COGITATIO

Media and Communication (ISSN: 2183–2439)
2020, Volume 8, Issue 4, Pages 197–207
DOI: 10.17645/mac.v8i4.3407

Article

Digital by Default: Children's Capacity to Understand and Manage Online Data and Privacy

Mariya Stoilova [1,*], Sonia Livingstone [1] and Rishita Nandagiri [2]

[1] Department of Media and Communications, London School of Economics and Political Science, London, WC2A 2AE, UK; E-Mails: m.stoilova@lse.ac.uk (M.S.), s.livingstone@lse.ac.uk (S.L.)
[2] Department of Methodology, London School of Economics and Political Science, London, WC2A 2AE, UK; E-Mail: r.nandagiri@lse.ac.uk

* Corresponding author

Submitted: 30 June 2020 | Accepted: 14 September 2020 | Published: 10 November 2020

Abstract
How do children understand the privacy implications of the contemporary digital environment? This question is pressing as technologies transform children's lives into data which is recorded, tracked, aggregated, analysed and monetized. This article takes a child-centred, qualitative approach to charting the nature and limits of children's understanding of privacy in digital contexts. We conducted focus group interviews with 169 UK children aged 11–16 to explore their understanding of privacy in three distinct digital contexts—interpersonal, institutional and commercial. We find, first, that children primarily conceptualize privacy in relation to interpersonal contexts, conceiving of personal information as something they have agency and control over as regards deciding when and with whom to share it, even if they do not always exercise such control. This leads them to some misapprehensions about how personal data is collected, inferred and used by organizations, be these public institutions such as their schools or commercial businesses. Children's expectation of agency in interpersonal contexts, and their tendency to trust familiar institutions such as their schools, make for a doubly problematic orientation towards data and privacy online in commercial contexts, leading to a mix of frustration, misapprehension and risk. We argue that, since the complexity of the digital environment challenges teachers' capacity to address children's knowledge gaps, businesses, educators, parents and the state must exercise a shared responsibility to create a legible, transparent and privacy-respecting digital environment in which children can exercise genuine choice and agency.

Keywords
children; digital environment; data; datafication; digital by default; media literacy; peer learning; privacy

Issue
This article is part of the issue "Children's Voices on Privacy Management and Data Responsibilization" edited by Ralf De Wolf (Ghent University, Belgium) and Mariek Vanden Abeele (Tilburg University, The Netherlands).

1. Introduction

Children's lives are traditionally conceptualized as part of the private sphere, supposedly protected from the public and commercial spheres by the actions of parents, teachers, and other carefully vetted adults. This is meant to ensure their safety and well-being, allowing them to 'just be children.' But today children are a major source of data in a hugely profitable data marketplace (Zuboff, 2019). Their lives are, arguably, becoming datafied—meaning that their possibilities for action, and the affordances of their lifeworld, are influenced by practices of data processing determined by commercial and political priorities far beyond the control or knowledge of a child (Barassi, 2019; Lupton & Williamson, 2017; Mascheroni, 2018). This raises urgent questions about their privacy (Barassi, 2019; Buitelaar, 2018).

UNICEF (2018) distinguishes several dimensions of privacy affected by digital technologies—physical, communication, informational and decisional privacy.

Physical privacy is violated in situations where the use of tracking, monitoring or live broadcasting technologies can reveal a child's image, activities or location. Threats to communication privacy relate to access to posts, chats and messages by unintended recipients. Violation of information privacy can occur with the collection, storage or processing of children's personal data, especially if this occurs without their understanding or consent. Finally, disruptions of decisional privacy are associated with the restriction of access to useful information or the operation of automated decision-making which limit children's independent decision-making or development.

We are reaching the point, especially in wealthier countries, where children's lives can be called digital-by-default: Even before birth they may have a digital profile generated by their parents, a health record produced by the state, and they may have attracted the interest of commercial actors. Thereafter, much of what they do and what happens to and around them will be digitally recorded, enriching that profile and potentially shaping their life chances. Digital-by-default is increasingly the policy of national governments, resulting in a shift away from (expensive) in-person state provision (for example, for paying taxes, claiming welfare or interacting with authorities) towards online-only services. Until recently, concerns with this policy focused on digital exclusion (Schou & Svejgaard Pors, 2019), but increasingly concerns arise for those who are digitally included—regarding their privacy, and the potential for discriminatory decision-making, as recently resulted from the algorithmic calculation of UK students' A-level results.

The more complex, risky and potentially exploitative the digital environment, and the powerful players that own much of its infrastructure, the greater the public call for stronger data protection regulation, privacy-by-design, data justice and a platform duty of care, as well as for digital literacy education for the public (LSE Truth, Trust and Technology Commission, 2018). Children are widely recognized as being among the most vulnerable, justifying calls for stronger privacy legislation. At the same time, children can benefit from digital literacy education, leading to hopes that they can be taught to be savvy and resilient in a digital world, and even to critically understand the modus operandi of the networked data economy (Buckingham, 2015; Culver & Grizzle, 2017; Livingstone, Stoilova, & Nandagiri, 2020). However, insofar as the digital environment is not designed or regulated to be legible and respectful of children's rights or best interests (Buitelaar, 2018) these hopes may be unrealistic.

Both regulation and education policies rely on assumptions about what children can understand or withstand. Our aim in this article is to examine what children can and do understand about online data and privacy, and how they learn about this, in order to inform the balance between regulatory and educational policies and to protect children's privacy online. Such a holistic approach, involving both regulatory and educational solutions aimed at empowering children and safeguarding their privacy and other rights, is increasingly advocated by a rights approach to privacy (Lievens, Livingstone, McLaughlin, O'Neill, & Verdoodt, 2018; Lupton & Williamson, 2017; UNICEF, 2018).

