Designing for GDPR - Investigating Children's Understanding of Privacy: A Survey Approach

John Dempsey
University of Central Lancashire
Preston, UK.
jpdempsey@uclan.ac.uk

Gavin Sim
University of Central Lancashire
Preston, UK.
grsim@uclan.ac.uk

Brendan Cassidy
University of Central Lancashire
Preston, UK.
bcassidy1@uclan.ac.uk

The General Data Protection Regulation (GDPR) places new obligations on businesses that collect and process data from children. It goes so far as to say that privacy notices should be presented in child-friendly and age appropriate formats. Fulfilling GDPR obligations will require designers to have a better understanding of how children understand privacy issues. This research aims to investigate children's understanding of privacy online. Thirty-two children from a UK primary school, aged between 8 years and 10 years old completed a survey to gauge their understanding of privacy. Eight different scenarios were presented to the children and they had to decide whether the information should be kept private or not and state the reason why. This work identifies that children do have an understanding of privacy, especially when related to online safety. However, children do not yet understand that their data has an inherent value, have misconceptions about data and what data should be protected. This highlights the challenges for designers of technology used by children to meet the GDPR obligations.

Children's privacy, privacy design, privacy knowledge, child computer interaction, children's survey.

1. INTRODUCTION

There has been growing interest in aspects of computer security within the Human Computer Interaction community (Schechter, 2013) including the challenges and issues associated with privacy (Conti & Sobiesk, 2007; Gill, Vasalou, Papoutsi, & Joinson, 2011; Silva, Silva, Silva, & Mourão, 2017). While humans are interacting with technology in different ways, data processors are collecting data entered directly into applications, but also quietly collecting usage data in the background (Dey, Ding, & Ross, 2013). Not only is this data being collected, it is being aggregated with other data and stored for some future use that has not yet been imagined (Conti & Sobiesk, 2007; Solove, 2006). It could be suggested that society and technology firms have not done enough to protect children, a vulnerable user group, from the potential dangers of intrusive technologies on the Internet.

Research into the harvesting of personal data has predominantly focused on the opinions of adult users of such systems, yet children also understand that they have special things that need protection; they are even capable of devising their own ways to protect those special things (J. Read & Beale, 2009). Despite this, they may not yet have the experience or knowledge to understand that their data is special and is a valuable commodity that needs protection. Children use computers in different ways and for different reasons than adults, and the privacy of their data may not be their main concern (J. C. Read, 2005).

Some privacy researchers suggest that children are not interested in privacy (Cranor, Reagle, & Ackerman, 2000) however it is important to position privacy work in a way that is usable by children, so that they are equipped with the knowledge and skills necessary to turn them into the digital citizens of the future, capable of using and creating internet-based services safely.

This study attempts to understand whether children already have an understanding of privacy, and whether it is actually possible to position privacy work in a child-friendly manner. People value privacy in different ways, so it is not possible to assess a child's understanding based purely around a question and response survey; however, it should be possible to design a survey that infers whether or not children have an interest in privacy by asking them to take decisions about privacy that relate to a situation they can engage with.

This paper is organised as follows. Section 2 discusses work that is related to this explorative study. Section 3 describes the design and justification of the study. Section 4 presents the results of the study. Section 5 analyses the results

and identifies how this work could be extended in the future.

2. RELATED WORK

New legislation has placed additional obligations on any business collecting or processing data within the European Union; the impact of which will be felt globally. This presents new challenges to the HCI community and organisations in establishing effective ways to communicate privacy issues to children. It is important for those designers to learn the lessons from the Child Computer Interaction community and ensure that any designs are child-centred. Thus the related work will focus on two aspects privacy and children.

2.1 Privacy

Despite the huge amount of research around privacy, there is still no unified definition of what privacy means. The one thing that privacy researchers tend to agree on, is that they cannot agree about the definition of privacy (Smith, 2014).

This study attempts to measure if children understand privacy concepts by asking the children questions that relate to privacy issues. The questions relate to an online setting where the child must make decisions about their "control over personal information" (Westin, 1967).

2.2 Privacy Laws

The General Data Protection Regulation (GDPR) is changing the way data processors deal with data and makes special requirements about child-related data. Among other things, privacy notices can no longer hide behind complex legal language and must be set out in a child-friendly manner that uses a child-friendly presentation (for example, icons, graphics, and cartoons) that explain the implications of sharing data with the data processor (Information Commissioners Office, 2018). If the data processor seeks consent from the child (or person with parental responsibility) then the data processor is responsible for ensuring the child can understand what they are consenting to; if they do not understand then they cannot give consent.

Research has shown that the majority of people do not read privacy policies and end-user license agreements (Bakos, Marotta-Wurgler, & Trossen, 2014), and even when they do they find them difficult to understand (Luger, Moran, & Rodden, 2013). GDPR intends for this to change.

Children are using tablet devices and installing apps not only for entertainment purposes but to facilitate their learning within schools (Henderson & Yeow, 2012; Mann, Hinrichs, Read, & Quigley, 2016). Currently, when a child accesses an app via an app store, they can install software without giving explicit consent for their data to be collected. Without reading the privacy notice, the child is likely to be unaware of the permissions they are giving away. Even if they do read the privacy notices they may struggle to comprehend the information (Stothard & Hulme, 1992). GDPR intends to make this process more explicit so that children understand the consequences of the agreements they make about the use of their data. The challenge for the HCI community is how this can be achieved.

The main aim of GDPR is to improve the protections offered by data processors, not to inform a data subject about how to maintain their privacy. The data processors want the data subject to share as much of their data as possible, and after GDPR, it is conjectured, they are likely to use language and presentation to entice children to share the data that they could have kept private.

While GDPR is European-based legislation, it claims jurisdiction over any EU-based data processor and EU-based data subject. Despite being based only in Europe, the consequences of GDPR will be felt around the world.

