Growing Up With Dark Patterns: How Children Perceive Malicious User Interface Designs

René Schäfer
rschaefer@cs.rwth-aachen.de
RWTH Aachen University
Aachen, Germany

Sarah Sahabi
sahabi@cs.rwth-aachen.de
RWTH Aachen University
Aachen, Germany

Annabell Brocker
a.brocker@cs.rwth-aachen.de
RWTH Aachen University
Aachen, Germany

Jan Borchers
borchers@cs.rwth-aachen.de
RWTH Aachen University
Aachen, Germany

Figure 1: In one part of our study, we asked fifth-graders (10–11 years old) to create cookie banners that would nudge people towards accepting cookies. For this, we gave them a reference cookie banner (see Figure 4 (F)) and asked them to redesign it while keeping the same goal. Children chose to, inter alia, use monetary compensation (P07), security promises (P10), false compromises (P13), different button sizes and colors (P38), unfair compromises (P42), smileys (P52), disguising the reject option as regular text (P57), reduced waiting times (P65), and forced decisions (accept or leave the application) (P66). Translations are provided in Figure 7 (Appendix A).

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
NordiCHI 2024, October 13–16, 2024, Uppsala, Sweden

© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 979-8-4007-0966-1/24/10
https://doi.org/10.1145/3679318.3685358

ABSTRACT

Apps and websites increasingly employ *dark patterns*, malicious interface design strategies that nudge people towards making decisions against their best interests. So far, dark patterns research has focused almost exclusively on adults. Today, however, children grow up with easy access to apps and online content, and they are particularly vulnerable to manipulation. Therefore, we aim to better understand how dark patterns impact children.

To this end, we conducted a triangulated elicitation study at a German school with 66 fifth-graders (10–11 years old) to start understanding how they perceive dark patterns. We found that many children understood the intentions behind simple dark patterns. When asked to actively search for manipulations, about half noticed overly complex wordings and color-based manipulations. About every fourth child spotted manipulative formulations. Most, however, completely missed *Bad Defaults* nudging them towards sharing personal data. This indicates that children may be particularly susceptible to bad privacy defaults.

CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in HCI**; • **Social and professional topics** → **Children**; • **Security and privacy** → *Social aspects of security and privacy*; • **Applied computing** → Psychology.

KEYWORDS

dark patterns, children, online manipulation, study

ACM Reference Format:
René Schäfer, Sarah Sahabi, Annabell Brocker, and Jan Borchers. 2024. Growing Up With Dark Patterns: How Children Perceive Malicious User Interface Designs. In *Nordic Conference on Human–Computer Interaction (NordiCHI 2024), October 13–16, 2024, Uppsala, Sweden.* ACM, New York, NY, USA, 17 pages. https://doi.org/10.1145/3679318.3685358

1 INTRODUCTION

The number of children using smartphones and accessing the internet has grown rapidly over the last decade [41, 44]. This has also increased their exposure to deceptive designs, which have become increasingly common on websites [20], in mobile apps [10, 13], and in games [52], and thus also increased the likelihood of children falling for such manipulations [12, 46]. As children are particularly vulnerable to media effects [48], it is crucial to investigate the impact so-called *dark patterns* have on children.

Dark patterns are malicious user interface design strategies that nudge people towards making decisions that may go against their best interests [31]. In cookie banners, for example, the visual emphasis of the *"Accept all cookies"* button while the *"Decline"* button looks greyed out is an application of the dark pattern *False Hierarchy* [16] and is supposed to nudge users into clicking this button. Another example of a dark pattern is *Confirmshaming* — using formulations that aim to make users feel ashamed when choosing the option that is less favorable for the owner of that particular service *("No thanks, I like paying full price").* Recently, research interest in this field has increased substantially [28]. Current research topics include legislation [27], automatic detection [31], privacy-friendly bright patterns [14], visual countermeasures [43], unifying pattern names and definitions [17], and user awareness [3]. Most studies targeting the impact of dark patterns on people were conducted with adults (e.g., [3, 10]). While this helps to learn more about malicious designs in general, this knowledge does not necessarily transfer to children. Because of this, we focus on investigating how *children* perceive dark patterns and the malicious intentions behind them. The concrete research questions for this work are:

- RQ1: Do children visually perceive interfaces containing dark patterns more negatively compared to a fair design?
- RQ2: Can children deduce how manipulative interfaces influence people?
- RQ3: Which dark patterns do children use when designing manipulative interfaces themselves?
- RQ4: After being made aware of dark patterns and manipulation, how well can children spot dark patterns themselves?