2. Theorizing Privacy in Relation to the Digital Environment

Westin (1967) explains privacy as the right of individuals, groups or institutions to determine if, when and to what extent information about them is shared with others. In popular discourse also, "privacy is understood almost universally as a matter of controlling one's own data" (Sarikakis & Winter, 2017, p. 1). However, the emphasis on individual control gives rise to many difficulties, not least because social life is relational (Solove, 2015) and subject to context-dependent norms (Nissenbaum, 2010). Mulligan, Koopman, and Doty (2016) argue that privacy is "essentially contested" because it must be persistently and adversarially debated and defended. Certainly, it is being intensely debated and defended in relation to the digital environment (Sarikakis & Winter, 2017), including in relation to children (Kidron, Evans, & Afia, 2018; Livingstone, 2018).

While the origins of the concept of privacy can be traced historically, Laufer and Wolfe (1977) offer a developmental account, tracing its meaning and importance to the early life of the infant, showing how privacy is vital to and inseparable from the individuation of the self during childhood. Consistent with contextual and relational accounts of privacy in legal theory, they offer an account of privacy in which the child's developing efforts to manage information are rooted in their growing capacity to manage social interaction. This capacity is always contextual and "it is not until long after the child has learned that he/she has choice that he/she can control access to himself/herself in a way that makes choice meaningful" (Laufer & Wolfe, 1977, p. 39). Positing a lag between the recognition of choice and the capacity to enact choice is particularly thought-provoking now that children spend so much time in a complex and opaque digital environment that offers them little genuine choice or control, and that substantially disintermediates their parents and other protective adults.

Neither a universalist approach centred on individual control nor a highly contextualist approach to privacy is practical when it comes to protecting children's privacy in the current commercialized digital environment. Hence, we work with a more practical classification that prioritizes three digital contexts, informed by Nissenbaum's (2010) idea of contexts as social spheres. Specifically, we propose that children's lives are primarily framed by three social spheres in which privacy matters: interpersonal (family, peers, community); institutional (such as the school or health service); and commercial (notably purchasing, marketing and data brokering). Building on the work of van der Hof (2016), we also distinguish three

types of data in the digital environment: data given (contributed by individuals about themselves or about others, usually knowingly, although not necessarily intentionally, during their participation online); data traces (which are left, mostly unintentionally and sometimes unknowingly, through online activities and captured via data-tracking technologies such as cookies, web beacons or device/browser fingerprinting, location data and other metadata); and inferred data (derived from analysing data given and data traces, often through the use of algorithms, possibly combined with other data sources, and also referred to as 'profiling').

We suggest that data functions rather differently in each of these three privacy contexts: In interpersonal contexts, data meaningfully given is the prototypical case; in institutional contexts, data is often collected—as children know from their school or medical records—although often not fully analysed (Selwyn, 2019); in commercial contexts, the really valuable data is not that given, nor even that taken, so much as the data that is then inferred, aggregated and used to generate profiles in order to target advertising or for other profitable purposes within the networked data ecology (Lupton & Williamson, 2017; Marwick & boyd, 2014; Stoilova, Nandagiri, & Livingstone, 2019). Given that our stated aim is to examine whether, what and how children understand online data and privacy, Laufer and Wolfe's emphasis on the development primacy for privacy of interpersonal contexts gains a new significance in a digital world in which institutional and commercial actors have far greater access to children's actions as mediated through data processing. Available evidence already suggests that children's knowledge of interpersonal contexts for privacy online exceeds that of other contexts (Barassi, 2019; Kumar et al., 2017; Stoilova et al., 2019). It seems that children largely regard the digital environment as a 'personal space' for self-expression and socializing and that, while children are often concerned about parental intrusion into their privacy, or with the interpersonal risks that arise when personal information circulates among peers without their consent, they have little awareness of future implications of data traces, particularly in relation to a distant future that is hard to predict or to conceive of (Bowler, Acker, Jeng, & Chi, 2017; Murumaa-Mengel, 2015; Pangrazio & Selwyn, 2018). Even by the time they reach adolescence, children have little knowledge of data flows or of infrastructure—they mostly see data as static and fractured, as when located on different platforms (Bowler et al., 2017), which can create a false sense of security.

3. Methodology

We conducted 28 mixed-gender focus groups with children aged 11–16 years from six UK secondary schools, two in London and one each in Essex, the Midlands, Wales and Scotland, selected to represent a mix of achievement and geographical area. The 169 participants (85 girls and 84 boys) were selected by their own schools from among those who volunteered after receiving an information sheet about the project, on the basis of diversity in background, grades and digital skills. The project was approved by LSE's Research Ethics Committee and consent was given by the children and one of their parents. The focus groups lasted 73 minutes on average and were held with three school year groups—aged 11–12 years, 13–14 years and 15–16 years.

We designed and piloted participatory research tools (visuals, games, pen-and-paper tasks, workshop activities) to engage students, using real-life scenarios and exemplar digital experiences. To allow children's understanding to emerge spontaneously, we structured the discussion using a 'ladder approach,' starting with more familiar issues and moving towards greater complexity as regards both the privacy contexts and the types of data we invited children to consider. We first invited children's spontaneous perceptions and practices (e.g., apps selection, checking age restrictions, reading terms and conditions), followed by a game to gauge their familiarity with relevant terminology (e.g., cookies, privacy settings, digital footprint, algorithms). Then we conducted exercises to explore the types of data children share in different contexts, gradually enabling discussion of less thought-of issues relating to data harvesting and profiling. Activities were conducted collectively, some in smaller groups, in order to generate conversation and avoid any perception of being tested. All sessions were recorded, transcribed and analysed using thematic analysis with NVivo.