The Children's Online Privacy Protection Act (COPPA) was US-based legislation that required those with parental responsibility to control the online privacy of children under the age of thirteen (Shmueli & Blecher-Prigat, 2011). GDPR places extra conditions on data processors that collect the data of children younger than thirteen years, including the requirement to have consent from both the child and the adult with parental responsibility. With iPads being adopted and used in facilities such as nurseries, with children as young as 3 years old (Flewitt, Messer, & Kucirkova, 2015), questions are raised about whether obtaining consent would even be viable.

While GDPR does not intend to help children make decisions about their privacy, it still requires the data processors to use language and presentation that is most appropriate for those age groups. Different age groups will communicate in different ways, so data processors must create presentations appropriate for specific age groups. The Child Computer Interaction community has long established methods to enable children to be design partners in the creation of technology (Druin, 1999). This is an opportunity to engage children in the conversation about the presentation of privacy related matters, to go further than GDPR and instead focus on the child data-subjects.

It is anticipated that GDPR will have a big impact on the way software is designed for children; not only with the presentation of privacy notices, but also in their delivery and operation. Despite every adult once being a child, they are not children, do

not think like children, and do not act like children; and it is imperative that children remain the central actor in any design exercises through participatory design methods (J. Read et al., 2002).

The data processors will employ designers to convey a message to children about consent and data protection; but unless those designers understand the context in which those children have read those privacy messages, then those designs may not be as child-friendly as the designer intended.

2.3 Measuring Privacy Knowledge

There is yet to be an established method for measuring another person's knowledge about privacy; and the work undertaken to understand a child's understanding of privacy is severely limited.

Issues relating to privacy do not start and stop at the keyboard, but extend into every part of an individual's life. The availability (or lack thereof) of data online might cause privacy problems in both the digital world and the real world.

For example, broadcasting the location of your birthday party might cause safety issues around stranger-danger; but might also be useful to inform parents of your physical location.

Privacy definitions take the form of tacit knowledge and has different meanings to different people (Dwyer, 2009). This lack of consistency has potentially confounded many privacy-related studies, where a limited definition of privacy has been assumed.

A good example of this can be found in a study which uses a comic book to engage children in the privacy conversation (Zhang-Kennedy, Baig and Chiasson, 2017). Part of their evaluation includes a test to discover how much about privacy has been understood by the participants by asking ten knowledge-based questions that the participants should learn by reading the comic book. At no point in the paper do the authors define their vision of privacy and instead it seems that privacy simply has an implied definition.

While Zhang-Kennedy et al (2017) do not define their vision of privacy, they seem to focus on the idea that privacy is something belonging to the individual that can be lost or taken away. They use the example of Jane uploading a photograph while walking her dog, and seem to suggest that Jane has lost her privacy because she has advertised her location on social media. This example would seem to be a classic definition of privacy where one has the "right to select what other people know about you" (Westin, 1967).

An alternative way to conceptualise privacy was proffered by Daniel Solove (Solove, 2006) where you consider the privacy problem using a taxonomy containing several identified potential problems; and then balancing those privacy problems against other competing interests. This method of analysing the privacy problem would not necessarily agree that Jane uploading photographs while walking her dog causes a privacy problem.

For example, is Jane using a secure method to upload the photograph? Does her social media account automatically remove geo-tagging of photographs? Does her social media account have access control preventing strangers from seeing her uploaded photographs? If Jane did not upload the picture and she went missing, then would her parents and the police have any information about her last known whereabouts? There must be a balance between the privacy concerns for uploading and for not uploading the photograph.

While it is important to engage children in the privacy conversation, it is also important to acknowledge that, while privacy is tacit knowledge, it will be informed by the many different experiences of the child participant (Adams & Sasse, 2001). For the child to participate in a valuable learning experience, any developed system must ensure that it does not attempt to pass off the researcher's personal values of privacy masquerading as the "one true definition".

Privacy has both a personal value and a cultural/societal value, and to make judgements about another person's understanding requires a certain amount of flexibility in assessing the descriptions given (Krasnova & Veltri, 2010). Just because somebody's definition of privacy does not match my definition of privacy, does not mean that they do not have a perfectly good grasp of the relevant concepts.

Zhang-Kennedy et al (Zhang-Kennedy, Mekhail, & Chiasson, 2016) stated that "young people (aged 7-11 years old) valued their privacy, yet only about half of them actually understood what it meant to remain private while online". How can these children value something that they do not actually understand? There seems to be a contradiction in these statements, yet it is not actually clear why this contradiction arises. Later, when discussing the children's views of privacy, all of the children gave descriptions of privacy that seemed completely reasonable.

The children gave descriptions of privacy that included "to be alone", "to hide secrets or special things", "to keep things to yourself" and "to not talk to strangers" (Zhang-Kennedy et al., 2016). Zhang-Kennedy et al then classified these to suggest that "only half of the children understood what it meant to remain private while online". It could be argued that each answer demonstrates some level of understanding of privacy, and while it may not fully

match the author's definition of privacy, they are certainly headed in the right direction.

When assessing a child's level of knowledge around privacy it will be essential to be clear which definition of privacy is used; however even if we are clear up-front about this then we should remain flexible in our assessment of privacy afterwards. A child should only be deemed to have little or no understanding of privacy concepts if their answers contain no characteristics that relate to privacy.

2.3 Learning from Adult-Based Lessons

There have been many studies examining an adult's understanding of the issues relating to privacy. A set of privacy harms or concerns were identified that relate to "privacy panic", that sinking feeling you get when you realise you've just informed the world about something that you really did not want to (Angulo & Ortlieb, 2015). This study was targeted at adults and not children, but one might have hoped to draw conclusions that would relate directly to children. The concerns and consequences of privacy panic simply are not felt in the same way with children. Adults have adult concerns and children have child concerns (J. C. Read, 2005), while the two may share an overlap they have mutually exclusive parts too.