To answer these questions, we conducted a study with 66 fifth-graders (10–11 years old) at a school in Germany. To better grasp children's understanding and knowledge base of dark patterns and manipulative designs, we triangulated several techniques commonly applied in research with children: *drawings,* as proposed by Doyle et al. [11], as well as *recognition* and *free-recall* tasks [36] targeting their understanding and awareness of the topic.

With our work, we want to highlight the importance of researching the impact of manipulative designs on children and motivate researchers, practitioners, and teachers to facilitate the education of children towards sensitizing them against the influence of such manipulations.

2 RELATED WORK

For an overview of the research relevant to studying children's perception and understanding of dark patterns, we first discuss related work on user awareness of dark patterns for adults and general information on children's data security and privacy awareness. Afterward, we review common techniques to elicit mental models, focusing on methods suitable for children.

2.1 Dark Pattern Awareness

The term *dark pattern* was introduced by UX researcher Brignull on his website[1] in 2010. That same year, Conti and Sobiesk [8] investigated users' self-reported frustration with and tolerance of common malicious interface designs, including spoofed interface elements, forced waiting, and coerced registration. Results suggested that users found all these malicious designs significantly frustrating. However, depending on the context and the task they were trying to accomplish, they demonstrated varying levels of tolerance towards such designs. For example, a higher tolerance for adult users regarding frustration was shown in gaming, shopping, and pornographic applications, while the lowest tolerance was shown on search, news, weather, and vendor support sites.

Moreover, Luguri and Strahilevitz [27] observed that the intensity of manipulation impacted users' attitudes. For example, dark patterns they classified as *aggressive* received fierce backlash and annoyance, while *mild* patterns did not. The researchers also identified differences between these two pattern types (aggressive and

[1] https://www.deceptive.design (formerly https://www.darkpatterns.org), *last accessed July 31, 2024*

mild) concerning *preference inconsistency*, i.e., the effectiveness of the manipulation to nudge a person into making choices they would not have made themselves [29]. While mild patterns were two times more effective than the standard user interface from the control conditions, aggressive patterns were four times more effective [27]. Furthermore, different user groups exhibited varying degrees of vulnerability to dark patterns. For instance, elderly and less educated users showed higher vulnerability. Time pressure also notably increased manipulation rates [29].

When assessing how accurately users could detect manipulative designs on apps and websites, findings diverge: Gray et al. [15] found that 79.3% of their participants were able to correctly detect designs that had been built to manipulate users. In contrast, other studies suggested that users frequently have difficulties recognizing them [10, 22, 29]. For example, Di Geronimo et al. [10] reported a recognition rate of only 25%. Subsequent investigations suggested that a primary reason for this poor detection was the prevalence of dark patterns, which had made them part of everyday interactions so that they passed unnoticed [10]. When exploring the relationship between users' dark pattern awareness and their capabilities to detect them, researchers also report contradictory findings: Keleher et al. [22] observed that users were still struggling with identifying patterns even when they were aware of them. In their experiment, however, participants only received a theoretical definition of dark patterns without further attempts to familiarize them with the topic, e.g., through practical examples. Hence, it remains unclear whether the observed effects validly reflect the relationship between awareness and detection or whether they stemmed from unsuccessful awareness-raising in the experiment. Indeed, Di Geronimo et al. [10] found that, once participants became more aware of and more knowledgeable about dark patterns, they performed better at detecting them [10] and were more likely to resist them [19].