4. What Do Children Know about Data and Privacy Online?

4.1. Interpersonal Contexts: Using the Familiar as a Model

It was immediately apparent that children find it easier and more obvious to focus on interpersonal aspects of online privacy. This is more familiar and understandable to children and it is also the sphere where they have more agency and control. Children were keen to describe their privacy strategies in terms of the way they handle the data they know they give—the pictures they post online, the links they share, the information they enter when registering for platforms—in order to protect their privacy, relationships and reputation. They told us how they remove unwanted information, untag content, use 'fake' personal data, switch between accounts and platforms and use assorted privacy settings and passwords to protect their devices and data. Children's actions of deciding what to disclose, where and to whom, and of negotiating with others what should be shared online, emphasize how they value individual control, and their nuanced appreciation of context, which results in considerable personalization of their choices and tactics.

In dealing with their interpersonal privacy, children acknowledge that they do not have full control over their

data because of what Marwick and boyd (2014) describe as 'networked privacy,' referring to the public-by-default nature of online communications. Thus, children realize that others could share information about them without permission or beyond the intended purpose or audience: Parents sharing embarrassing pictures with relatives or friends is a frequent example of how children feel their privacy is breached. Data traces and inferred data appear to be much less significant for interpersonal privacy contexts, although children sometimes mention these in relation to how their information will be perceived or used by others—their parents might track their location when they are late home from school, a burglar could see that they are not at home when they check in to a holiday destination, some of their distant friends will figure out that they were not invited to a birthday party.

Perpetrators of privacy risks are also thought of in interpersonal terms—the stalker, the hacker, the bully, the kidnapper, the 'paedo' (Livingstone, 2014), and children's thinking often revolves around 'what's the worst that can happen':

> People could find out where you go. So they could try and find you and wait for you there. (Boy, Year 7, Essex)

> The only thing that worries me is weird folk, like stalkers. (Girl, Year 11, Scotland).

Indeed, interpersonal risks seem to them much more salient than institutional risks, or long-term risks associated with the commercial data ecology, as already pointed out by previous studies (Barassi, 2019; Bowler et al., 2017; Kumar et al., 2017; Livingstone, 2019; Lupton & Williamson, 2017; Selwyn, 2019). Fewer studies have as yet explored children's understanding of institutional or commercial privacy, so we devote more attention to these in what follows. As we seek to show, because children learn first about interpersonal privacy, it appears that they extend interpersonal assumptions to institutional and commercial contexts. Specifically, we observed how they tend to apply attitudes to privacy—along with their analysis of privacy risks and privacy strategies—from the interpersonal context to a context which is quite different from their initial point of reference. Importantly, as we also demonstrate in what follows, drawing on interpersonal notions of privacy leaves children at a disadvantage in an institutional or commercial environment.

4.2. Institutional Privacy: Symbolic Boundaries and The Role of Trust

When asked about privacy online, children rarely think about institutional contexts or the data that their school, doctor, government, or future employer might hold. Similarly, when talking about personal data, children rarely refer to their immunization or dental records, or school academic achievement or attendance records. We found children rather bewildered when we first mentioned such data, as this is neither information they choose to share nor something they could refuse to give. Perhaps because they have very little control of what is collected by institutions, or of how or when this is collected, they do not grasp immediately that, beyond the data they have knowingly given, these data records are also personal data with significant privacy implications. After all, such information is collected about everyone and some of it—their dental records, for instance—may not immediately seem very telling about who they are as a person. Yet, over the course of our conversations, children realized that the handling of such data could have significant repercussions (for example, if it is stolen or used outside its intended purpose) and, thereby, that a range of institutions that they had not previously given much thought to gather considerable amounts of sensitive data about them.

In relation to institutional contexts, children find it easiest to understand that of the school, which they know holds a lot of their personal information—provided by students and parents or collected by schools. Even the youngest children we spoke to (11–12 years) could list data such as their names, photographs, attendance records, fingerprints (in schools that use this for lunch payment), health information and, on reflection, what they eat for lunch. Rarely a cause for concern, this institutionalized data collection is viewed as justified for ensuring children's health, safety, learning or wellbeing, provided its use is limited to the original purpose. As one boy (Midlands, aged 11–12) explained, "they're my school, they're going to keep my data safe." This comment reflects the importance of trust: The negotiation of trust is traditionally emphasized by theories of privacy (Petronio, 2002), although it is noteworthy that children learn to trust in the context of interpersonal relations (Davis & James, 2013; Kumar et al., 2017) and only later extend this to certain institutional or commercial contexts.

Institutional data collection typically occurs within a broader regime of monitoring, supervision and surveillance. Because it is, in effect, part of 'ordinary' practice, it rarely provokes privacy considerations. In other words, children's understanding of the school's data processing is embedded in their everyday social relations with teachers and school administrators, as well as in the implicit sanctioning of such relations by their parents and peers. For example, children expect to be monitored both digitally and offline to ensure their compliance with an established code of behaviour. Children talked of how their searches and the websites they visit on school computers are monitored:

> The teachers will tell us, they're watching what you're doing. (Boy, 15–16 years, Essex)

> If I need to put sexual health clinic or something and then it blocks it, which is annoying. (Girl, 15–16 years, Essex)

This 'privacy licence,' however, is not without limitations. Children expect institutional monitoring to occur within certain physical and symbolic boundaries—on school premises and in relation to educational activities. Or, in relation to health, in the doctor's surgery or other relevant places. Beyond these boundaries, children expect to retain their privacy. The same teacher who can check what children search for online on the school computer cannot follow them on social media or monitor their social life in their 'private time': "Teachers can't talk to students outside, on social media and so on" (boy, 15–16 years, Midlands).