For example, "possible loss of employment" and "money going missing or financial harm" are not the concerns of a child. It may be that "embarrassment or damage to my reputation" is relevant to some children, but their reputations probably cannot be damaged in the same way as that of an adult. Perhaps their embarrassment and reputational damage can be caused by not having access to the privacy risking service in the first place.

One study considered the balance between privacy concerns and how a business uses collected information and presented a model to describe the "privacy leverage point" (Culnan & Armstrong, 1999). Children are an important user group that are being targeted for commercial purposes, their information is being collected, aggregated and used for targeted marketing purposes, but would this model and would this study be able to draw conclusions suitable for the child market it targets?

The model depicts how the difference between customer expectation and business practice can have a bearing on customer retention and the ability to attract new customers (Culnan & Armstrong, 1999). A child's concerns about using a service might extend only so far as being able to use the same tools as their friends or family, and while the literature does suggest children are concerned about their privacy, it does not suggest that privacy is their primary concern (Zhang-Kennedy et al., 2016).

When a child is playing freely available computer games via a mobile phone app store they may be unwittingly taking part in a "privacy calculus" where they give away their data for the benefit of receiving access to the computer game for free. A child may not yet have expectations for how they expect the business to handle their data; thus the business can take advantage of this lack of leverage point.

A "privacy benefit" is based on the idea that some people use their privacy as a commodity that can be swapped for certain economic benefits (Pavlou, 2011). While it could in fact be argued that this is entirely true for children, is this a conscious decision taken by the child? Can they accurately estimate the value of their own data and use it proportionally? Even an adult could not accurately estimate the future value of a child's data, so it is not likely that a child would be able to.

The three main privacy benefits for adults include "a) financial rewards, b) personalisation, and c) social adjustment" (Pavlou, 2011). We know that some children are giving away their data in exchange for reward; freely downloadable apps are collecting usage statistics which is being collected and aggregated for later use; but the privacy benefits here are not the same as for adults.

Much of the privacy literature has focused on adults within a commercial context and many of these papers provide conclusions that could possibly be applied to children, but only after they have been viewed from the child's point of view. The Child Computer Interaction (CCI) community quite clearly views that only children are the experts in being children, and adults, no matter how hard they try, are not children, and do not have the same actions, behaviours or concerns as children (Read, 2005).

2.4 The Challenge of Children

The Child Computer Interaction community has identified lots of challenges associated with working with children when they are used as part of the design team or for the evaluation of technology. As a result of these challenges methods have had to be adapted to meet the needs of children.

Often, children have a skewed perception of technology, and see that all technology is good, even when it is not. Childhood is supposed to be the happiest time of our lives and therefore the instruments used to measure this must be appropriately designed for children (Hall, Hume, & Tazzyman, 2016).

During their study Zhang-Kennedy et al (2016) used 30-minute interviews with children, who they kept separated from their parent (although within eye sight). There are many problems associated with interviewing children, some of which have been addressed by using techniques such as group

interviews, or using stories to help the children articulate their answers (To et al., 2016).

Horton, Read and Sim (2011) identified another problem when the child attempts to satisfice the questions being asked. Here the child gives an answer that is 'good enough' (Hox, J.J.; Borgers, 2001) rather than give full effort into working out an answer. To avoid concerns related to satisficing different techniques must be employed that ensure any questions are easy to understand and to ensure the answers are easy to give (J. C. Read, 2007). Asking the same question and using the same words will make the question easier to answer as the children will have more experience of what the question actually means; providing a selection of possible answers will also make answering the questions more easy.

Metzger, Flanagin and Nekmat (2015) report on 'comparative optimism' which postulates that people believe good things will fall on oneself, while worse things will fall onto others. This means that it is easier for people to think about the safety of others, rather than themselves. Because privacy is loosely related with safety and security, it is important to make sure that children are not thinking about their own privacy, but the privacy of somebody else who they can help protect.

If children are to be empowered to make decisions about their privacy, then it is also important to acknowledge those that influence them, such as family and friends (Minkus, Liu, & Ross, 2015). Whatever work is undertaken must consider the fact that an empowered child must be able to influence the decisions taken by others. "Sharenting" is when a child has little or no influence over the information that a parent shares about their child (Steinberg, 2017).

It is difficult to explore the negative aspects of technology when a child's outlook is through a positive lens.

3. STUDY

In order to gain a better understanding of whether or not children understand privacy, and how they understand privacy, a survey study was undertaken with children. This study attempted to take into account the issues that have been identified in the literature.

3.1 Design objectives

The objectives of this study was to attempt to understand if children in the age range of 7-11 years old understood concepts behind online data privacy. This study did not set out to force the participants into a predefined concept of privacy designed by the researcher beforehand, but set out to be flexible and to allow the children to articulate their own understanding of privacy. This study did not look to make definitions of privacy from a child's perspective, but would look to the data collected and look for themes or characteristics that might indicate some kind of expressive understanding of privacy-related understanding.

Objective: to see if children demonstrated an understanding of privacy-related matters.

3.2 Participant

As the study was aimed at children in the age range of 7-11 years old, Piaget's theory of cognitive development was taken into consideration when designing the material. This theory suggests that children in their concrete operational stage (roughly 7-11 years old) are able to "work things out in their heads" (Lee, 2000). At this age children should be competent enough to use digital devices on their own, and may risk revealing private information.

The study was undertaken by 32 children from a UK primary school aged between 8 years and 10 years old. There were 17 boys and 15 girls. Two researchers facilitated the discussion and assisted the children in completing the survey along with two teachers from the children's school. The two researchers had prior experience of conducting research with children.

The children were invited to come to the university as part of a MESS day (Mad Evaluation Session with Schoolchildren) (Horton, Read, Mazzone, Sim, & Fitton, 2012). During the day the children took part in three different activities all lasting approximately 30 minutes each.

