

*Ethics in educational technology research: informing participants on data sharing risks*

**Marc Beardsley, Patricia Santos, Davinia Hernández-Leo, and Konstantinos Michos**

Department of Information and Communication Technologies, Universitat Pompeu Fabra, Barcelona

*Marc Beardsley is a PhD candidate in the ICT Department of Universitat Pompeu Fabra, Barcelona. His research interests lie in the fields of Neuroeducation, Learning Design, Lifelong Learning.*

*Patricia Santos is a PhD and senior researcher in the ICT Department of Universitat Pompeu Fabra, Barcelona. Her interests include understanding how to use technologies for Learning Design, Situated Learning and Smart Learning Environments.*

*Davinia Hernández-Leo is Associate Professor and Serra Hunter Fellow at the ICT Department of Universitat Pompeu Fabra, Barcelona. Her interests include transversal topics in technology-enhanced learning research, technologies for learning design, collaborative learning, analytics and architectures for educational environments.*

*Konstantinos Michos is a PhD candidate in the ICT Department of Universitat Pompeu Fabra, Barcelona. His research interests lie in the fields of Learning Design, Learning Analytics and Computer Supported Collaborative Learning.*

Address for correspondence: Mr. Marc Beardsley, Universitat Pompeu Fabra, Roc Boronat 138, 08018 Barcelona. E-mail: [marc.beardsley@upf.edu](mailto:marc.beardsley@upf.edu)

**Abstract**

Participants in educational technology research regularly share personal data which carries with it risks. Informing participants of these data sharing risks is often only done so through text contained within a consent form. However, conceptualizations of data sharing risks and knowledge of responsible data management practices among teachers and learners may be impoverished – limiting the effectiveness of a consent form in communicating such risks in a manner that adequately supports participants in making informed decisions about sharing their data. At two high schools participating in an educational research project involving the use of technology in the classroom, we investigate teacher and student conceptions of data sharing risks and knowledge of responsible data management practices; and

introduce a communication approach that attempts to better inform educational technology research participants of such risks. Results of this study suggest that most teachers have not received formal training related to responsibly managing data; and both teachers and students see the need for such training as they come to realize that their understanding of responsible data management is underdeveloped. Thus, efforts beyond solely explaining data sharing risks in an informed consent form may be needed in educational technology research to facilitate ethical self-determination.

## **Keywords**

Ethics, data risks, data sharing, teachers, learners, informed consent, educational technologies, learning analytics

## **Practitioner notes**

What is already known about this topic

- Learning technology research involves the collection of personal data from teachers and learners.
- Sharing personal data carries with it risks.
- Written consent forms are the most commonly used ethical procedure for informing participants of data sharing risks.

What this paper adds

- A survey study in two schools about the prior knowledge of teachers and learners regarding data management and data sharing risks.
- A communication approach aimed at increasing educational technology research participant understanding of data sharing risks.
- Results suggesting that explicit information and training about data sharing risks is needed for teachers and learners to understand ethically the implications of sharing data in educational research contexts.

Implications for practice and/or policy

- Efforts beyond solely explaining data sharing risks in an informed consent form would increase the ethical commitment in responsible educational technology research.
- The communication approach introduced can help guide the creation of materials used to inform research participants.
- A video, as an example of material based on the communication approach, can improve participant conceptualizations of data sharing risks.

## **Introduction**

Educational technology research often requires that both teacher and learner participants consent to sharing personal data as the technology used to support teaching and learning facilitates the collection of data about participants, data generated directly by participants and data generated by observations of participants (Ferguson, 2012; Drachler & Greller, 2016). This sharing of personal data carries with it risks. Risks that can lead to reputational damage, financial loss, and the (ongoing) manipulation of one's behaviour (Shao et al., 2017). It can be argued that advancements in the field of educational technology research have altered the amount of risk that participants are facing. As examples, the greater accessibility of technologies enabling the collection and integration of learner's physiological data begins to blur the lines between educational technology data and biomedical data; the wider adoption of data aggregation and machine learning techniques can alter the threshold for what can be considered low-risk data; and the increased capability to collect data from activities that extend beyond classrooms can provide greater insights into the private lives of participants (Slade & Prinsloo, 2015; Lohr, 2012).

Despite the technological and methodological advances in the field, few publications have looked into the ethical implications of the changing risk exposure in educational technology research. In a structured literature review, completed by conducting keyword searches ("ethics" AND "educational technology", "ethics" AND "learning analytics") in SCOPUS we found 66 related articles between 2002-2018 (see Tables A1-A2 in Appendix A). Out of the 66 articles, 30 articles were related to data privacy, data management by teachers and student and only one article addressed the topic of informed consent (Prinsloo & Slade, 2015). In this article, the authors argue that users usually do not read terms and conditions and privacy notices and may lack background knowledge to understand informed consent. In comparison, a similar search involving informed consent and bioethics, for a single calendar year (2018), resulted in 24 relevant articles (see Table A3 in Appendix A).

Yet, the most common method of informing participants in educational technology research of data sharing risks is via written text in an informed consent form. However, empirical studies in clinical settings have demonstrated that participants often do not sufficiently understand informed consent (Mandava et al., 2012; Schenker et al., 2011). Research in reading comprehension often attributes a lack of comprehension to levels of relevant background knowledge. In a study involving undergraduate students, Kendeou and van den Broek (2007) found that "readers with misconceptions generated more incorrect inferences and fewer correct inferences than did readers without misconceptions." Thus, if the prior knowledge and resulting conceptualisations of data sharing risks among teachers and learners are deemed to be lacking, the effectiveness of a consent form in informing participants of risks may be inadequate.