In short, children's trust in their schools gives them confidence that their teachers—and the digital apps they deploy at school (homework apps, learning software, etc.)—store their data securely within the school and would not share it further without their permission. However, when we inquired further, children realized they had little knowledge of how their schools use their data, or whether it is shared with third parties, or stored in ways that could not be hacked or breached. Their willingness to trust the school—albeit without much alternative—and their acceptance that they could hardly challenge the decisions made by the school, sets up a particularly problematic reference framework insofar as children apply this to the even more instrumental relationships that characterize the commercial domain (Steeves & Regan, 2014). We explore this below.

4.3. Commercial Privacy: A One-Dimensional Relationship

Commercial contexts are the least likely to be on children's radar when they think about privacy (Davis & James, 2013). Children are less familiar with how processes operate in these contexts and less able to conceive of their personal implications. Further, in discussing the activities of businesses, some children appeared more able than others to grasp the complex ways in which their data flows online, as well as more aware of the implications arising from the commercialization of personal data. This was partly a matter of age and maturity (Kumar et al., 2017), and also of digital expertise; we also saw hints that critical discussion at school or in the home make a difference to how well children navigate the more complex features of data and privacy online, particularly in relation to commercial contexts.

We found that most children understand that they are targeted with advertising content. They know companies are trying to gain their attention for commercial purposes and are beginning to grasp the mechanics behind content personalization (Davis & James, 2013). For example, most children see connections between the information they are shown and their previous actions—past searches, pages visited, content engaged with. While some understand the functionality behind personalization of commercial content and how cookies support that, more are puzzled:

> Sometimes I get stuff that I've already searched. (Girl, 15–16 years, Scotland)

> I'm not entirely sure what it [cookies] means. But, I think, it's, like, a good thing if you agree though. (Boy, 11–12 years, Essex)

Some children are critical and disapprove of the online business model, but most see it as a necessary compromise in exchange for 'free Internet' or just how things are. Some even suggested that personalized advertising content creates a better online experience, presumably trusting commercial messages. Few children made the 'jump' from giving an account of targeted advertising to recognizing the algorithmic reshaping of the online environment. Nor did most consider how the same principles of personalization might have wider implications, biasing their online experience or differentiating it from that of their peers or others. In short, children tend to miss the 'bigger picture,' as most are not told or taught how such processes might influence their learning, exposure to diversity, choices or decision-making.

Children also struggle with the idea that their online activities produce data traces that permit further inferences about them. Many understand that their actions leave traces online, but what that data is and where it goes is perplexing and made even more complicated by technological innovation, differences across platforms, and non-transparent policies:

> [Talking of a map app]…it will know what trips I'm taking, without me saying. (Girl, 15–16 years, Essex)

> It's when you go onto websites and stuff and you leave traces, like what you looked up. Like a footprint, but it's digital. (Girl, 11–12 years, London)

Even though their search history is not data that they have given voluntarily, it is still related to their actions online, so children have a sense of what these actions are and sometimes use strategies to remove or protect such data and their privacy, for example by using incognito tabs or deleting their history. It is much harder for them to grasp the harvesting of data that is not connected to intentional activities—for example, IP address, device information, browser type, time spent on a page. Most children are surprised to learn that such data is gathered, or that it has value for companies.

Understanding how data flows are merged into a digital footprint is too complex for children, as for most adults. The depth and extensiveness of data profiling within a commercial data ecology is too far removed from their experience. Children have only a rough sense

of being monitored online, and our focus group discussions led them to raise many questions about how, why and for what purpose this occurs. For instance, asking whether their own deleting of their data means that it is removed permanently from the Internet sparked many debates as children struggled to grasp the idea of a growing digital footprint that is durable, searchable and virtually undeletable. Yet, some experiences give them hints that their online activities leave a permanent trace, as one girl explained: "If you deactivate your [Instagram] account, you just log in and it reactivates it….All your posts come back, and people you follow" (girl, 15–16 years, Scotland).

Most experiences, however, teach them that they have little power to manage the commercial environment. With their privacy settings, data protection choices or other design options, children find in practice that they have little choice but to consent to default options. Each app and platform functions differently from the next, privacy settings change when software updates are implemented, and protecting one's privacy becomes like a game of tag. Tricked by deceptive design (Kidron et al., 2018), children tend to assume that providing personal information when asked is mandatory. Incomprehensible terms and conditions that need to be accepted in order to use a service and cookies that are hard even for a diligent adult to disable teach children that exchanging data for access is unavoidable.

Although the news media make children increasingly aware of data breaches and fraud—children were keen to share recent instances in the focus groups—for the most part, their lack of meaningful choices or forms of redress undermines children's agency in the digital environment. It is in this context that we observe their willingness to trust companies. In practice, they have little option if they wish to use an online service, but in their talk they appeared to draw on their familiarity with interpersonal relationships in explaining why they trusted a company with their personal information:

> If you have friends who have it, then…you trust the app. (Girl, 15–16 years, Scotland)

> If it's like a trusted website that I like, I visit often and trust and all. If it's a big corporation like Adidas or Apple for instance. (Boy, 13–14 years, London)

In other examples, children talked of 'the people' at Instagram, or a friend's father in the tech industry, assuming that the company would act with the same values as would someone they know personally. Or, because they themselves feel offended that 'others' collect their 'private' data, they assumed that those others, be they individuals or companies, would feel it improper to keep or share their data. Or, again, they talked as if the privacy tactics, workarounds and deceptions that they use to protect their online privacy from their friends or parents (such as giving a false name or age, searching 'incognito,' or switching devices) would also protect them from online businesses (or, indeed, institutions).