3.3 Designing the Activity Workbook

An activity booklet was created to capture the extent to which children understood concepts related to privacy. It was based on a story where they were asked to help another child, from a developing country, to choose what information she should publish on her new website. The story talked about 'Opaline' from Mauritius who wanted to create her own website, so that she could talk to and make friends with children in the UK. Opaline was not sure what information other children thought should be published.

A key objective, when designing the story, was to position the participant as an "external consultant", enabling them to consider the safety of others. A scenario containing children from a developing nation would encourage the participants to think they had more experience than Opaline.

Various countries were considered such as India or Pakistan, but these countries had a high chance that some of the children (or their families) may actually come from there. Care had to be taken not

to cause distress or opportunity to bully. In the survey the child was named "Opaline" for several reasons: a) based on common female Mauritian names, b) the name did not contain any special characters; c) the name was different and special enough to remember.

Some statistics about Mauritius were included to enable the researchers the opportunity to talk about what a 'developing nation' was; and highlight differences with the UK. The discussion included the fact that most homes in the UK have broadband access with high data transfer speeds which is not the case in Mauritius. It was expected that the participants would be surprised at how much of the Mauritian population did not have broadband access. Choosing Mauritius also gave the researchers the opportunity to engage participants by talking about the Dreamworks film, Madagascar.

The survey was designed in three sections.

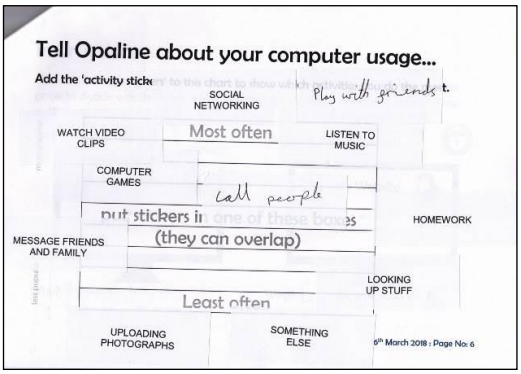


Figure 1: *Activity Booklet front page*

At the start of the workbook (see Figure 1), Opaline asks the participants to tell her about how they use computers in their lives. This data was collected to position the participants as 'typical' computer users and the list of activities provided were identified from the literature (Livingstone, Davidson, & Bryce, 2017). This data was never going to be used to exclude any data; however it would be used to check the sample used computers for the same typical reasons reported in the literature.

The survey required the participants to give advice to Opaline about whether or not to 'keep something private' or to 'make it public'. The rationale for using this technique was the fact that people tend to believe that good things will land on themselves while bad things will land on others (Metzger et al., 2015) and by providing advice to Opaline, it externalised the risk away from the participants, allowing them to consider the risks of something bad happening.

The second part of the survey consisted of eight survey questions which the children had to answer. Each of the eight survey questions had one of two answers, either "make it public" or "keep it private", and the participants were asked to tell Opaline if she should make several scenarios public or private. Underneath each question was a space for the child to write down the reason why they chose that option (see figure 2). The survey questions were designed not to ask the participants outright questions to try and encourage truthful answers and mitigate against satisficing. Instead, the children were provided with a set of sticker books that they could use to answer the questions.

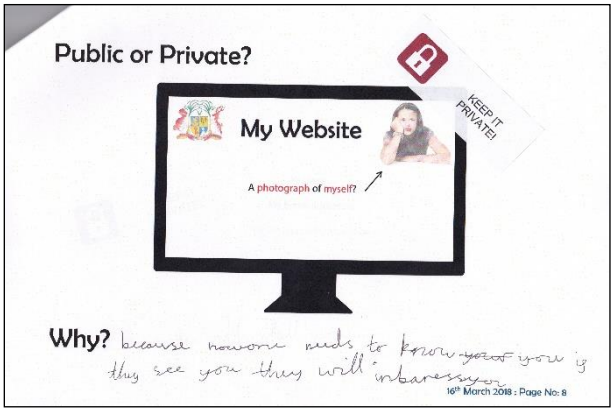


Figure 2: *One of the survey question pages*

The survey questions were all about publishing information on her new website. Each of the questions was designed to test if the participant would be able to identify a privacy concern. Westin's definition of privacy was used to design the questions, and each question revolved around controlling access to information within an online setting (Westin, 1967). Concerns were grouped around personal safety, personal reputation (e.g. risk of being bullied), permission (e.g. doing something without permission), and data (e.g. that their personal data has a value). Privacy issues were aimed at Opaline herself, somebody else (e.g. the best friend) or an object that was owned by Opaline. Some of the questions had obvious privacy concerns and others were more subtle.

The first two questions related directly to the publishing of personal information about Opaline. These scenarios were selected due to their obvious relationship to privacy and were expected to help the children settle down and understand the task. The first question asked about uploading a photograph of herself, and the second question asked about publishing her home address.

The third question asks about uploading a photograph of Opaline's best friend. Similar to the first question, although this question intended to understand whether a child understood that others were also entitled to privacy.

The fourth question has Opaline asking people to vote on if they like her new mobile phone. This question was no longer about the data subject and instead focused on a device; however, this device

still provided information about Opaline, and could be used to invade her privacy.

The fifth question asked whether it was OK to share a YouTube channel with her favourite band. At first glance this looks like a perfectly innocent web page; however, under closer examination it revealed information about Opaline and her whereabouts. This question would attempt to understand if children looked at the entire scenario, or if they took their first impression only.

The sixth question asked whether it was OK to publish a book review for a homework assignment. This was chosen to represent a situation without any obvious privacy concerns.

The seventh question advertises the location of a birthday party on social media. This question seeks to understand if children understand the risks of social media, and specifically advertising their whereabouts to lots of people they may not know.

The last question provides a product review of a brand new smart watch. It was designed to see whether the children would understand that Opaline was advertising that she had recently spent money on a device that would be easy to steal.

When designing the questions, the researcher had their own viewpoint on privacy, and the tacit nature of privacy means those concerns might be interpreted differently by the children answering the questions. For example, the fourth question talks about mobile phones and shows a picture of an iPhone. Some people may be concerned that this tells others that they own an iPhone; others may be unconcerned because many people have iPhones but the cultural context of Mauritius may also be taken into account.