Similarly, related work examining how user awareness impacts manipulation resistance yields no consistent results. While early findings suggest that awareness increases resistance [19], Bongard-Blanchy et al. [3] found that user awareness would not always induce a change in behavior or resistance. As the authors speculate, this could be traced back to the fact that users were unaware of the actual harm and dangers that emanated from the manipulations. Consequently, one approach to counteract dark patterns would be to educate users not only about dark patterns in general but also about possible risks and how to resist them. This potential of awareness has already been established in related disciplines, such as information security awareness: There, higher awareness effectively helps users identify potential vulnerabilities and avoid falling into the trap of disclosing sensitive data [1, 40].

2.2 Security and Privacy

Ensuring the security and privacy of children's data is an important issue. The proliferation of digital devices among children, coupled with their increased access to the internet [41, 44], introduces numerous potential avenues for privacy breaches. Andrews et al. [2] cited a study[2] from 2014 indicating that approximately two-thirds of children have not optimized their privacy settings to safeguard their data effectively. Moreover, around 50% of children consistently enabled GPS tracking, exposing their real-time locations, while 14% disclosed their home addresses online [2]. Another study from 2017[3] revealed that over half of the applications commonly utilized by children exhibited deficiencies in user data protection. Oates et al. [35] let children draw *privacy* and found that children under the age of ten might not consider digital spaces when thinking about privacy. White et al. [51] demonstrated that individuals aged 13 to 17 are notably more inclined to divulge their data online compared to their counterparts aged 18 to 25. One explanation for this can be derived from the description–experience gap by Hertwig and Erev, which argues that individuals tend to make decisions based on their accumulated experiential knowledge. Consequently, a greater propensity to favor decisions aligned with positive experiences arises when an individual has encountered fewer negative instances thus far [21]. Another explanation stems from the natural developmental trajectory, wherein a specific brain region integral to inhibitory control and objective risk assessment undergoes maturation towards the conclusion of childhood [37]. Consequently, children often exhibit a proclivity for engaging in risky decision-making due to the incomplete development of this crucial neural substrate [37].

Furthermore, a parallel evolutionary rationale, characterized by the rapid development and remodeling of various brain areas during childhood, elucidates the heightened impulsivity, self-confidence, and susceptibility to risk in children compared to adults [37]. Consequently, this increased vulnerability renders children more prone to potential harm [37]. Andrews et al. [2] investigated various approaches, such as an educational video or a quiz with feedback, to enhance children's awareness of their data security and privacy. They found that quizzes and comprehensive feedback mechanisms increased children's awareness and sensitivity and thus reduced their tendency to disclose data. This outcome, among others, underlines the importance of raising children's awareness regarding data security and privacy from an early age, thus making them less likely to disclose data unintentionally [1, 2, 23, 40]. Some initiatives have begun to establish such measures. Examples include the *Media Literacy Framework North Rhine-Westphalia*[4] and the *Media Education Orientation Framework Lower Saxony*[5], which define the media literacy curriculum for schools in different German states. To increase children's awareness of relevant dark pattern-related issues, it is essential to effectively introduce and enhance such curricular initiatives over the long term. This requires a comprehensive understanding of how children react to and perceive certain risks.

[2] https://www.businesswire.com/news/home/20140602006734/en/Cyberbullying-Triples-According-to-New-McAfee-"2014-Teens-and-the-Screen-Study", *last accessed July 31, 2024*
[3] https://www.washingtonpost.com/news/the-switch/wp/2017/07/27/we-tested-apps-for-children-half-failed-to-protect-their-data/, *last accessed July 31, 2024*
[4] https://medienkompetenzrahmen.nrw/, *last accessed July 31, 2024*
[5] https://bildungsportal-niedersachsen.de/fileadmin/2_Portale/Medienbildung/medienbildung_vorgaben/Orientierungsrahmen_Medienbildung_Niedersachsen.pdf, *last accessed July 31, 2024*