This leads to our research questions which are aimed at gauging the background knowledge of educational technology research participants in relation to data sharing risks: How are teachers and students currently conceptualizing data sharing risks? Do such conceptualizations suggest impoverished understandings of data sharing risks in educational contexts?

#### *Prior knowledge deficits*

Glenberg (2011) writes that “understanding a situation or a text means that the understanding can be used to guide effective action... whether one is understanding situations, dialogue, or text.” For participants to be able to take effective action such as deciding on whether to participate in a study and assume the corresponding risks, they must understand the situation. Further, comprehension through reading “involves the construction of a coherent mental representation of the text in the reader’s memory” and includes “textual information and associated background knowledge interconnected via semantic relations (e.g., causal, referential, and spatial relations)” (Kendeou et al., 2014). For an individual to attain a “good comprehension of textual material, we generally mean that they have successfully integrated the information from text with their existing knowledge” (Lipson, 1982).

This existing knowledge can both facilitate and interfere with comprehension (Kendeou & O’Brien, 2015) as faulty prior knowledge affects the accuracy of what is retained from a text (Lipson, 1982). Such findings have been demonstrated in studies involving both children and adults (Bransford & Johnson, 1972; Glenberg, 2011; Kendeou et al., 2013; Kendeou & Van Den Broek, 2007; Lipson, 1982). If the knowledge base of readers is “rich and correct” then a concise text can efficiently build understanding on top of that. However, “if the knowledge base contains inaccurate information or misconceptions, then the text must contain information that addresses that in a manner that decreases its future influence” (Kendeou & O’Brien, 2015). Alternatively, if the knowledge base is “accurate but weak” then the text should be “written in a manner that reduces a reader’s reliance on their knowledge and allows for increased reliance on the text” (Kendeou & O’Brien, 2015). In a randomized trial evaluating the recall of consent information, Campbell et al. (2004) found that reading comprehension ability “was the overriding factor that predicted the degree to which persons could restate what they learned from these presentations of informed consent information.”

#### *Ethical obligations*

One can argue that educational researchers have an ethical obligation that extends beyond their legal obligation to minimize the potential negative footprint and maximize the social benefits resulting from their work. Whereas the legal obligation requires evidence that a participant has been made aware of and has signed off on the potential risks; the ethical obligation requires that participants have adequately understood the risks (Drachsler & Greller, 2016). Fulfilling the ethical obligation, as expressed by

social researchers participating in interviews and focus groups, is “complex because it is difficult to determine the participant’s level of understanding beforehand” (Parsons et al., 2016) – and, as previously stated, research in reading comprehension suggests that readers “... who lack background knowledge are likely to construct impoverished representations of the texts they read and, as a result, fail to grasp their meaning” (Kendeou et al., 2014).

In reference to background knowledge on personal data, Eggers, Hamill and Ali (2013) have written that, “Citizens lack a clear picture of what is being collected about them, by whom, or to what end.” This blurred picture of personal data is often attributed to a combination of fast evolving technology and analysis techniques such as natural-language processing, pattern recognition and machine learning with the expanding contexts in which technology is being integrated into our lives (Slade & Prinsloo, 2015; Lohr, 2012). It is difficult to keep up with the breadth of what is being collected and the risks associated with the collection of such data both on its own and in aggregate (Nissenbaum, 2011; Drachler & Greller, 2016). As Lohr (2012) states, “Data is not only becoming more available but also more understandable to computers.”

#### *Enhanced approaches to informed consent*

Miller and Wertheimer (2011) argue that “subjects often do not adequately comprehend what they are consenting to” which undermines the validity of informed consent. Hence, in contexts where participants lack background knowledge possibly leading to impoverished understandings of the risks, participants should be afforded the opportunity to acquire the necessary background knowledge to understand such risks prior to making the decision on whether to participate in the research or not. Enhanced consent forms, multimedia interventions, extended discussions, and test/feedback techniques have been shown to improve comprehension in informed consent in the biomedical field (Schenker et al, 2011). Additionally, specific formative efforts including whole class sessions taken to inform educational participants have been described (David et al., 2001; Mayne et al., 2016; Parsons et al., 2016; Moore et al., 2018). However, as educational technology research is often of a low-risk nature, a determination needs to be made as to what level of enhancement is needed. Enhanced consent forms, in which steps have been taken to improve readability, along with enhanced discussions were found to be the most effective in improving understanding (Nishimura et al., 2013). Enhanced consent forms may also align with research in the field of reading comprehension that show that the addition of causal explanations and simple refutations can reduce the impact of deficient prior knowledge on reading comprehension (Kendeou et al., 2013). The present study looks to investigate teacher and learner conceptions of data sharing risks and their data management practices.

## **Approach**

### *Context of Study*

Survey research was conducted to examine teacher conceptions of responsible data management and student perceptions of data sharing risks. This study was completed within the context of a research project, *Communities of Teaching as a data-informed design science and contextualized practice* (CoT), which involves the use of technology in the classroom with educational purposes. Data includes the designs of learning activities, self-assessments and perceptions (e.g. using questionnaires), and actions tracked by a community platform, and learning design and enactment tools. For example, one of the tools being used is the PyramidApp, a web tool that orchestrates collaboration scripts with pyramidally-incremental groups of learners across activities designed by teachers (Manathunga & Hernández-Leo, 2018). The objective of the project is to study how this data can be used to support learning design community awareness (Michos & Hernández-Leo, 2018) and teacher-led inquiry into their student's learning (Michos, Hernández-Leo, & Albó, 2018) so teachers can iteratively improve their teaching practices. As teachers and students were to be exchanging data, a need arose to ensure that teachers and students were able to responsibly share and manage data.