3.4 Procedure

The children arrived at the University and were taken to a lab for an induction. They were informed about the concept of data, and consent was explained to them. The children were told who was paying for the MESS day and given the opportunity to not take part.

The class was split up into three smaller groups; each of which would have 30 minutes with one of the three research exercises taking place in different locations.

For this study four tables were arranged so that all the children could see the researcher while he was talking. Each participant had their own chair to sit on, and was given a workbook and pen, and each workbook came with a book of stickers within.

The researcher introduced himself and explained what he was interested to find out and invited the participants to complete the activity book.

Opaline was introduced and a discussion was had about favourite characters from the film Madagascar. This was to put the children at ease with the researchers and encourage them to communicate. Internet connectivity in Mauritius was introduced and the participants asked if they could imagine not having Internet access.

The researcher went through the activity book page by page and explained what each page was asking for. The participants worked at different speeds, but the researcher continued to move onwards to ensure that steady progress was made, and so that the study would finish within the allotted 30 minutes. All participants managed to complete the exercise within the allocated time, although some left blank answers.

After they had completed the eight questions, they were debriefed about what they had just done, and reminded that they could keep their data. There were activities (a word search and a cypher quiz) that the participants could take with them.

3.5 Data Analysis

For the purpose of this paper just the 8 questions that related to Opaline's website were analysed. Due to the tacit nature of privacy within these questions it might be difficult to really gauge whether a child's answer demonstrates complete understanding, thus a ranking scale was proposed along with two separate classifications for the responses.

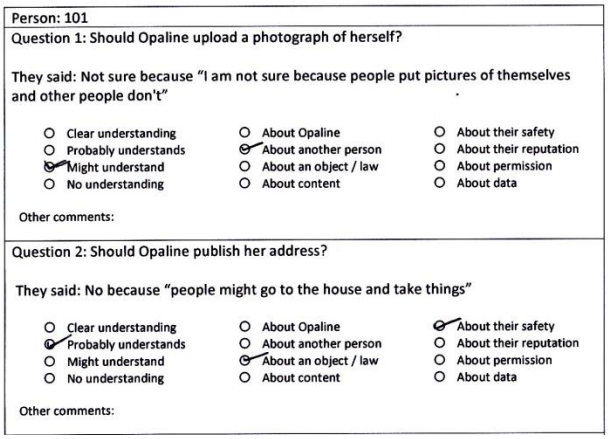
The aim of the ranking scale was to identify whether or not the child had understood an issue relating to privacy. They either demonstrated some understanding or did not demonstrate an understanding. However, it was expected that the answers written by the participants would not be so clear-cut, and instead a four point rating scale was used to help categorise the child's understanding of privacy. The scale was a) clear understanding b) probable understanding c) they might have understood d) no understanding. Within this ranking system a and b were judged to demonstrate understanding whilst c and d were judged to demonstrate no real understanding. These rating scales were to help the analyst to determine whether the participant had understood.

Based on the questions, the second part would categorise who the child was concerned about. For example, their concern might be about Opaline herself, somebody else (e.g. a parent or friend), an object that relates to an object, or the content of the website itself. As the participant will be free to write anything they could also record other information.

The final part related to what their concern was about. This was broken down into the categories of safety, reputation, permission and data. Safety dangers relate to stranger-danger and physical

harm that might befall a victim; reputation danger might be linked with bullies or actions taken to save face; permission dangers linked to what might happen if permission is not received; and data dangers relate to the value of the person's data.

The information from each activity booklets was copied into a spreadsheet and subsequently mail merged into a document that would be analysed, see Figure 3.


Person: 101

Question 1: Should Opaline upload a photograph of herself?

They said: Not sure because "I am not sure because people put pictures of themselves and other people don't"

○ Clear understanding	○ About Opaline	○ About their safety
○ Probably understands	✓ About another person	○ About their reputation
✓ Might understand	○ About an object / law	○ About permission
○ No understanding	○ About content	○ About data

Other comments:

Question 2: Should Opaline publish her address?

They said: No because "people might go to the house and take things"

○ Clear understanding	○ About Opaline	✓ About their safety
✓ Probably understands	○ About another person	○ About their reputation
○ Might understand	✓ About an object / law	○ About permission
○ No understanding	○ About content	○ About data

Other comments:

Figure 3: *Data Analysis mail merge*

Three analysts were used to categorise the data. Analysts included one of the authors of this paper, a post-graduate research assistant and a public health officer from local government. The research assistant and public health officer had not been involved in any part of the research design or running of the study. They were instructed what they should do, but were not told about how to interpret the child's understanding of privacy. Each analyst categorised the participant's responses independently of each other. There were a few instances in which the analysts were unable to rate the children's responses and these were omitted from the overall analysis.

The results of this analysis were then summarised into a spreadsheet, allowing a comparison between the different analysts to be performed. A Cronbach Alpha analysis was conducted on the responses to the 8 questions, $\alpha = .875$ indicating a high level of agreement between the analysts. The three analysts appear to have similar views as to whether the children understand privacy or not based upon their responses.

4 RESULTS

The results will be presented based on the three forms of analysis, the ranking and the two categories.

4.1 Understanding of Privacy

Table 1 shows the responses for each of the questions and whether the child thought it should be made public or private.

Table 1: *Number of children who stated the information should be made public or private*

Q	Number Public	Number Private	Unsure
1	*3*	*27*	*1*
2	*1*	*30*	*0*
3	*10*	*21*	*0*
4	*26*	*3*	*2*
5	*16*	*14*	*1*
6	*27*	*4*	*0*
7	*4*	*27*	*0*
8	*17*	*9*	*5*

As can be seen from the results there was lack of agreement between the children on whether something should be kept private or made public especially relating to questions 3, 5 and 8.