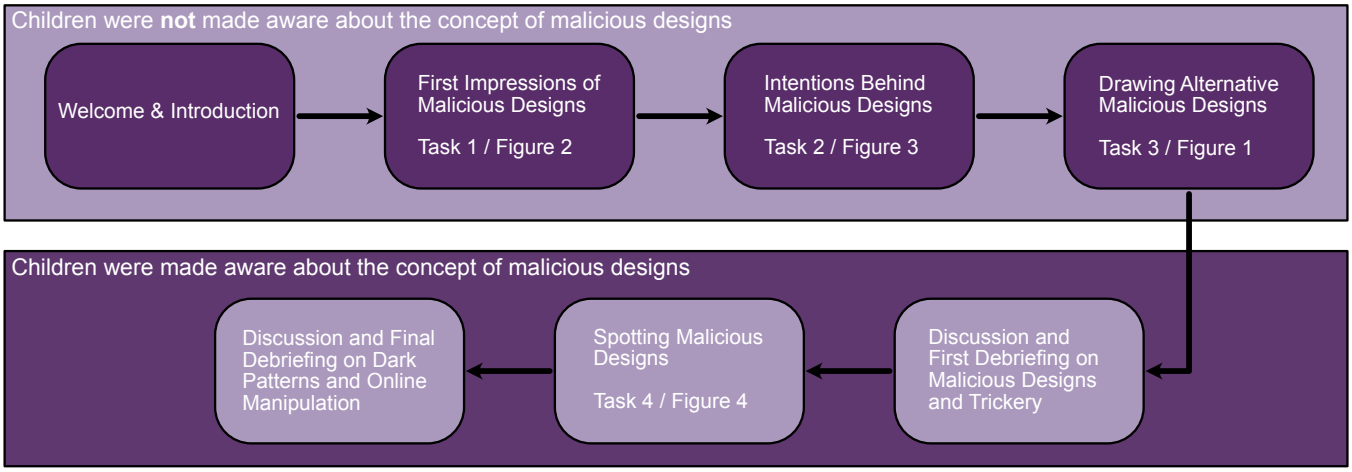

Figure 2: The schedule for our study. Overall, the study consisted of an introduction, four tasks, and two discussions & debriefings. Until the first discussion, we did not make children aware of the concept of dark patterns (tasks in the top row).

2.3 Research Methods with Children

To better understand how children perceive and understand dark patterns, we make use of several techniques commonly used in research with children. The most commonly used technique to elicit mental models from people is interviewing, which is relatively easy to implement and analyze [11]. However, it may produce inaccurate and incomplete models when used with children, as their lack of self-awareness [24, 30] and appropriate terminology [9, 47], as well as potential stress due to the test-like atmosphere of interviews, often prevent them from expressing their mental models appropriately [9, 47].

Drawing is a technique that solves these issues and is used frequently with children [11]. Mental models of concrete, substantial constructs can be elicited directly from drawings (e.g., what the inside of a computer looks like under the hood [9]). Corresponding drawing instructions must be designed with care and formulated precisely, as they can yield different mental models [24, 36]. Since some children may lack the manual dexterity to depict what they want to express [30], another common technique is to include textual annotations in the drawings [38].

However, Panagiotaki et al. [36] criticize that drawings only reveal children's naïve mental models. They propose using recognition tasks instead of free-recall techniques for scientifically more accurate responses. Such recognition tasks include the *Arrange Cards* technique [30], in which participants are asked to spatially arrange pieces of concepts written on a set of cards in a way that matches their representation of a construct.

Overall, Grenier and Dudzinska-Przesmitzki [18] conclude that triangulating insights using multiple elicitation techniques is a solid approach to counteract the challenges that come with each method. We used this approach in our study design by using *drawing, recognition* and *free-recall* tasks.

3 STUDY

To investigate our research questions, we conducted an elicitation study with fifth-graders, most aged 10 to 11 years old, at a local school in western Germany. We ran a pilot study with three children of similar age to test the scope of the study and whether children were able to understand the given tasks. We then adapted our study design accordingly. During the full study, we collected quantitative and qualitative data. Overall, the study contained four tasks and two discussion & debriefing sessions. Figure 2 shows the order in which these were scheduled. Since we conducted the study in German, we translated all screens for this paper in Figures 3–5 for better understandability. The original screens can be found in Appendix B.