#### *Approach for Communicating Data Sharing Risks in Research*

David, Edwards and Alldred (2001) suggest that researchers should provide research participants the information that is “adequate to their competencies in order for them to make informed decisions.” To guide the creation of materials used to support research participants in making informed decisions, an *approach for communicating data sharing risks in research* grounded in ethical theories of autonomy and reflecting principles from both social contract theory (Martin, 2016) and the fair transaction model of informed consent (Miller & Wertheimer, 2011) was conceived. The purpose of defining a communication approach was to lay the groundwork for a conceptual framework through which the benefits and harms of sharing private data could be framed and communicated to potential educational technology research participants in an understandable manner.

Ethical theories of autonomy relate to the conditions necessary for self-determination of an individual person (Delany, 2008). Obtaining informed consent refers to communication between researchers and participants in research; and the elements of this communication are guided by law and the moral obligation of respecting the self-determination principle. Furthermore, ethical theories of autonomy align with recent changes in the law (e.g., General Data Protection Regulation (EU) 2016/679) (Wachter, 2018; Rouvroy & Poullet, 2009). Social contract theory advances privacy as “a mutually beneficial agreement within a community about sharing and using information” and attaches the notion of privacy “primarily to the relationship rather than to a piece of data or location” (Martin, 2016). Although social contract theory has been proposed in the context of business ethics, its principles are relevant to educational technology as the field incorporates greater tracking of learners and educators over

extended periods of time. Martin (2016) elaborates that a key tenet of the theory is in directing the focus toward agreements around privacy expectations which shifts the responsibility from adequate notification to the responsibility of maintaining a mutually beneficial and sustainable solution. Similarly, the fair transaction model of informed consent states that “valid consent transactions should prove fair to both parties in light of their preferences” and that researchers should “treat prospective subjects fairly in light of the key values of personal sovereignty (reflecting its negative and positive dimensions) and well-being” (Miller & Wertheimer, 2011). The emphasis on the formation and maintenance of fair, mutually beneficial and sustainable data sharing relationships aligns with the direction of educational technology research, such as learning analytics, which regularly involves ongoing rather than discrete sharing of data.

### *An Approach for Communicating Data Sharing Risks in Research*

The *approach for communicating data sharing risks in research* relates the collection and use of participant data to that of a transaction – with the sharing of data and its risks evaluated and communicated in terms of a net benefit that represents the data sharing relationship to the individual. The relative value of the risks and benefits are determined by the individual. Thus, informed consent involves presenting participants with not only potential risks and harms but also the benefits of participation – and presenting them in a manner that facilitates a net benefit determination by the participant. This premise is consistent with the core principle of justice within the *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans (CIHR, 2014)* in which researchers are obliged to treat people fairly while considering the distribution of the benefits and burdens of research participation. The policy statement must be complied with by individuals and institutions in order to receive funding from federal agencies in Canada including for research related to engineering, social sciences and humanities.

Five principles underlie the communication approach and aim to rationalize defining the duty of the researcher as one of transparently presenting the immediate and ongoing risks and benefits of the relationship to the participant. These five principles are:

1. Personal data holds value.
2. Data collection is a transaction.
3. An ethical transaction implies a fair exchange between parties.
4. A burden exists on the powerful party to ensure fairness of the transaction.
5. Fairness should be explicit and maintained over time.

#### *1) Personal data holds value*

The value of personal data from an economic perspective is apparent “in successful innovative business models—for example, social networking sites, search engines, behavioural advertising companies, and so on—that thrive on personal data” (Purtova, 2015). Eggers, Hamill and Ali (2013) argue that the economic value is demonstrated when we make use of services such as Facebook, Twitter, or Google, as we pay for the

privilege of using these services “by divulging personal information.” Thus, personal data can be considered “a medium of exchange, something that can be ‘cashed out’ for goods and services, or used to pay debt or to store value for future use.” In research, the value is realized by the generation of knowledge achieved by researchers. Further, when research involves the use of commercial tools, there can also be an associated economic value for the tool providers.

## *2) Data collection is a transaction*

Pardo and Siemens (2014) write that, “Whenever data are collected from users, a transaction takes place. A business or institution obtains something valuable from a user.” Purtova (2015) claims that there is debate on whether personal data should be a legal property right, however, there is less debate on whether personal data is an economic property right as “economic property rights signify one’s de facto ability to enjoy a resource and exclude others from that resource.”

## *3) An ethical transaction implies a fair exchange between parties*

Gundlach and Murphy (1993) offer that, “The notion of equity or fairness is widely recognized as essential for mutually satisfying exchanges” and such mutually beneficial exchanges are what underlie social contract theory (Martin, 2016).

## *4) A burden exists on the powerful party to ensure fairness of the transaction*

In referring to the use of learning analytics, Slade and Prinsloo (2013) write that there exists an “inherently unequal power relations in the use of data generated by students in their learning journey.” It can often be inferred that one party has a greater ability to influence the other outside of a study context as a considerable amount of educational technology research is completed within student-teacher or teacher-institution contexts (David et al., 2001; Gallagher et al., 2010; Drachsler & Greller, 2016). In the context of research, Hughes (2014) argues that “A transaction is exploitative if one party to the transaction does not benefit as much as fairness requires.”