5. Children's Capacity to Learn about Data and Privacy Online

How can children gain a deeper and more critical understanding of their privacy online not only in interpersonal contexts but also in institutional and commercial ones? In a rapidly changing technological environment, digital literacy—broadly, the knowledge that children need in order to act effectively in relation to the digital environment—is a moving target. Children must summon all their resources to keep on top of new developments, devices, functionality, policies and regulations. Formal education is an important source of information for them, but it is only one form of learning—in many cases children are working it out on their own. However, in trying to put the pieces of the puzzle together from diverse sources of information, children acquire fragmented knowledge, including some misconceptions. Insofar as both children's capacities and the practice of digital literacy education face real limitations, regulatory and/or design solutions for the protection of children's privacy in relation to the digital environment will be necessary.

5.1. Learning by Doing: Working It Out

In their engagement with technologies, children take a hands-on approach—trying things out, learning by doing, and quickly moving from one app to another in pursuit of something new and exciting, while speculating amongst themselves as to where the risks are and what the possibilities may be (Livingstone, 2014). Their digital lives are dynamic and so is their approach to learning about privacy. Children sense—or are working out—that everything they do online may be tracked and recorded for whatever purpose by businesses, parents and schools. While children might ask a parent, a sibling, or a knowledgeable friend to help them, many expect to learn on their own by trial and error or by searching online for information when needed. Children are quite confident about their own abilities to navigate Internet-related issues, while asking adults for help is reserved for 'really serious' situations. Children enjoy exploring new opportunities, following up on things they have heard about from friends, checking out new gossip or trends, or following their favourite popular online figures. These practices are fun and informative and a great way to learn actively. They are also a coping mechanism in a rapidly-changing, hard-to-predict environment with few knowledgeable authority figures in children's immediate surroundings:

> I think ourselves are our best teachers because we learn, and we kind of know. (Boy, 15–16 years, Essex)

> It's so new that no one really knows what's going to happen it. No one knows where it's going to go. (Girl, 15–16 years, Essex)

Children also learn from widely debated 'privacy buzz cases.' For example, high-profile privacy breaches (such as Cambridge Analytica) or public discussions of new regulations (such as the European General Data Protection Regulation) are examples that even the youngest children brought up in their discussion of privacy:

> Facebook sold the information of their users to a different company who made things to put people off of voting for someone. (Boy, 11–12 years, Essex)

> Mark Zuckerberg, he's always watching. (Boy, 15–16 years, Essex)

Legal and policy changes are harder to grasp than front-page privacy breaches, but their repercussions attract children's attention as well. Children had noticed that they are asked about cookies on all the sites they visit, that notifications about changes to privacy policies of social media platforms had started to pop up, while their schools had sent letters home asking for consent for data collection. Such developments serve as learning opportunities for children, even though not all can follow the debates or fully understand the issues. Not integrating such 'privacy buzz moments' or new regulatory changes into children's (formal or informal) digital literacy education seems like a wasted opportunity, especially when they are intrigued by the topics that everyone seems to be talking about, and that are affecting their daily online experiences.

However, left on their own, most children do not learn more complex digital skills or engage in the full spectrum of online opportunities (Livingstone, 2019). Data literary is not a competence which is easy to master and such knowledge is hard to come by without a scaffolded learning process. Hence, it is not surprising that, in spite of their active approach to learning, children have many gaps and misconceptions. Terms are misleading—why is it consent if you must agree to use the service? Why is it called deletion if nothing is really gone permanently? Policies are illegible to children because "there is quite weird language on purpose to trip you up" (girl, 13–14 years, London). It is perplexing to them why some apps request information that seems irrelevant to the services they provide, and children find it counterintuitive that companies want to keep data that quickly becomes outdated or that they know to be wrong because they have provided misleading information. At the same time, as we have seen, children are often trusting of companies, expecting them to protect the privacy of their customers, to collect data to improve the user experience and to follow the rules so as not to jeopardize their own reputations.

Trying to make sense of how the data ecology works, children create their own hypotheses, myths and resolutions, drawing on their familiar interpersonal experiences although these may be inappropriate to the circumstances. Notably, children find it hard to imagine why any company would be interested in their data and why this might have privacy implications, when they have 'nothing to hide':

> I just don't think that what the ordinary everyday person does on the Internet is really that interesting to companies and even if they take data, I don't think that anything bad will happen to me. (Girl, 13–14 years, London)

> I don't really do any sensitive stuff on the Internet….Why would somebody want to track me down? (Boy, 11–12 years, London)

> I don't see what they'd get out of it [selling my data], to be honest. (Girl, 15–16 years, Essex)

5.2. Can These Gaps Be Addressed by Media Literacy Education at School?

Formal education is an important source for learning about online privacy, be this as part of the curriculum on media education, computing, citizenship, or elsewhere. In our discussions, children often mentioned learning about different privacy issues—online contact, sharing information, privacy settings, cookies, or geolocation—in class or following teachers' advice on how to avoid online risks. They also acquire practical technical skills in tasks ranging from easier ones like using email to much harder ones like learning a programming language. But how realistic is to expect that all gaps in children's knowledge and skills related to data and privacy online can be addressed in educational settings?