3.1 Tasks

We focused on minimizing the cognitive load for the children. Following the recommendation by Mertala [33] and Punch [39], we designed the tasks around a playful theme. For this, we placed all tasks in the context of a fictional mobile game as touchscreen interfaces are well-known environments for children [5]. We created several "screenshots" of this game and included common manipulations from mobile games in them, such as the option to watch an advertisement to skip waiting times and a lock-in mechanism when a user wants to quit the game. For the sake of consistency, we based all dark pattern names on the ontology by Gray et al. [17]. All tasks were formulated using simple language to make them easier for the children to understand, and they were completed using pencil and paper. Before the study, the participating children were told that our study was about children's media literacy.

3.1.1 Task 1: Initial Impressions. With **Task 1**, we aimed to gather insights into children's first impressions when seeing interfaces that contain different dark patterns (RQ1). Participants received four similar screenshots of a fictional mobile game with varying numbers and intensities of dark patterns (see Figure 3). To counteract order effects, we used a balanced 4x4 Latin square for the "screenshots".

We wanted to incorporate dark patterns into our designs that children would have likely encountered before, to include potential effects due to prior experience. For this, we compared statistics

Figure 3: For our tasks, we created four different designs of a fictional mobile game containing different intensities of dark patterns: *Fair (A)* is a fair design without dark patterns. *Confirmshaming (B)* uses an influential formulation in the buttons. *False Hierarchy (C)* has one button less compared to A and B. Additionally, it makes the Close button in the corner less visible. *Multiple (D)* combines *False Hierarchy* with additional *Confirmshaming* below the button. For Task 1, children rated all designs on three scales (beauty, complexity, and trustworthiness). Task 2 was about the visual and conceptual differences between *Fair (A)* and *False Hierarchy (C)*. The original versions that the children received during the study are provided in Figure 8 (Appendix B).

about the most popular mobile apps among German 10- to 11-year-olds [41] and the dark patterns that occur most prominently in those apps [10, 13]. Because of the prevalent role mobile devices play in children's social lives [10, 41], we decided to only consider mobile scenarios. Furthermore, we limited our selection to dark patterns that were representable by one static and non-interactable screenshot each, excluding manipulations like *Nagging*, which usually occur over time. Finally, to test whether the intensity of the dark pattern would affect children's initial impressions, we chose one mild and one aggressive dark pattern out of our selection, according to the classification by Luguri and Strahilevitz [27]. This yielded the following designs: *Confirmshaming (B)*, which contained manipulative formulations within the buttons, creating a mild dark pattern, and *False Hierarchy (C)*, where the reject button was removed and the Close button was altered, as an aggressive dark pattern. To cover a wider range of intensities, we further added *Fair (A)*, which contained no dark patterns, and *Multiple (D)*, which combined the mild *Confirmshaming (B)* and the aggressive *False Hierarchy (C)*.

Children rated each screen according to perceived *beauty* (not beautiful to very beautiful), *trustworthiness* (not trustworthy to very trustworthy), and *complexity* (not complex to very complex) using 21-point semantic-differential ratings. We utilized an unusually high number of tick marks for the semantic differential scale to approximate a continuous scale. This approach increased the granularity and discrimination of participants' responses, allowing children to express stronger opinions without overwhelming them with indefinite options [7].

We based these three scales on the three essential factors to judge websites, identified by Lindgaard et al. [26]: *visual appeal*, *perceived usability*, and *trustworthiness*. A teacher was present and helped ensure that the children understood the task and the terminology.

3.1.2 Task 2: Understanding Malicious Intents. In **Task 2**, we wanted to explore children's abilities to understand malicious designs and deduce the intentions behind using them (RQ2). This task was split into three parts:

(1) The children marked and described differences between *Fair (A)* and *False Hierarchy (C)* (see Figure 3).
(2) They were asked which of the two designs they would choose to make more users watch advertisements and justify their decision in writing.