## *5) Fairness should be explicit and maintained over time*

As Miller and Wertheimer (2011) suggest, a lot of the research involves a “relational” rather than “discrete” exchange of data. A “*relational* exchange involves transactions linked together over an extended time frame. These exchanges trace back to previous interactions and reflect an ongoing process.” Therefore, it is important to take into consideration the ongoing benefits and risks involved in the data sharing relationship. In reference to the laws governing the processing of personal data, Cormack (2016) describes how balancing tests are used to “... demonstrate that continued processing generates sufficient benefit to balance the decreased benefit and increased risk to the individual’s rights and interests.”

## *Support Materials*

Flory & Emanuel (2004), Schenker et al. (2011), and Nishimura et al. (2013) describe efforts of presenting risks to participants beyond standard consent forms in systematic reviews of innovative approaches used to achieve informed consent in biomedical research. The initial review by Flory & Emanuel (2004) found limited improvements to participant understanding with multimedia interventions, however the latter review suggests that multimedia interventions have positive effects (Nishimura et al., 2013). As part of the CoT Project, a video was produced that was shown to teacher and student participants. The video (<https://bit.ly/2Fkl8it>) script was constructed using the proposed communication approach. It presented both risks and benefits and asked participants to weigh the harms and benefits of participation prior to making their decision on participation. Further, separate workshops for teachers and students were held at both of the participating schools. Content for the workshops was developed using the communication approach and both the teacher workshop (see Table 1) and student workshop (see Table 2) involved the showing of the CoT video.

*Table 1. Teacher workshop training activities*

<i>Activity</i>	<i>Description</i>
Introduction	Introduction to the CoT project
Pre-questionnaire	Pre-questionnaire to understand the prior knowledge of teachers about responsible data management
CoT video	CoT video about data sharing
Learning design methodologies	Explanation of learning design methodologies
Conceptualization	Hands on activity in ILDE. Conceptualization of a learning scenario
Post-questionnaire	Post-questionnaire to understand the impact of the workshop on teachers' understanding of responsible data management

*Table 2. Student workshop training activities*

<i>Activity</i>	<i>Description</i>
Introduction	Introduction to the data privacy workshop
Pre-questionnaire	Pre-questionnaire about student perception of data privacy
Personal data	Activity about types of personal data
Data benefits	Examples of data sharing benefits
Data risks	Examples of data sharing benefits

Minimizing data sharing risks	Activity about minimizing data sharing risks
CoT video	CoT video about data sharing
Final summarizing activity	Summary activity about data privacy
Post-questionnaire	Post-questionnaire about student perceptions of data privacy

---

## Methodology

### *Instruments and research method*

Similar to other survey studies (Wang & Heffernan, 2010; Toprak et al., 2010; Arnold & Sclater, 2017) which evaluated both teacher and student perceptions of data privacy in an educational technology context, we conducted a survey study to investigate teacher and student conceptions of responsible data management and data sharing risks. The survey study corresponded with the formative workshops for teachers and students. Two questionnaires (pre and post) including open and closed questions, were prepared for teachers. The pre-questionnaire was distributed at the beginning of the workshop and aimed to investigate teacher prior knowledge and daily behaviours related to the management of student data. The post-questionnaire aimed to understand if the CoT workshop resulted in teachers thinking differently about their data management practices in a school context. A pre-questionnaire was also prepared for students. The student pre and post-questionnaires included open and closed questions and were designed to investigate both the conceptions of students related to data sharing risks and their daily behaviours related to data sharing outside of an educational context. Descriptive statistics have been used to report results of the closed questions in the questionnaires. An inductive thematic analysis based on our research questions (see Table 3) has been used to report the results of the open questions.

*Table 3. Questions used in the thematic analysis*

---

<i>How are <b>teachers</b> currently conceptualizing data sharing risks?</i>	<i>How are <b>students</b> currently conceptualizing data sharing risks?</i>
--	--

---

What do teachers think they know about responsible data management (e.g. collection, analysis, sharing)?	What risks do teachers see in sharing their data?	What do students think they know about responsible data management (e.g. collection, analysis, sharing)?	What risks do students see in sharing their data?
What are the data management practices of teachers?	What risks do teachers see in sharing the data of their students?	What are the data sharing practices of students?	
See Appendix B Table B1	See Appendix B Table B2 and Table B3	See Appendix C Table C1	See Appendix C Table C2 and Table C3

### *Participants*

A total of 31 teachers and 104 students from a public high school and a semi-private vocational school in Spain participated in the study. Teacher ages ranged between 20-60 and teaching experience between 1-25 years. Main subject areas taught by participating teachers included Science, Maths, Social Science, Languages and Art/Music. Student ages ranged between 16-18, and students were either in the first or second level of upper secondary education (*bachillerato* in Spanish).

## **Results**

### *Teacher Prior Conceptions and Practices*

A pre-questionnaire was completed by 25 teachers at the beginning of the formative CoT workshop. Results are related to our first research question about teacher conceptualizations of data sharing risks and responsible management of student data.

***The results of the questionnaires suggest that the majority of teachers (68%) think they have a reasonable to very high level of knowledge about responsibly managing student data (see Table B1 in Appendix B). However, none of the participating teachers (0%) have received formal training on responsible data management:*** In rating their knowledge of responsible data management in education, a third of the teachers (32%) think they have very little prior knowledge about data privacy topics; 28% believe they have a reasonable amount of knowledge; and 40% rate their knowledge as good or very good. In rating the importance of knowing responsible data management practices, 56% of respondents think it is very important for teachers to receive formal training in responsible data management; 28% believe it to be important; while 8% think it is reasonably important; and 8% not important. Yet, none of the

teachers reported receiving formal training (e.g. attending a course or workshop conducted by an expert) in data management practices in education and have mainly acquired the knowledge on their own.

