critical data literacy are limited in formal educational settings (Pangrazio & Sefton-Green, 2020). The learning environment, developed for upper elementary students, was designed following a design-based research approach (Barab & Squire, 2004; Design-Based Research Collective, 2003), and implemented in two suburban schools. It was revised through two design cycles, which we report next. One of the main objectives of this learning environment was to develop activities that scaffold students in engaging in critical discussions and reflections of their own personal data. Towards this goal we developed a learning environment that employed the use of activity trackers by students. Activity trackers were chosen because they are popular smart devices with these ages and allowed students in our studies to unobtrusively generate their own personal digital data through physical activity. At the same time access to activity trackers is not usually prohibited in schools, as social media often are, thus making the alignment with the elementary school curriculum easier. Each student was provided with an activity tracker adjusted to his/her personal information (such as pseudonym, sex, year of birth, height and weight) after receiving signed informed consent by their legal guardians. Figure 1 presents an overview of the design cycle of the learning environment.

Design cycle 1

<table>
<thead>
<tr>
<th>1st iteration</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pedagogical approach:</strong></td>
</tr>
<tr>
<td>- Socio-scientific inquiry-based learning (SSIBL)</td>
</tr>
<tr>
<td>- Problem-based scenario</td>
</tr>
<tr>
<td><strong>Implementation:</strong> Two 5th grade classes (21 students, 10 female)</td>
</tr>
<tr>
<td><strong>Aim:</strong> Explore students’ prior knowledge and beliefs to better understand how to scaffold them in cultivating critical data literacy</td>
</tr>
</tbody>
</table>

Design cycle 2

<table>
<thead>
<tr>
<th>2nd iteration</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pedagogical approach:</strong></td>
</tr>
<tr>
<td>- Inquiry-based learning</td>
</tr>
<tr>
<td>- Gamified approach</td>
</tr>
<tr>
<td>- Two hypothetical scenarios</td>
</tr>
<tr>
<td><strong>Implementation:</strong> One 5th grade class (19 female)</td>
</tr>
<tr>
<td><strong>Aim:</strong> Test the learning materials and the data collection protocols</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>3rd iteration</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pedagogical approach:</strong></td>
</tr>
<tr>
<td>- Inquiry-based learning</td>
</tr>
<tr>
<td>- Gamified approach</td>
</tr>
<tr>
<td>- Three hypothetical scenarios</td>
</tr>
<tr>
<td><strong>Implementation:</strong> One 5th grade class &amp; two 6th grade classes (65 students, 32 female)</td>
</tr>
<tr>
<td><strong>Aim:</strong></td>
</tr>
<tr>
<td>- Understand students’ stances towards their personal digital data</td>
</tr>
<tr>
<td>- Investigate students’ awareness of personal data monetization by third parties</td>
</tr>
</tbody>
</table>

Figure 1: Design-based process of the learning environment.

Design cycle 1

**Aim:** Gain insights into students’ initial understanding of the ethical dimensions and privacy issues related to using activity trackers and generating a vast amount of self-tracking data. We wanted to identify students’ prior knowledge and beliefs to better understand how to scaffold them in cultivating a critical stance about personal digital data.

**Implementation:** The learning environment was implemented in two fifth grade classes with a total of 21 students (10 female). During the enactment students worked both individually and collaboratively, divided in four groups.

**Approach:** Socio-scientific inquiry-based learning, SSIBL (Levinson, 2018). The learning environment was structured around a driving question with students assuming roles of different stakeholders.

**Learning sequence:** An educational intervention that employed activity trackers and offered students the opportunity to collect and critically consider data from their own physical activity, was designed by a co-design team, with input from collaborating teachers. The module was designed for four 80-minute lessons. The first
lesson begins with a short, animated story ending with the driving question “Would you use an activity tracker to monitor your physical condition?” Each student was provided with an activity tracker, adjusted to his/her personal information, and performed a sequence of physical activities in the school yard. Students kept the activity trackers for three days to be used during school hours. The second lesson was about physical education and nutrition. During the third lesson, students discussed the graphical representations of their own physical activity data and assumed roles to investigate the pros and cons of using activity trackers as presented by the following stakeholders: companies who develop activity trackers; scientists and doctors; users of activity trackers; and insurance companies. The lesson ended with a debate between representatives of each stakeholder. During the fourth lesson, students discussed how the personal data from activity trackers are provided to third parties and what ethical aspects this might involve. Examples of the privacy policy and terms and conditions of various websites, including of the activity trackers’ platform, were discussed.