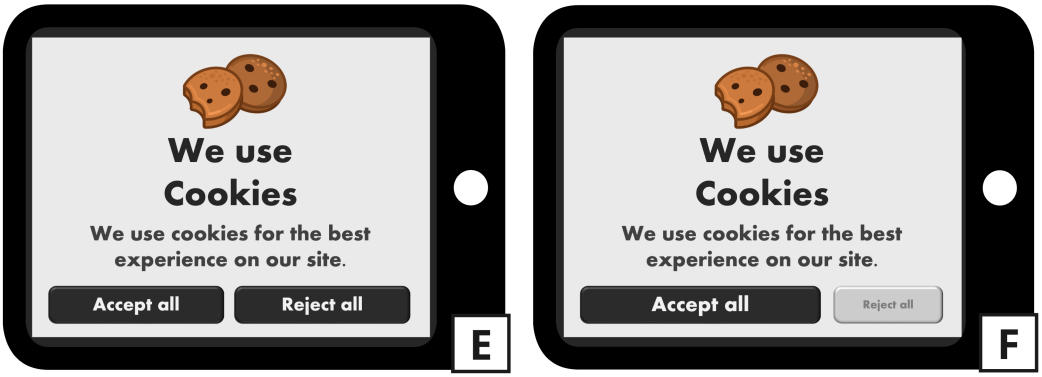

Figure 4: The cookie banners used in Task 2. The left design *(E)* provides two fair choices. The right design *(F)* uses *False Hierarchy* to nudge people towards accepting cookies. Children were asked to deduce the goal behind using *(F)* over *(E)*, which 81.5% accomplished successfully. Cookie illustration provided by Alexandr Martinov. The original versions that the children received during the study are provided in Figure 9 (Appendix B).

(3) They were presented with two different versions of cookie banners (see Figure 4): *Fair (E)* without dark patterns and one using *False Hierarchy (F)* to make users accept cookies. The children were asked to speculate which goal one could achieve by choosing *False Hierarchy (F)* over *Fair (E)* and to justify their answer in a text box.

We included the first part of this task to draw the children's focus to the manipulative parts of the design. Therefore, our goal was not to see whether children could spot dark patterns but rather if they could understand the implications of using malicious designs.

3.1.3 Task 3: Drawing Cookie Banners. In **Task 3**, we used drawings to investigate whether children already intuitively incorporate manipulative dark patterns themselves when designing cookie banners (RQ3). For this, we asked children to think of an alternative design for *False Hierarchy (F)* that would achieve the same goal they identified in the previous task. We let children draw their redesigns for this task, following Doyle et al.'s elicitation method for children's mental models [11]. In addition to their drawing (or if they did not want to draw), children could explain their thoughts in a text box.

3.1.4 Task 4: Finding Manipulations. **Task 4** explored whether children can spot manipulative designs (RQ4) after being educated about the topic of deceptive designs (Section 3.2). The task showed an image of a screen in which a user wants to quit the game from the first task (see Figure 5). This image contained four dark patterns:

(1) a *Bad Default* in form of a *preselection* of an option that shares data,
(2) a *Confirmshaming* in the other option stating that the user played too badly to share their score with others,
(3) a *Trick Question* using a double-negation in the text of the cookie banner,
(4) and *Emotional and Sensory Manipulation*, which used switched color conventions for the buttons to trick the user into clicking the wrong button and keep playing. Here, green keeps the user in the app, while red quits the app.

The task clearly stated that the image contained four manipulations, and children were asked to find these and then to justify their selection. By providing a number, we aimed to prevent children from simply marking everything and instead have them focus on the elements that were most likely to be manipulative.

3.2 Discussions and Debriefings

Overall, there were two sessions of discussion & debriefing together with all children. The **first session** took place **before Task 4**. In that session, we discussed their answers and opinions on all previous tasks. We concluded that deceptive design tricks could be used to manipulate people in their decision-making by, e.g., the look of a given website or app. We demonstrated this using the deceptive designs from previous tasks without naming any specific dark patterns. This discussion was needed because children had to actively search for manipulative designs in **Task 4**, and we did not make children aware of the context of malicious designs before.