Question 3 related to uploading a picture of her best friend. There were various reasons why children thought this should be made public including, "because her friend said she could", "you can see your best friends" and "people expose their best friends". Whilst some reasons for keeping private included "people will then know about you and your friend", "they might not be able to go online" and "people will know your friend and you".

Question 5 was about advertising her favourite band via a YouTube channel and some of the reasons for keeping this information private were "people can make fun of her" and "because the youtuber could edit it and turn it into something rude". Whilst the reasons for making the information public included "It's not that important" and "other people might like it too".

Whilst for question 8 which was about writing a review of a brand new smart watch the public responses included "showing what she bought", "so other people can buy them" and "it is just a watch". Whilst the reasons for keeping it private included "people could rob you people they think it is valuable; advertising smart watches" and "someone might steal it and connect to her phone".

Many of the reasons stated would suggest that these children were mostly aware of privacy-related matters when those issues related to personal safety or stranger-danger, but not necessarily when the risk/danger was something different.

To establish whether the children's responses to the questions inferred any understanding of privacy the results of the three raters were then analysed. Table 2 shows the percentage of children that the

raters thought understood privacy based on their responses to each of the questions.

Table 2: % of Understanding vs Non-Understanding

Q	% Shows Understanding	% Does Not Show Understanding
1	*83*	*17*
2	*95*	*5*
3	*78*	*22*
4	*57*	*43*
5	*48*	*52*
6	*56*	*44*
7	*78*	*22*
8	*58*	*42*

The first two questions related directly to publishing personal information about Opaline, and they were supposed to be the "easy" questions that helped the children settle into the task. For both questions the majority of the children demonstrated some understanding of privacy. However for the first question one child stated "I am not sure because people put pictures of themselves and other people don't" demonstrating a lack of understanding.

Surprisingly, the second question revealed that one participant thought it was OK to advertise their home address. It is possible that the participant had misunderstood the question, however their written comment suggested it had been understood because they said it was OK to publish your home address just in case "others wanted to go and see her". This then suggested a potentially dangerous lack of understanding.

The third and first questions were very similar to each other and involved the uploading of photographs. The participants had clearly understood the privacy issue while the photograph contained Opaline, but the privacy issue was not so clearly understood when it contained a picture of the best friend, and as can be seen in table 2 only 78% of the sample said that the photograph should be kept private. 83% of the sample said Opaline should keep the photograph of herself private, but the children felt having permission from the best friend removed their privacy concern. The reasons given included "it's nice to have a little picture" and "people expose their best friends".

Question four asked people to rate her new iPhone. Here the participants were split and only a little over half demonstrated an understanding of the privacy concerns with publishing information about the mobile phone that you own. Most of the participants thought it was OK to publish this information, but participants were uncertain about the reason why with the reason distributed across safety, reputation and data. For example children gave reasons such as "she is just trying to show people what she has but it is a bit of a show off" and "she is not sharing any information".

Question five was about advertising her favourite band via a YouTube channel. At first glance this did not have many issues, however if the participants had taken the time to examine the image in more detail they would find that Opaline's name and where she lived (Mauritius) could be discovered. The analysts felt that there was a lack of understanding demonstrated in these answers, and this is the only question where the lack of understanding fell below the 50% margin. Most felt that it was OK to make it public, and their concerns were distributed quite evenly across all the different available concerns. This question seems to have caused confusion, with some commenting on the fact that it "wasn't dangerous to promote her favourite band", or that "people could make fun of her musical tastes" or that "the YouTube channel tells us her name and her whereabouts". Where the concern is not immediately obvious there were a wide range of concerns demonstrated.

Question six was about writing a book review and again there were a low number of participants demonstrating an understanding of privacy concerns. Most said it was OK to make public with comments quite commonly used like "it is only a book review" or "it is advertising David Walliams". Looking at the qualitative data suggested that the participants were not demonstrating a lack of understanding; but they were not demonstrating an understanding either.

Question seven was about advertising a party over social media; and this was probably the most concerning set of answers. Almost a quarter of the participants did not demonstrate an understanding of privacy, and with 13% of them saying that it was OK to make this information public. While some participants commented that "strangers can come and kill you!!!" others commented that it would enable "people to come to the party".

Question eight was about writing a review of a brand new smart watch. Opinions were divided with just over half demonstrating an understanding of the relevant privacy concerns; with most saying it is OK to make public and 16% unsure of whether to make it public or remain private.

4.2 Children's Concerns

The children's responses were analysed and the privacy concerns were categorised into one of four predefined themes. The results can be seen in Table 3. Children could identify safety concerns which were evident in the scenarios and the children were also concerned with their reputation. Most felt it was not OK for Opaline to upload a photograph of herself, but fewer felt it was not OK for Opaline to upload a photograph of her best friend without their permission. The participants did not have the same privacy concerns for Opaline as they did Opaline's best friend.

Table 3: *Distribution of Concern Ratings %*

Q	% Safety	% Reputation	% Permission	% Data
1	*55*	*36*	*5*	*4*
2	*99*	*0*	*1*	*0*
3	*21*	*9*	*63*	*7*
4	*25*	*29*	*7*	*39*
5	*29*	*17*	*24*	*29*
6	*13*	*21*	*10*	*56*
7	*82*	*5*	*5*	*7*
8	*33*	*18*	*6*	*43*

For half the questions (1,2,3,7) children did not have any concerns relating to personal data being shared. Some of the reasons relating to personal data included "she could get hacked".

4.1 Gender Differences

Gender did seem to make a difference with girls demonstrating a marginally better understanding of privacy than the boys.

Table 4: *Difference between genders*

Q	% Boy Understood	% Girl Understood
1	*74*	*74*
2	*94*	*98*
3	*56*	*70*
4	*36*	*48*
5	*34*	*48*
6	*47*	*55*
7	*55*	*91*
8	*39*	*58*

In table 4 you can clearly see that the boys almost always demonstrate a lesser understanding of privacy. It is quite disturbing that for questions 3 to 8 almost half of the boys regularly did not demonstrate an understanding of privacy.