Data collection and analysis: Data were obtained from video-recordings of three student groups during the third and fourth lessons which featured their own personal digital data discussions. Videos were carefully reviewed based on the critical incident technique (Flanagan, 1954) to identify incidents in which students had been engaged in critical discussions about the ethical dimensions and the privacy issues related to using activity trackers.

Findings: Students engaged in critical discussions about their self-tracking data. They believed that their data can be used by others, but most of them could not explain who these people might be. Students also believed that access to their self-tracking data by others is less important or serious as compared to someone having access to their conversations on social media. Even though we observed incidents of critical discussions in all groups, those reflected the opinions of some students. Not all group members participated in the discussions and this raised the question of how to scaffold each student in voicing their opinions on the matter. Similarly, during the debate activity, students identified arguments supporting their assigned stakeholder role, but these arguments did not reflect students' personal opinions and stances towards the use of their self-tracking data. We also noticed that students had difficulties in interpreting the graphical representations and in understanding their own data visualizations and needed further scaffolding to engage in such discussions.

Implications for design
The SSIBL pedagogical approach used in the first iteration was successful in helping students recognize the plurality of stakeholders involved in the issue in real life, and in helping us understand how students engage in collaborative discussions about the uses of personal digital data. In our next iterations, though, the aim was to develop learning activities that would be more personally meaningful to students and to give the opportunity to each student to shape the collaborative discussion. To address this goal, the SSIBL ethical dimension remained as the core of the learning environment, but we proceeded with the following revisions:

1. We decided that it would be better to have students at this age reflecting on their personal data rather than assuming roles of different stakeholders. In combination with their experience of using the activity trackers during the implementation, this would provide students with the opportunity to express their own personal opinions about the use of their self-tracking data.

2. Students were less engaged in some activities, so we decided to embed gamification elements in the learning environment to increase their motivation to participate. Using the ABC model of attitudes, activities were re-designed to attend to affective, behavioral and cognitive components of the task.

3. Based on findings indicating that students were not aware of who might use their personal data, we decided to design activities that would allow students to contextualize these ideas using multiple scenarios that resembled situations they may encounter in their personal lives. For this reason, we developed three hypothetical scenarios of specific companies asking for access to students’ self-tracking data. We used these activities to help us gain additional insights about how students perceived the privacy of their personal data in actions (vs. based only on self-reported interview data).

4. We decided to keep the collaborative aspect of the learning environment and have students discuss in groups as the debate was important for externalizing ideas and highlighting the complexity of the situation; however, it was also important to encourage the participation of all students in the discussion. We, therefore, decided to scaffold the group decision making process in each scenario, by introducing an activity during which all students would first reflect on and articulate their personal opinion to the question raised in each hypothetical scenario.

5. We simplified the complex data visualizations and embedded them in the hypothetical scenarios to provide a meaningful context for interpretation.
Design cycle 2

Aim: With the second design attempt we aimed at gaining a deeper understanding of students’ stances towards their own personal data in relation to online privacy and data monetization by third parties by increasing student engagement with the ideas both at an individual and at the group level.

Implementation: The learning environment was pilot tested in one fifth grade class of 19 students (8 female). After making necessary improvements, a revised version (third iteration) was implemented in one fifth and two sixth grade classes with 63 students in total (32 female) from the same school. During the enactments, students were divided in four groups per class.

Approach: Inquiry-based learning enhanced by gamification elements. Hypothetical scenarios were embedded in the learning environment to personalize the experience and invite students’ reflections and critical discussion.