***Most teachers (64%) believe they are managing sensitive data with the vast majority (80%) expressing high levels of confidence in their data management practices (see Table B1 in Appendix B). Yet, many are unaware of data risk minimization strategies as evidenced by the lack of formal minimization strategies used and elaborated on:***

Most teachers (92%) believe they are collecting ‘sensitive data’ from students while 4% are not sure and 4% do not believe the data they collect is sensitive. In rating their level of confidence in responsibly managing student data, the vast majority (80%) selected high (44%) or very high (36%); whereas only 12% rated their confidence level as reasonable and 8% as low. Yet, when asked to describe the data risk minimization strategies used, 28% of respondents replied that they did not follow a specific risk minimization strategy; 64% replied that they followed strategies they had created on their own; and 8% did not reply to the question. Further, when asked to elaborate on specific data management strategies they are aware of with regard to protecting student data, 64% of respondents referred to the importance of saving data in a protected repository (with restricted access); 8% mentioned the process of consent; 8% mentioned the encryption of data; 8% replied with the name of the current data privacy law; and 12% offer a specific strategy.

***Teachers essentially only mention the use of restricted-access repositories as a data management strategy; and identify private data becoming public as the risk of collecting student data (see Table B2 in Appendix B):*** Almost all teachers (96%) responded that the main risk of sharing personal data is that private data becomes publicly accessible and to minimize this risk the data collected should be saved in repositories with private access – only accessible by designated professionals from the school. However, 36% of respondents were not sure if they followed adequate protocols with regard to the saving and protection of student data.

***In general, (88%) teachers believe that an inadequate data management strategy carries with it risks with problematic consequences (see Table B2 in Appendix B). However, when asked to elaborate and describe specific consequences of poor data management, their answers lacked specificity and mainly fell into two broad categories: lack of trust (affecting reputations and relationships) and legal problems:*** When asked to elaborate on the possible consequences of inadequate data management, 40% of teachers refer to legal punishments (but none of them know the specific legal consequences), 12% do not know the possible consequences. In their comments, teachers also describe the potential effects on levels of trust as they relate to a teacher’s professional reputation and the relationships among students and between teachers and students: *"Lack of trust in the teachers’ professionalism ...Personal conflicts with and*

*between students"; "... students can lose confidence in their privacy"; "Possible work responsibilities or even penalties + reputational risks".*

#### *Teacher Post-Conceptions*

After participating in the CoT workshops, 31 teachers from two educational centers responded to the post-questionnaires to better understand how their conceptualizations of responsible data management (e.g. collection, analysis, and sharing) changed. Of 31 teachers who responded in the post-questionnaire, 25 responded to the pre-questionnaire. This discrepancy was due to six teachers not being able to attend the opening of the workshop in which the pre-questionnaire was given. Results are related to our second research question about impoverished understandings of data sharing risks in educational contexts.

***Following the session all 31 teachers (100%) recognized the importance of receiving formal training to understand how to responsibly manage student data (see Table B3 in Appendix B). Furthermore, participants commented that they were able to think differently about the topic and came to realise that previously they had not reflected sufficiently about the topic nor known enough about appropriate protocols:*** All respondents (100%) believe that after the workshop they are more aware of data sharing risks and strategies to minimize such risks. Prior to the workshop, 84% of teachers had responded that it is important/very important to be informed on data management strategies, after the workshop this had slightly increased to the 89,4% of teachers. In detail, 78,9% think it is very important (this number has increased were initially was a 56% before the workshop) to be informed on data management strategies; 10,5% think it is important; 5,3% says it is not important; and 5,3% believe it is not very important. An analysis of the elaborated comments of teachers, reveals how teachers have become more conscious of the importance of establishing clear data management strategies after the workshop: *"I had not thought much about this topic but with the change in data protection law (referring to the new European GDPR law), I believe it is very important that we learn and become more aware of the data we collect, how we make use of them, as well as the risks that exist in relation to third parties."*; *"It is an aspect on which I am careful, but nothing systematic. I have realized of the importance of generating protocols related to this topic."*; *"I have become aware of new ways of working."*; and *"I am now more aware of the risks."*

***Participants showed a deeper knowledge and awareness of responsible data management after the session as they were able to more accurately identify relevant strategies, risks and terms (see Table B3 in Appendix B):*** Following the workshop, participants were able to identify more risk management strategies. The majority of participants (73,7%) listed the anonymization of data as a data management strategy; 31,5% indicated the importance of not saving sensitive data; and 26,3% cited student consent forms as instruments for responsible data management. Respondents were also

able to more accurately describe data management risks and protocols after the session: “Risks appear when I share data in communities or open platforms, and the consequences may appear in short term, but also over the long term (identity theft, economic theft, discrimination)”; “Risks on sensitive data are minimized by not using them, if it is not necessary. In addition, we can make sure that third parties will not see them”; “YES, all the collected and shared data is sensitive material. This (referring to potential risks) can be minimized by giving anonymity (by means of codes, pseudonyms ...)”.

### *Student Conceptions and Practices*

104 students from two educational centers responded to questionnaires to better understand how they were managing their own personal data and conceptualizing data sharing risks in the context of responsibly managing data. Results are related to our first research question about students’ conceptions of data sharing risks.