While there were differences in how the girls and boys categorised the privacy concerns, these differences were mostly minor. Perhaps the most interesting was with the first question (uploading a photograph) where more boys considered the concern to be stranger danger, and more girls considered the concern to be about reputation.

The only real difference was in question 5 (favourite band on YouTube) which had a different spread across all concern types.

5 DISCUSSION AND CONCLUSIONS

Participants demonstrated their understanding of privacy clearest when they had safety concerns. The participants clearly understood that keeping some of their information private could have an impact on their safety.

This demonstrates that the message about 'stranger danger' is getting through to children.

When the concerns were not as clear cut as safety, the answers became more evenly distributed and it did not seem as though the participants always understood the context of the potential risk.

The GDPR recognises that a data subject's data is worth something; and it encourages those data subjects to take more care of their data by obliging data processors to seek informed consent. All of the scenarios presented in the activity workbook related directly to information and data that had been uploaded onto the Internet; however, this fact was not always recognised by the participants.

The data harvesting scandal facing Facebook during March 2018 demonstrates the power of collecting and aggregating user data. It is claimed that the data collected influenced both the 2016 US Presidential election and the 2016 Brexit referendum. By building profiles from the harvested data, they were able to produce targeted political advertising, to take advantage of a person's inner demon (Cadwalladr & Graham-Harrison, 2018).

Perhaps it is time to start educating children (as well as their parents and their teachers) more about the value of their data, in a digital world they are digital entities, with digital footprints and for their own benefits need to start managing their own data in a way that they will not regret at some point in the future.

5.1 Future Work

There is not much evidence of work being carried out about children and privacy, and the majority of this work does not tend to identify that privacy information is tacit information. Children are clearly capable of understanding privacy related matters especially where safety is concerned. We now need to move towards investigating these design issues, which will empower children to make informed decisions about their online digital footprint, and the associated privacy concerns that may arise from these things.

Children are the experts in being children. Adults are not children, no matter how hard they try to be; and as such children should be involved in the design issues that are evolving around online data privacy for children.

6. ACKNOWLEDGEMENTS

Thanks go to Kirkham St Michael's primary school for attending the MESS day and providing access to all of this interesting data.

5. REFERENCES

Adams, A., & Sasse, M. A. (2001). Privacy in multimedia communications: Protecting users , not just data. *People and Computers XV—Interaction without Frontiers.*, 1–17.

Angulo, J., & Ortlieb, M. (2015). "WTH..!?!" Experiences, reactions, and expectations related to online privacy panic situations. *Symposium on Usable Privacy and Security*, 19–38.

Bakos, Y., Marotta-Wurgler, F., & Trossen, D. R. (2014). Does Anyone Read the Fine Print? Consumer Attention to Standard-Form Contracts. *The Journal of Legal Studies*, *43*(1), 1–35. https://doi.org/10.1086/674424

Cadwalladr, C., & Graham-Harrison, E. (2018). Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. Retrieved March 25, 2018, from https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election

Conti, G., & Sobiesk, E. (2007). An Honest Man Has Nothing to Fear: User Perceptions on Web-based Information Disclosure. *Proceedings of the 3rd Symposium on Usable Privacy and Security - SOUPS '07*, 112. https://doi.org/http://doi.acm.org/10.1145/1280680.1280695

Cranor, L., Reagle, J., & Ackerman, M. (2000). Beyond concern: Understanding net users' attitudes about online privacy. In *The Internet Upheaval: Rasing Questions, Seeking Answers in Communications Policy* (pp. 47–70). Retrieved from https://books.google.co.uk/books?hl=en&lr=&id=wSfvdWIm3ykC&oi=fnd&pg=PA47&dq=related:PhmTcii7vjEJ:scholar.google.com/&ots=_WrVgJ6zK8&sig=y1_kykItnVOEzsMNIWBMwMBiRzE#v=onepage&q&f=false

Culnan, M. J., & Armstrong, P. K. (1999). Information Privacy Concerns, Procedural Fairness, and Impersonal Trust: An Empirical Investigation. *Organization Science*, *10*(1), 104–115. https://doi.org/10.1287/orsc.10.1.104

Dey, R., Ding, Y., & Ross, K. W. (2013). Profiling High-school Students with Facebook: How Online Privacy Laws Can Actually Increase Minors' Risk. *Proceedings of the 2013 Conference on Internet Measurement Conference*, 405–416. https://doi.org/10.1145/2504730.2504733

Druin, A. (1999). The role of children in the design technology. *Behaviour & Information Technology*.

Dwyer, C. A. (2009). The Inference Problem and Pervasive Computing. *SSRN Electronic Journal*, 1–11. https://doi.org/10.2139/ssrn.1508513

Flewitt, R., Messer, D., & Kucirkova, N. (2015). New directions for early literacy in a digital age: The iPad. *Journal of Early Childhood Literacy*, *15*(3), 289–310. https://doi.org/10.1177/1468798414533560

Gill, A. J., Vasalou, A., Papoutsi, C., & Joinson, A. N. (2011). Privacy Dictionary : A Linguistic Taxonomy of Privacy for Content Analysis. *CHI '11 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 3227–3236. https://doi.org/10.1093/pan/mps028

Hall, L., Hume, C., & Tazzyman, S. (2016). Five Degrees of Happiness: Effective Smiley Face Likert Scales for Evaluating with Children. *Proceedings of the The 15th International Conference on Interaction Design and Children*, 311–321. https://doi.org/10.1145/2930674.2930719

Henderson, S., & Yeow, J. (2012). iPad in Education: A Case Study of iPad Adoption and Use in a Primary School. *2012 45th Hawaii International Conference on System Sciences*, 78–87. https://doi.org/10.1109/HICSS.2012.390

Horton, M., Read, J. C., Mazzone, E., Sim, G., & Fitton, D. (2012). School friendly participatory research activities with children. In *Proceedings of the 2012 ACM annual conference extended abstracts on Human Factors in Computing Systems Extended Abstracts - CHI EA '12* (pp. 2099–2104). https://doi.org/10.1145/2212776.2223759

Horton, M., Read, J., & Sim, G. (2011). Making your mind up?: the reliability of children's survey responses. *British HCI 2911*, 437–438. Retrieved from http://dl.acm.org/citation.cfm?id=2305393

Hox, J.J.; Borgers, N. (2001). Item nonresponse in questioinnaire research with children. *Journal of Official Statistics*.