Learning sequence: The learning environment “Who’s looking at my data? Learning using activity trackers”, consists of three 80-minute lessons; the activity sequence progresses from familiarizing students with activity trackers through experiential activities (Lesson 1) to the exploration of three hypothetical scenarios presented and explored through a gamified approach (Lessons 2 & 3). In Lessons 2 and 3, students were presented with their mission: the director of a company that develops and sells activity trackers is asking for their help to reply to three requests, presented in three scenarios, about the release of students’ personal tracking data. All three scenarios included information on: (a) the type of data being collected, (b) the requesting stakeholder, (c) the purpose of usage, and (d) the perceived benefits for kids. Our aim was to scaffold students in taking a position on each hypothetical scenario; we framed each scenario so as to highlight both the positive and the negative outcomes of conceding data ownership to empower students’ impartial decision. The three scenarios were: (A) a sports company is requesting access to students’ personal tracking data so that they can send them discount coupons and promotions based on their physical activities; (B) a health insurance company is asking for a girl’s heart rate data to decide whether to cover medical expenses related to treating her heart disease, and (C) a children’s gym owner gave each gym member a free activity tracker and is now asking for their permission to post their activity data on a members-only website for a monthly competition, with a free, one-month subscription to the gym as the prize.

Students in each class worked in four groups of six students each (12 groups in total). A progress map, showing start, middle and end points which the students should reach, was used to scaffold students’ inquiry in each scenario activity. Each group had a unique logo chosen at the first lesson, and the progress map was updated with each group’s achievements at several points during the investigation, by moving each group’s logo to the current point. At the start line, students were asked to read the letter sent to the director. At the first station of the progress map, each student had to make an individual decision on what to advise the director. The individual decisions were followed by group discussion leading to an initial written group decision, as follows: “We the (name of the team), advise the director of the company to respond (positively/negatively) to the sports company/insurance company/gym owner, because…” At the second station, students were provided with different data visualizations of the self-tracking data, according to each scenario, to scaffold their initial group decision. At the third and final station students had to declare their final decision, after examining the data visualizations they accumulated from the second station. After each scenario, all groups had to announce their decisions and provide evidence-based arguments. Plenary discussion followed to explore each group’s arguments and explain the unanimity or the diversity of their decision.

Data collection and analysis: Data obtained from video-recordings from three student groups in each class (nine groups in total) during the second and third lesson, and semi-structured interviews obtained by all students before and after the learning interventions. Videos were carefully reviewed following the critical incident technique (Flanagan, 1954) to identify incidents in which students had been engaged in critical discussions about allowing or denying the companies in each of the scenarios to access and use their self-tracking data. During these episodes, we focused on how the learning materials and the resources made available to students had scaffolded or hindered their critical discussion. The interviews were analyzed using two coding schemes, reported in Agesilaou and Kyza (2021). The first coding scheme, which was based on the ABC model (Chi, Jeng, Acker & Bowler, 2018) and on Pangrazio and Selwyn’s (2019) work, was used to identify students’ Affective, Behavioral, and Cognitive states towards personal data privacy practices and to help us understand students’ awareness of the commercial use of their personal digital data. The second coding scheme was used to identify students’ conceptualization of ownership of their personal digital data.

Findings: The analysis of the video-recordings indicated that students were more engaged and motivated to participate in the collaborative discussions during the second design cycle which employed a gamified approach using a series of hypothetical scenarios. The hypothetical scenarios provided the opportunity to approach the matter from multiple angles; since they were neutrally framed they allowed students to express different points of view leading to interesting discussions. In contrast to the implementation of the first design iteration, students were more confident to collaboratively discuss the three scenarios and to provide their experience as users.
During the second design attempt, we guided students to examine the three hypothetical scenarios following the same steps, which were visible in the progress map placed on the blackboard. By the second and third scenario, students already knew how to proceed. This enabled each student team to continue on its own pace and allowed the teacher the opportunity to provide differentiated scaffolding to each team. Allowing students to track their progress, through the progress map, restructured the classroom organization between the teams and the teacher, but also between the members of each team. Even though students knew that there was no prize if they finished their inquiry steps first, they were monitoring the other teams’ progress, and on many occasions asked their teammates to avoid off-task discussions and to contribute to the collaborative work.