***The majority of students (74%) share data everyday through mobile apps without having read the privacy policies of such apps:*** More than half of the students have registered for 10 or more app accounts (see Table C1 in Appendix C). Over 90% use WhatsApp and Instagram every day. However, 74% have never read user terms and conditions and privacy policies before using an app.

***Almost all students (96%) agree on the importance of being formally taught about responsible data management. Although more than half (56%) have received formal training, only a third (33%) are confident with what they know about the topic*** (see Table C1 in Appendix C): Most students (96%) agree that it is important to learn about responsible data management. More than half of students (56%) stated that they had learned about data privacy and the risks of sharing personal data at school; 35% answered that they had learned about the topic on their own; and 9% responded that they had no knowledge of the topic. Although more than half (56%) indicated that they had some prior knowledge about the topic, only 33% of students indicated that they felt they knew enough about responsible data management. Despite the lack of knowledge, 85% had decided at some point not to share data because it was too risky.

***Third parties gaining access to and misusing one’s personal data is the main risk identified by the majority of students (87.5%)*** (see Table C2 in Appendix C): The main risks that students associated with sharing personal data (in order of frequency of being mentioned): (1) third parties and companies can gain access to your personal data; (2) hackers can exploit your data (e.g. loss of identity); and (3) cyberbullying. Examples of student comments regarding data sharing risks are: “Your data is sold to companies and they send you specific ads. Moreover, they know how you think”; “Problems, like hijacking or hacking accounts”; “You can be catfished, they can take your identity”; “People can use your data against you, taking your money, threatening you... “; and

*“That finally all your life would be in a cloud and people with bad intentions could access it. Like cyberbullying or like controlling your life without your knowing. They can delete all your data and you could become nobody.”*

***When asked to elaborate on ways to minimize data sharing risks, student responses were general and lacking in detail. Further, 18.9% of students responded that they did not know how to minimize the risks of sharing data*** (see Table C2 in Appendix C): Examples of student comments with regard to listing ways to minimize data sharing risks are: *“Be more conscious of the risks and know how to use social networks well”*; *“Read the instructions of apps and make your profile private”*; *“Be aware of the information that you are sharing (know where it goes, etc.)”*; and *“Do not share your private information on a public account and create a private account for sharing with friends and family.”*

### *Summary of Findings*

Study results show that none of the participating teachers received formal training related to responsibly managing data; and both teachers and students see the need for such training (see Appendix B: Table B3, Appendix C: Table C3), especially as they came to realize that their understanding of responsible data management was underdeveloped. In the evaluation of the use of apps by students, it is clear that students are actively sharing personal data with peers, companies, and, often, publicly. Yet, the overwhelming majority have never read user terms and conditions nor privacy policies. Students are aware of the general riskiness of their behaviour but most lack awareness of specific consequences of their behaviour and almost a fifth of respondents did not know of any strategies for minimizing their risks. Responses from students also suggest that students know they lack knowledge of responsible data management and recognize that formal training in the subject matter would be beneficial.

Despite not receiving formal training, the pre-questionnaire responses of teachers suggest that teachers are fairly confident in the knowledge that they hold about responsible data management and in the adequacy of the data management protocols they currently follow. However, the pre-questionnaire also revealed limits in teachers' conceptions teachers with regard to specific data sharing risks, consequences of such risks, and strategies to minimize data sharing risks. Almost half of teachers indicated that legal punishments are a risk but none of them knew of any specific legal consequences. Over a quarter of teachers did not follow any specific risk minimization strategy and over a third were not sure if they were following adequate data management protocols. Teachers mainly associated data management practices with the security protocols imposed by the school. The post-questionnaire given to teachers suggests that, after a formal training session, teachers came to realize the limitations of their prior knowledge and practices. Teachers were better able to describe specific data sharing risks and data management strategies such as anonymizing data and avoiding the

collection of sensitive data. All participating teachers rated formal training on responsible data management for teachers as being important. Overall the results suggest that there are deficits in the conceptualisations of data sharing risks among teachers and students prior to interventions involving formal training.

## **Discussion and Conclusion**

Our present study identifies potential deficits in conceptualizations and practices of teachers and learners with regard to data sharing and data management that should be considered when preparing such interventions as enhanced consent forms; details a communication approach conceived to more transparently support participants in making informed decisions that can be further elaborated upon in the future; and aims to open up discussions on informed consent in educational technology research to account for changing dynamics of relationships in terms of risk with participants.

### *Conclusion*

It is clear that educational technology research is advancing what can be measured and analyzed in educational contexts. These advancements may be shifting the data sharing relationship among researchers and participants from discrete, activity-focussed exchanges of data in classroom settings to ongoing monitoring of behaviour and physiology both within and outside of the classroom. Present research often falls somewhere between these two extremes. In parallel to shifting dynamics of relationships, recent changes in privacy law suggest a greater emphasis on autonomy and informational self-determination. Both are critical components of informed consent. For informed consent to be valid it must be freely given and informed (Faden & Beauchamp, 1986; Gallagher et al., 2010; Miller & Wertheimer, 2011; Cormack, 2016; Drachler & Greller, 2016). Consent requires understanding and “may no longer provide adequate protection and guidance either for individuals or for organizations” if such understanding is not achieved (Cormack, 2016). Faden and Beauchamp (1986) summarize the autonomous authorization (AA) model of consent by writing that, “X acts autonomously only if X acts 1) intentionally, 2) with understanding, and 3) without controlling influences.”