Information Commissioners Office. (2018). Guide to the General Data Protection Regulation: Children. Retrieved March 14, 2018, from https://ico.org.uk/for-organisations/guide-to-the-general-data-protection-regulation-gdpr/applications/children/

Krasnova, H., & Veltri, N. F. (2010). Privacy calculus on social networking sites: Explorative evidence from Germany and USA. *Proceedings of the Annual Hawaii*

International Conference on System Sciences, 1–10. https://doi.org/10.1109/HICSS.2010.307

Lee, K. (2000). Piaget's Theory of Cognitive Development. In *Childhood Cognitive Development: The Essential Readings*. Blackwell Publishes Ltd.

Livingstone, S., Davidson, J., & Bryce, J. (2017). Children's online activities, risks and safety A literature review by the UKCCIS Evidence Group, (October). Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/650933/Literature_Review_Final_October_2017.pdf

Luger, E., Moran, S., & Rodden, T. (2013). Consent for All: Revealing the Hidden Complexity of Terms and Conditions. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13*, 2687. https://doi.org/10.1145/2470654.2481371

Mann, A.-M., Hinrichs, U., Read, J. C., & Quigley, A. (2016). Facilitator, Functionary, Friend or Foe?: Studying the Role of iPads within Learning Activities Across a School Year. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*, 1833–1845. https://doi.org/10.1145/2858036.2858251

Metzger, M., Flanagin, A., & Nekmat, E. (2015). Comparative Optimism in Online Credibility Evaluation Among Parents and Children. *Journal of Broadcasting and Electronic Media*, *59*(3), 509–529. https://doi.org/10.1080/08838151.2015.1054995

Minkus, T., Liu, K., & Ross, K. W. (2015). Children Seen But Not Heard: When Parents Compromise Children's Online Privacy. *Proceedings of the 24th International Conference on World Wide Web*, 776--786. https://doi.org/10.1145/2736277.2741124

Pavlou, P. A. (2011). State of the Information Privacy Literature: Where Are We Now and Where Should We Go? *Misq*, *35*(4), 977–988.

Read, J., & Beale, R. (2009). Under my Pillow – Designing Security for Children's Special Things. *HCI 2009 - 23rd Annual Conference on Human-Computer Interaction*, 288–292.

Read, J. C. (2005). The ABC of CCI. *Interfaces*, *62*(Spring 2005), 8–9. Retrieved from http://www.bcs.org/upload/pdf/interfaces62.pdf

Read, J. C. (2007). Validating the Fun Toolkit: an instrument for measuring children's opinions of technology. *Cognition, Technology & Work*, *10*(2), 119–128. https://doi.org/10.1007/s10111-007-0069-9

Read, J., Gregory, P., MacFarlane, S., McManus, B., Gray, P., & Patel, R. (2002). An investigation of participatory design with children-informant, balanced and facilitated design. *Interaction Design and Children*, (JANUARY), 53–64. https://doi.org/10.1.1.109.1324

Schechter, S. (2013). The User IS the Enemy, and (S)he Keeps Reaching for that Bright Shiny Power Button! *Proceedings of the Workshop on Home Usable Privacy and Security (HUPS)*. Retrieved from http://research.microsoft.com/apps/pubs/default.aspx?id=194484

Shmueli, B., & Blecher-Prigat, A. (2011). Privacy for Children. *Columbia Human Rights Law Review*, *42*(3), 759–795. https://doi.org/10.1525/sp.2007.54.1.23.

Silva, C. S., Silva, I. S., Silva, T. S., & Mourão, F. (2017). Privacy for Children and Teenagers on Social Networks from a Usability Perspective : A Case Study on Facebook, 63–71. https://doi.org/10.1145/3091478.3091479

Smith, H. J. (2014). Information Privacy Research : An Interdisciplinary Review . I NFORMATION P RIVACY R ESEARCH : A N I NTERDISCIPLINARY R EVIEW 1, (December 2011).

Solove, D. J. (2006). A Taxonomy of Privacy. *University of Pennsylvania Law Review*, *154*(3), 477–564.

Steinberg, S. B. (2017). Sharenting: Children's Privacy in the Age of Social Media. *Emory Law Journal*, *66*, 839–884.

Stothard, S. E., & Hulme, C. (1992). Reading comprehension difficulties in children. *Reading and Writing*, *4*(3), 245–256.

To, A., Fan, A., Kildunne, C., Zhang, E., Kaufman, G., & Hammer, J. (2016). Treehouse Dreams: A Game-Based Method for Eliciting Interview Data from Children. *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts*, 307–314. https://doi.org/10.1145/2968120.2987718

Westin, A. (1967). *Privacy and Freedom*. Ig Publishing.

Zhang-Kennedy, L., Baig, K., & Chiasson, S. (2017). Engaging Children About Online Privacy Through Storytelling in an Interactive Comic, 1–12. Retrieved from http://chorus.scs.carleton.ca/wp/wp-content/papercite-data/pdf/zhang-kennedy2017childrencomics-bhci.pdf

Zhang-Kennedy, L., Mekhail, C., & Chiasson, S. (2016). From Nosy Little Brothers to Stranger-Danger : Children and Parents ' Perception of Mobile Threats. *ACM SIGCHI Conference on Interaction Design and Children 2016*, 388–399. https://doi.org/10.1145/2930674.2930716