The analysis of the interviews indicated that students would feel anger, fear, stress and be offended if someone gained access to their data without their permission. Our analysis revealed that students employ privacy and security mechanisms to their personal digital data to avoid data violations from hackers, hazardous communication with strangers and disclosure of geolocation information. While students’ answers before the learning intervention, as to who might be interested in their personal digital data, were more general (e.g. hackers, strangers, for bullying purposes), after the learning intervention more (but still not all) students became more specific and reported that their data can be used by companies, highlighting the potential of our intervention to positively affect students’ awareness of personal data mining. We also examined students’ conceptualization of ownership of their personal digital data. The majority of students believed that they are the only owners of their data; only few students were aware that when giving permission to a commercial entity to access your data, one also agrees to shared ownership and shared rights about those data.

Implications for design
These results indicate that future efforts should focus on helping students acquire a better understanding of the transfer of data ownership to third parties, along with what happens with the data analysis and profiling of their personal data, in order to develop the necessary competencies to decide which privacy and security practices towards their personal digital data they wish to employ. The analysis of students’ interviews and in-class group discussions indicated that students’ willingness to concede data ownership to third parties is dependent on the nature of the personal data being shared, how they will be used, by whom and for what purposes. These criteria shaped students’ decisions in the three hypothetical scenarios. Students were positive in conceding data ownership to the sports company (9/12 groups) and the insurance company (12/12 groups), but negative to conceding data ownership to the gym owner (9/12 groups). These results highlight that context is important on how students evaluate data privacy and how they perceive the commercial utilization of their personal digital data.

Discussion and conclusions
Design-based research serves twin and complementary goals: While it seeks to validate learning theories in real-world practice, it also contributes to the development of these theories (Hoadley, 2004). The goal of this design-based research effort was to design a learning environment for upper elementary students to support their understanding of the data economy and the commercial exploitation of their personal digital data. Informed by critical theories of data literacy, we designed and implemented a learning environment in order to understand how children and youth can be empowered through scaffolded inquiry experiences to reflect on the use of their own smart, self-tracking devices. Lessons learned from the two design cycles described in this study, show that the following design principles are important in shaping students’ development of critical data literacy: (1) Contextualization is an important factor which can shape students’ understanding of the exploitation of their personal digital data and foster engagement in critical discussions. There are four determinants of students’ willingness to concede data ownership to third parties: the nature of the personal data being shared, how they will be used, by whom and for what purposes. The results of this study suggest that students evaluate shared ownership of personal digital data based on different criteria and this impacts their willingness to disclose or protect personal information, and their willingness to share data ownership with third parties. Future efforts should consider these criteria when designing learning materials aiming to foster critical data literacy. (2) Setting up a scaffolded, personally-motivating, critical inquiry context increases students’ willingness and ability to engage with the ideas. In our study, it was important to support students’ engagement with the, otherwise, abstract ideas of data privacy by connecting experiential activities with individual and group-level reflection. The gamified approach and task-driven activities fueled engagement, transforming the activity-trackers and the scenarios that were linked to them to evocative objects (Turkle, 2007).

Our ongoing work, of which this study is a part, also contributes to theory building about students’ conceptualization of the psychological ownership of digital data, which has been less explored in the literature, and critical data literacy education efforts. Attending to the affective-behavioral-cognitive (ABC) components of the task increased students’ critical discussions. Yet still, students in our study believed they were the only owners...
of their self-tracking data, even in the cases when they agreed to give access to their data to third parties, such as the companies presented in the three hypothetical scenarios. In order for students to develop critical data literacy, they have to gain a deeper critical understanding of the digital infrastructure, and how data flows from the individual devices and applications, to second and third parties managing those data. Such an understanding is necessary for students to realize the subsequent data transfers and data rights to others.

Efforts to design learning materials aiming to foster critical data literacy in formal education have been reported as limited in the literature (Pangrazio & Sefton-Green, 2020; Pangrazio & Selwyn, 2019). Our study is among the first efforts to understand how students can be scaffolded in developing the conceptual and practical competencies needed to gain a deeper understanding of the digital environment and the political economy. Such initiatives are important in contributing to the discussion of how critical data literacy can be theorized in formal educational settings and supported through empirically validated research-based curricula.

References