If teacher and learner knowledge of data sharing risks is deficient, then the effectiveness of a consent form in communicating such risks in a manner that adequately supports participants in making informed decisions about sharing their data is called into question. The ideal scenario, Miller and Wertheimer (2011) argue, would involve a comprehension test to confirm participant understanding. However, this is unrealistic in terms of practical costs for most research. The authors suggest that researchers should evaluate individual risk profiles under consideration and demonstrate that efforts that match the level of risk have been made to ethically inform participants. Enhanced consent forms offer an attainable upgrade on current efforts. Ultimately, the decision on what practices are permissible and how informed consent is achieved is left up to ethics

committees. Few in the field want to complicate the consent process and raise additional barriers to obtaining research participation. On the other hand, it can be argued that responsible research includes field research that can inform ethics committees in establishing the ethical standards upon which their decisions are based.

### *Limitations of the study*

This paper is positioned within educational technology research relating to a project focussing on community and learning analytics as reported in the publications, *Teacher-led inquiry in technology-supported school communities* (Michos, Hernández-Leo, Albó, 2018); and *Supporting awareness in communities of learning design practice* (Michos & Hernández-Leo, 2018). The project and our survey study involved a small number of students and teachers in high school and vocational school settings which limits the generalizability of the findings. Further, due to the multidisciplinary of the topic and limitations in time, the literature search conducted was not exhaustive. The search was restricted to specific keywords in SCOPUS and did not include other search engines or databases. SCOPUS was selected as it is the largest citation database of peer-reviewed literature. The keywords (ethics AND educational technology, learning analytics) were selected based on the context of the project. However, literature relating to the area of ethics in educational technology research is not always easily discoverable. It may be defined using alternative keywords to those used in our search or may be positioned in nearby fields of study rather than explicitly in the field of educational technology. Thus, the literature search could be enriched by searching in related educational technology journals and databases and other closely related fields which discuss the issues surrounding adequately informing participants, see for example social media research (von Benzon, 2018) and online research ethics (Dawson, 2014). Finally, future studies that evaluate the effectiveness of interventions on participant comprehension of informed consent in educational technology research would be welcomed.

### **Acknowledgements**

The authors would like to thank all teachers and students who participated in the study, as well as Laia Albó for producing the data privacy video. This work has been partially funded by “la Caixa Foundation” (CoT project, 100010434) and FEDER, the National Research Agency of the Spanish Ministry of Science, Innovations and Universities MDM-2015-0502, TIN2014-53199-C3-3-R, TIN2017-85179-C3-3-R. DHL is a Serra Hünter Fellow.

### **Statements on open data, ethics and conflict of interest**

The ethics procedure was approved by the ethics committee of Universitat Pompeu Fabra Barcelona. Consent was obtained from participants. Anonymized data excerpts

from the teacher and student questionnaires are available in Zenodo. There are no potential conflicts of interest in the work.

## References

Acquisti, A. (2010). The economics of personal data and the economics of privacy.

Arnold, K. E., & Sclater, N. (2017). Student perceptions of their privacy in learning analytics applications. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (pp. 66-69). ACM.

Beardsley, M., Santos, P., & Hernández-Leo, D. (2018). Ethics in educational technology research: informing participants on data sharing risks [Data set]. *Zenodo*. doi.org/10.5281/zenodo.1486056

Bransford, J. D., & Johnson, M. K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of verbal learning and verbal behavior*, 11(6), 717-726.

Campbell, F. A., Goldman, B. D., Boccia, M. L., & Skinner, M. (2004). The effect of format modifications and reading comprehension on recall of informed consent information by low-income parents: a comparison of print, video, and computer-based presentations. *Patient Education and Counseling*, 53(2), 205-216.

Canadian Institutes of Health Research, Natural Sciences and Engineering Research Council of Canada, and Social Sciences and Humanities Research Council of Canada, *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans*, December 2014.

Cormack, A. N. (2016). Downstream Consent: A Better Legal Framework for Big Data. *Journal of Information Rights, Policy and Practice*, 1(1).

David, M., Edwards, R., & Alldred, P. (2001). Children and school-based research: 'informed consent' or 'educated consent'? *British educational research journal*, 27(3), 347-365.

Dawson, P. (2014). Our anonymous online research participants are not always anonymous: Is this a problem?. *British Journal of Educational Technology*, 45(3), 428-437. doi:10.1111/bjet.12144

Delany, C. (2008). Making a difference: incorporating theories of autonomy into models of informed consent. *Journal of Medical Ethics*, 34(9), e3-e3

Drachsler, H., & Greller, W. (2016, April). Privacy and analytics: it's a DELICATE issue a checklist for trusted learning analytics. *In Proceedings of the sixth international conference on learning analytics & knowledge* (pp. 89-98). ACM.

Eggers, W. D., Hamill, R., & Ali, A. (2013). Data as the new currency: Government's role in facilitating the exchange. *Review of Deloitte*, 13, 18-31.

Faden, R. R., & Beauchamp, T. L. (1986). *A history and theory of informed consent*. Oxford University Press.

Ferguson, R. (2012). Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, 4(5/6), 304-317.

Flory, J., & Emanuel, E. (2004). Interventions to improve research participants' understanding in informed consent for research: a systematic review. *Jama*, 292(13), 1593-1601.

Gallagher, M, Haywood, SL, Jones, MW, Milne, S (2010) Negotiating informed consent with children in school-based research: a critical review. *Children and Society* 24(6): 471–482.

Glenberg, A. M. (2011). How reading comprehension is embodied and why that matters. *International Electronic Journal of Elementary Education*, 4(1), 5-18.

Gundlach, G. T., & Murphy, P. E. (1993). Ethical and legal foundations of relational marketing exchanges. *The Journal of Marketing*, 35-46.