Talking to children revealed many challenges and gaps in the current curriculum. Most of schools' emphasis is on e-safety, including attention to interpersonal privacy and data given, but offering them little understanding of institutional or commercial data practices. We also found that children's knowledge relates predominantly to their current or near future situation, but rarely encompasses the possible long-term consequences of their digital footprint for employment or further education. Yet, there are many things that children want to learn more about, extending beyond their interpersonal experience to encompass also how the Internet works and how their data flows.

Indeed, when we asked them what they want to learn, children quickly assembled a list of questions, many of which we could not ourselves answer with certainty. Children want to know where their data goes, who keeps it and why, for how long their data is stored, its use and with whom it is shared. They are puzzled by the bigger picture, asking about how the Internet works, who controls it and who makes decisions about the consequences of selling personal data. Many of their ques-

tions showed their desire to have more control over their privacy—how to change their data and digital footprint and how to have better privacy without having to stop using social media or other digital resources. Some seemingly naïve questions, like "What do they do with your face when you use facial recognition?" tap important issues about the future of datafication and the dangers arising from the endless surveillance possibilities of governments and corporations. Children are prepared to do the work to gain this knowledge and want schools and parents to step up to teach them about these issues, but they also want companies to make things easier for them to understand on their own. This seems to open up an opportunity for schools, given children's enthusiasm to learn more.

But these issues are not easy to teach about, and this would require further training of educators and updates to the school curriculum. Children and their teachers discussed the difficulties of keeping the curriculum up to date and sufficiently engaging:

> What about the people who still don't know how to send emails or anything like that? Because I still struggle with sending emails. Like, I just still….I can't get my head around it. (Girl, 13–14 years, Wales)

> We just get bored and don't listen. (Girl, 13–14 years, Essex)

> What they're trying to say is just, like, oh yes, don't do this, don't do that, don't do this. When it's, like, basically the whole point of that thing. (Girl, 13–14 years, Scotland)

Differences in the competence of children, even of the same age, can be quite pronounced, and these are likely to increase further as more privileged parents gain more knowledge and can support their children differentially. This can make teachers' task complicated, but also opens up possibilities for encouraging peer learning and makes the role of schools all the more important in improving equity in privacy and data literacy. In some of the schools we visited, we found that concerted efforts to offer a more comprehensive curriculum seem to show positive results, with children at some of the research locations appearing notably more knowledgeable than at others. Yet, even the most competent children struggle with some aspects of datafication, commercialization of personal information or data flows that are simply beyond their comprehension, and in many cases also beyond that of parents and educators.

In spite of the many challenges faced by digital literacy education at present, our research also demonstrates the unique position of schools as institutions tasked simultaneously with educating students and with managing their personal data. The trust that children and parents place in schools, the access that schools have to all children equally, and the fact that children spend years in school, means that schools have a rare opportunity to deploy their own data protection and management practices as a pedagogical strategy extended over real time and to teach children about privacy, the mechanics of data gathering and protection, and the rights and responsibilities associated with sustaining standards of transparency, security, fairness and data justice (Gangadharan & Niklas, 2019). In schools, therefore, the theory and practice of online privacy and data protection could be productively aligned, thereby offering children an example of best practice that would enable them to view the practices of other organizations critically where merited (Stoilova et al., 2019).

Arguably, however, regardless of how good their education is or becomes, children cannot be expected to fully comprehend and manage their data and privacy online in the current and ever-innovating digital environment. Children are trying to engage with an environment that is generally not designed with their interests or capacities in mind and that is fully comprehensible neither to children nor to many adults. Moreover, the design and operation of digital services continue to be largely 'age-blind,' without regard for whether the person in front of the screen is a minor, and to innovate in highly complex ways led by an incentive structure that rarely prioritizes human rights or ethics (Lievens et al., 2018). Hence, there are growing calls for educational efforts to be supported by greater regulation of the technology sector, including for legislation mandating privacy-by-design solutions (Barassi, 2019; Culver & Grizzle, 2017; Kidron et al., 2018; UNICEF, 2018; van der Hof, 2016).

6. Conclusions

The more children's lives become digital-by-default, the more the design and functioning of the digital environment matters, as do children's understanding of and capacity to manage their data and privacy online. Children are involved, one way or another, in all interpersonal, institutional and commercial privacy contexts, each with its own distinctive logic and outcomes. Our child-centred qualitative study of children's understanding of these contexts revealed that children primarily conceptualize privacy, including their own data online, in relation to interpersonal contexts. As expected, children are most familiar with the contexts where they play an active role in how their data is shared, rectified, used and removed. Significantly, they draw on this understanding to generalize about privacy and to guide their data protection tactics in other contexts.

Some aspects of how privacy works in institutional contexts are also familiar, but here children rely on existing regulations and build relationships of trust to manage their privacy. This accords them a fairly passive role within an environment where they are heavily monitored and regulated (Steeves & Regan, 2014) and are accorded little knowledge or choice. Children's expectation of agency, and their tendency to trust familiar institutions, make for

a doubly problematic orientation towards data and privacy online in commercial contexts, leading to a mix of frustration, misapprehension and risk. Finally, children find the commercial domain perplexing and manage to grasp only some aspects of how it operates. Again, they have little choice but to adopt a fairly passive approach to privacy because of the choice architecture (Thaler, Sunstein, & Balz, 2013) of digital systems, which offers the user only superficial alternatives but no real ways to manage their privacy, while still benefiting from the services. This has important implications for digital literacy, media education and for child rights in a digital-by-default age (Lievens et al., 2018).