Hughes, R. C. (2014). Justifying community benefit requirements in international research. *Bioethics*, 28(8), 397-404.

Kendeou, P., Smith, E. R., & O'Brien, E. J. (2013). Updating during reading comprehension: Why causality matters. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39(3), 854.

Kendeou, P., & O'Brien, E. J. (2015). Prior Knowledge. *Handbook of individual differences in reading: Reader, text, and context*, 151.

Kendeou, P., Van Den Broek, P., Helder, A., & Karlsson, J. (2014). A cognitive view of reading comprehension: Implications for reading difficulties. *Learning disabilities research & practice*, 29(1), 10-16.

- Kendeou, P., & Van Den Broek, P. (2007). The effects of prior knowledge and text structure on comprehension processes during reading of scientific texts. *Memory & cognition*, 35(7), 1567-1577.
- Lipson, M. Y. (1982). Learning new information from text: The role of prior knowledge and reading ability. *Journal of Reading Behavior*, 14(3), 243-261.
- Lohr, S. (2012). The age of big data. *New York Times*, 11(2012).
- Manathunga, K., & Hernández-Leo, D. (2018). Authoring and enactment of mobile pyramid-based collaborative learning activities. *British Journal of Educational Technology*, 49(2), 262-275.
- Mandava, A., Pace, C., Campbell, B., Emanuel, E., & Grady, C. (2012). The quality of informed consent: mapping the landscape. A review of empirical data from developing and developed countries. *Journal of medical ethics, medethics-2011*.
- Martin, K. (2016). Understanding privacy online: Development of a social contract approach to privacy. *Journal of Business Ethics*, 137(3), 551-569.
- Mayne, F., Howitt, C., & Rennie, L. (2016). Meaningful informed consent with young children: looking forward through an interactive narrative approach. *Early Child Development and Care*, 186(5), 673-687.
- Michos, K., Hernández-Leo, D. (2018) Supporting awareness in communities of learning design practice, *Computers in Human Behavior*, 85 (Aug. 2018), 255-270.
- Michos, K., & Hernández-Leo, D., Albó, L. (2018). Teacher-led inquiry in technology-supported school communities. *British Journal of Educational Technology* 49(6), 1077-1095.
- Miller, F. G., & Wertheimer, A. (2011). The fair transaction model of informed consent: an alternative to autonomous authorization. *Kennedy Institute of Ethics Journal*, 21(3), 201-218.
- Moore, T. P., McArthur, M., & Noble-Carr, D. (2018). More a marathon than a hurdle: towards children's informed consent in a study on safety. *Qualitative Research*, 18(1), 88-107.
- Murphy, R. S. (2017). Property rights in personal information: An economic defense of privacy. In *Privacy* (pp. 43-79). Routledge.

Nissenbaum, H. (2011). A contextual approach to privacy online. *Daedalus*, 140(4), 32-48.

Nishimura, A., Carey, J., Erwin, P. J., Tilburt, J. C., Murad, M. H., & McCormick, J. B. (2013). Improving understanding in the research informed consent process: a systematic review of 54 interventions tested in randomized control trials. *BMC medical ethics*, 14(1), 28.

Pardo, A., & Siemens, G. (2014). Ethical and privacy principles for learning analytics. *British Journal of Educational Technology*, 45(3), 438-450.

Parsons, S., Sherwood, G., Abbott, C. (2015) Informed consent with children and young people in social research: is there scope for innovation? *Children and Society* 30(2): 132–145.

Prinsloo, P., & Slade, S. (2015). Student privacy self-management: implications for learning analytics. In *Proceedings of the fifth international conference on learning analytics and knowledge* (pp. 83-92). ACM.

Purtova, N. (2015). The illusion of personal data as no one's property. *Law, Innovation and Technology*, 7(1), 83-111.

Rouvroy, A., & Poullet, Y. (2009). The right to informational self-determination and the value of self-development: Reassessing the importance of privacy for democracy. In *Reinventing data protection?* (pp. 45-76). Springer, Dordrecht.

Schenker, Y., Fernandez, A., Sudore, R., & Schillinger, D. (2011). Interventions to improve patient comprehension in informed consent for medical and surgical procedures: a systematic review. *Medical Decision Making*, 31(1), 151-173.

Shao, C., Ciampaglia, G. L., Varol, O., Flammini, A., & Menczer, F. (2017). The spread of fake news by social bots. *arXiv preprint arXiv:1707.07592*, 96-104.

Schwartz, P. M. (2003). Property, privacy, and personal data. *Harv. L. Rev.*, 117, 2056.

Slade, S., & Prinsloo, P. (2013). Learning analytics: Ethical issues and dilemmas. *American Behavioral Scientist*, 57(10), 1510-1529.

Toprak, E., Ozkanal, B., Aydin, S., & Kaya, S. (2010). Ethics in E-Learning. *Turkish Online Journal of Educational Technology-TOJET*, 9(2), 78-86.

von Benzon, N. (2018). Informed consent and secondary data: Reflections on the use of mothers' blogs in social media research. *Ethics in/of Geographical Research*, 51(1), 182-189.

Wang, S., & Heffernan, N. (2010). Ethical issues in Computer-Assisted Language Learning: Perceptions of teachers and learners. *British Journal of Educational Technology*, 41(5), 796-813.

Wachter, S. (2018). Normative challenges of identification in the Internet of Things: Privacy, profiling, discrimination, and the GDPR. *Computer law & security review*, 34(3), 436-449.