Struggling to make sense of how the data ecology works, children attempt to learn actively—trying out, searching and figuring things out on their own. Creating their own hypotheses and resolutions as a way of coping with a rapidly changing environment, children sometimes fall into the trap of misconceptions and have many competence gaps, particularly in institutional and commercial contexts. Insofar as education is part of the solution, these challenges, and the continued pace of technological innovation, raise the bar for children's digital literacy, which is the fastest-changing part of media literacy (Livingstone et al., 2020). At present, schools tend to teach a combination of e-safety and computer programming, but attention to the digital economy and its technological and business operations is rarely included in the computer science or media education curricula (Polizzi, 2020). Our findings suggest that children not only need better digital skills to manage their data and privacy online, but they also need a more comprehensive understanding of how the digital environment works—in terms of its technology, political economy, business and governance. This is challenging to teach, both because of its complexity and pace of change, and because the digital infrastructure of modern societies shares the character of all infrastructures—they are routine, taken-for-granted, noticed only when they break down (Lievrouw & Livingstone, 2009). Moreover, what is needed is a flexible educational approach that recognizes differences among children and promotes their understanding of their rights as digital citizens and data subjects. This should provide particularly for vulnerable or disadvantaged children, given the potential for abuses of sensitive data and for discrimination in relation to automated decision-making.

Not only do children need and want to play an active role in decision-making about online participation and privacy protection, but businesses, parents and the state have a shared responsibility to create a legible and transparent online environment where children have real choices and agency. Specifically, the technology industry needs to take greater steps to respect children's rights and well-being, including through supporting privacy-by-design, data justice and a platform duty of care (Lievens, et al., 2018; LSE Truth, Trust and Technology Commission, 2018; Lupton & Williamson, 2017). Also important is the need for stronger data protection regulation and enforcement. As the policy climate shifts to reconsider rebalancing the responsibility for managing privacy in a digital world between provider and consumer, along with redesigning services and developing accessible systems of redress, democratic politics requires that citizens' voices must be heard on their opinions and concerns. This applies to children as much as to adults, as stated in Article 12 of the UN Convention on the Rights of the Child (UN, 1989). At present, in most policy consultations on data and privacy online, the 'data subject' is treated as ageless, and there is little consultation with children or specific regulatory provision for children's data rights and the protection of their privacy (Livingstone, 2018). An exception is the recent introduction of the UK's Age-Appropriate Design Code, part of the 2018 Data Protection Act—itself based on a consultation with children among other stakeholder groups (Information Commissioner's Office, 2019; Revealing Reality, 2019).

In the societal effort to transcend the too-simple binary choice of education or regulation, it is important to hear children's voices, and to recognize their desire to exercise agency but not to face overwhelming risks in relation to the digital environment. While children wish to take responsibility for their own digital lives, this must rest in part on understanding, and in part on the design and operation of the digital environment: if the latter is opaque, highly technical and fast-changing, children's understanding (and that of the adults who support them) will continue to be challenged and their privacy at risk.

Acknowledgments

We thank the UK Information Commissioner's Office for funding this project, accessible at www.myprivacy.uk. We are also grateful to Julian Sefton-Green for the concept of 'digital-by-default.'

Conflict of Interests

The authors declare no conflict of interests.

References

Barassi, V. (2019). Datafied citizens in the age of coerced digital participation. *Sociological Research Online*, *24*(3), 414–429.

Bowler, L., Acker, A., Jeng, W., & Chi, Y. (2017). "It lives all around us": Aspects of data literacy in teen's lives. In S. Erdelez & N. K. Agarwal (Eds.), *Proceedings of the association for information science and technology* (pp. 27–35.) Hoboken, NJ: Wiley.

Buckingham, D. (2015). Defining digital literacy: What do young people need to know about digital media? *Nordic Journal of Digital Literacy*, *10*, 21–35.

Buitelaar, J. (2018). Child's best interest and informational self-determination: What the GDPR can learn from

children's rights. *International Data Privacy Law*, *8*(4), 293–308.

Culver, S., & Grizzle, A. (2017). *Survey on privacy in media and information literacy with youth perspectives*. Paris: UNESCO.

Davis, K., & James, C. (2013). Tweens' conceptions of privacy online: Implications for educators. *Learning, Media and Technology*, *38*, 4–25.

Gangadharan, S., & Niklas, J. (2019). Decentering technology in discourse on discrimination. *Information, Communication & Society*, *22*(7), 882–899.

Information Commissioner's Office. (2019). *Consultation on age-appropriate design: Summary of responses*. Wilmslow: ICO. Retrieved from https://ico.org.uk/media/about-the-ico/consultations/aadc/2616996/summary-of-responses.pdf

Kidron, B., Evans, A., & Afia, J. (2018). *Disrupted childhood*. London: 5 Right Foundation. Retrieved from https://5rightsfoundation.com/uploads/5rights-disrupted-childhood-digital-version.pdf

Kumar, P., Naik, S., Devkar, U., Chetty, M., Clegg, T., & Vitak, J. (2017). No telling passcodes out because they're private. *Proceedings of the ACM on Human–Computer Interaction*, *1*, 1–21.

Laufer, R., & Wolfe, M. (1977). Privacy as a concept and a social Issue: A multidimensional developmental theory. *Journal of Social Issues*, *33*(3), 22–42.

Lievens, E., Livingstone, S., McLaughlin, S., O'Neill, B., & Verdoodt, V. (2018). Children's rights and digital technologies. In T. Liefaard & U. Kilkelly (Eds.), *International children's rights law* (pp. 1–27). Berlin: Springer.

Lievrouw, L., & Livingstone, S. (2009). Introduction. In L. Lievrouw & S. Livingstone (Eds.), *New media: Sage benchmarks in communication* (pp. xxi–xl). London: Sage.

Livingstone, S. (2014). Developing social media literacy: How children learn to interpret risky opportunities on social network sites. *Communications: The European Journal of Communication Research*, *39*(3), 283–303.

Livingstone, S. (2018). Children: A special case for privacy? *InterMedia*, *46*(2), 18–23.

Livingstone, S. (2019). Are the kids alright? *Intermedia*, *47*(3), 10–14.

Livingstone, S., Stoilova, M., & Nandagiri, R. (2020). Data and privacy literacy: The role of the school in educating children in a datafied society. In D. Frau-Meigs (Ed.), *Handbook on media education research* (pp. 413–425). London: Wiley Blackwell.

LSE Truth, Trust and Technology Commission. (2018). *Tackling the information crisis: A policy framework for media system resilience*. London: London School of Economics and Political Science. Retrieved from http://www.lse.ac.uk/media-and-communications/truth-trust-and-technology-commission/The-report

Lupton, D., & Williamson, B. (2017). The datafied child. *New Media & Society*, *19*(5), 780–794.

Marwick, A. E., & boyd, d. (2014). Networked privacy: How teenagers negotiate context in social media. *New Media & Society*, *16*(7), 1051–1067.

Mascheroni, G. (2018). Researching datafied children as data citizens. *Journal of Children and Media*, *12*(4), 517–523.

Mulligan, D. K., Koopman, C., & Doty, N. (2016). Privacy is an essentially contested concept: A multi-dimensional analytic for mapping privacy. *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences, 374*(2083), 1–17.

Murumaa-Mengel, M. (2015). Drawing the threat: A study on perceptions of the online pervert among Estonian high school students. *Young*, *23*, 1–18.

Nissenbaum, H. (2010). *Privacy in context: Technology, policy, and the integrity of social life*. Stanford, CA: Stanford University Press.

Pangrazio, L., & Selwyn, N. (2018). 'It's not like it's life or death or whatever': Young people's understandings of social media data. *Social Media and Society*, *4*(3), 1–9.

Petronio, S. (2002). *Boundaries of privacy: Dialects of disclosure*. New York, NY: SUNY Press.

Polizzi, G. (2020). Digital literacy and the national curriculum for England. *Computers & Education*, *152*, 1–13.

Revealing Reality. (2019). *Towards a better digital future*. Wilmslow: Information Commissioner's Office. Retrieved from https://ico.org.uk/media/about-the-ico/consultations/2614763/ico-rr-report-0703.pdf

Sarikakis, K., & Winter, L. (2017). Social media users' legal consciousness about privacy. *Social Media & Society*, *3*(1), 1–14.

Schou, J., & Svejgaard Pors, A. (2019). Digital by default? A qualitative study of exclusion in digitalised welfare. *Social Policy Administration*, *53*, 464–477.

Selwyn, N. (2019). What's the problem with learning analytics? *Journal of Learning Analytics*, *6*(3), 11–19.

Solove, D. J. (2015). The meaning and value of privacy. In B. Roessler & D. Mokrosinska (Eds.), *Social dimensions of privacy: Interdisciplinary perspectives* (pp. 71–81). Cambridge: Cambridge University Press.

Steeves, V., & Regan, P. (2014). Young people online and the social value of privacy. *Journal of Information, Communication and Ethics in Society*, *12*(4), 298–313.

Stoilova, M., Nandagiri, R., & Livingstone, S. (2019). Children's understanding of personal data and privacy online: A systematic evidence mapping. *Information, Communication and Society*. Advance online publication. https://doi.org/10.1080/1369118X.2019.1657164

Thaler, R., Sunstein, C., & Balz, J. (2013). Choice architecture. In E. Shafir (Ed.), *The behavioral foundations of public policy* (pp. 428–439). Princeton, NJ: Princeton University Press.

UN. (1989). Convention on the rights of the child. New York, NY: UN. Retrieved from https://www.unhcr.org/uk/4aa76b319.pdf

UNICEF. (2018). *Children's online privacy and freedom of*

expression. New York, NY: UNICEF.
van der Hof, S. (2016). I agree, or do I? A rights-based analysis of the law on children's consent in the digital world. *Wisconsin International Law Journal*, *34*(2), 409–445.
Westin, A. (1967). *Privacy and freedom*. New York, NY: Atheneum.
Zuboff, S. (2019). *The age of surveillance capitalism*. London: Profile Books.

About the Authors

Mariya Stoilova is a Postdoctoral Researcher at the Department of Media and Communications, London School of Economics and Political Science. Her work falls at the intersection of child rights and digital technology, focusing particularly on the opportunities and risks of digital media use in the everyday lives of children and young people, data and privacy online, digital skills, and pathways to harm and well-being. Mariya's work incorporates multi-method evidence generation and cross-national comparative analyses. For projects and publications see here: https://www.lse.ac.uk/media-and-communications/people/research-staff/mariya-stoilova

Sonia Livingstone (FBA, OBE) is a Professor in the Department of Media and Communications, London School of Economics and Political Science. Her 20 books include *Parenting for a Digital Future: How Hopes and Fears about Technology Shape Children's Lives*. She directs the projects Children's Data and Privacy Online, and Global Kids Online (with UNICEF) and has advised the European Commission, European Parliament, Council of Europe, ITU, OECD and others on children's risks and rights in a digital age. See www.sonialivingstone.net

Rishita Nandagiri was a Research Assistant on the Children's Data and Privacy Online project. Her research focuses on gender and reproduction in low-and-middle-income countries. She is currently an ESRC Postdoctoral Fellow, Department of Methodology, London School of Economics.