The **second discussion session** took place **after Task 4** and completed the study. We used this time to educate children on dark patterns and manipulative designs, highlight possible threats, and explain what they could do to protect themselves. Finally, children could ask questions and talk about their personal experiences.

3.3 Study Procedure

The schedule of our study is shown in Figure 2. We conducted the study three times, each time with a different class of fifth graders (10–11 years old) from the same school. The study took place in the regular classroom of each class and was co-supervised by their class teacher. With this familiar setting, we wanted to reduce possible stress and anxiety. Also, the teacher had the pedagogical expertise to intervene in case of difficult situations, e.g., if children felt too much pressure. In addition, at least two researchers were present for each class. This helped us answer upcoming questions quickly without leaving the other children unsupervised. Together with the teacher, we embedded the study in a regular double school lesson of each class. This way, every child participated, although

Figure 5: In Task 4, children had to spot four different dark patterns on the given screen. Overall, 47.1% of the children spotted the *Trick Question* in the cookie banner and *Emotional and Sensory Manipulation* regarding the button colors. 25.5% realized the *Confirmshaming* in the text of the second checkbox. Only three children (5.9%) found the data-insecure *Bad Defaults* in the first checkbox. The original version that the children received during the study is provided in Figure 10 (Appendix B).

we only analyzed answers for which we were given explicit consent as described in Section 3.4.

During the study, the children received only one task at a time, which enabled us to control the pace of the study and make sure that everyone could take the time they needed. We explicitly did not introduce children to the concept of malicious designs for **Tasks 1–3** to not bias their first impressions and thoughts when looking at our "screenshots". **After Task 3**, we carried out the first debriefing & discussion session. During this time, we discussed the children's thoughts and ideas about the previous tasks, including the implications of the manipulative elements presented. Explicitly bringing up terms like "manipulation" and "trickery" was essential, as the last task required the children to search for manipulations on a specific screen (see Figure 5). **After Task 4**, there was a final discussion round on the manipulations the children had found and how these designs influence people. To provide the children with educational compensation for their time, we closed the study with a debriefing on dark patterns and online manipulations and discussed how one could detect and resist them. Overall, the study took approximately 90 minutes, including breaks, which matches the duration of regular double lessons in that school.

3.4 Ethical Considerations

Studies with children have to be designed and conducted with extra care [39]. We followed the *ACM Code of Ethics*[6] and the standards established by the *Ethical Research Involving Children*[7] project to guarantee a respectful and ethical treatment of all children in our study as our institution does not have an internal review board. Furthermore, we worked closely with a social worker from the participating school. Following this, we only evaluated answers from children who gave explicit consent themselves, in addition to their parents or legal guardians. Children could skip any task, and answers were anonymized to protect the identity of all participating children. To create a pleasant environment, we tried to minimize stress among the children using, inter alia, a playful design for our questionnaire [33, 39]. Additionally, a teacher the children knew was present throughout the study to reduce stress and anxiety. Finally, all children received a debriefing regarding dark patterns and online manipulation to strengthen their ability to recognize such malicious designs themselves better in the future.

3.5 Participants

Overall, 66 children participated in our study. Most children were 10–11 years old ($M = 10.5, SD = 0.5$), with only one child being 12 years old. 56.1% identified as female, 43.9% as male. The self-reported smartphone and tablet usage was: *at least once a day* (68.2%), *several times per week* (22.8%), *at most once per week* (4.6%), *not at all* (3%), and *no answer given* (1.4%). School officials told us that the children had some knowledge regarding media competence and internet threats such as cyberbullying. However, manipulation, and in particular dark patterns, had not been part of the syllabus.

4 RESULTS

We first analyzed all qualitative answers from tasks 2–4 using thematic analysis to develop an in-depth understanding of the children's understanding and perception. To create the codes and themes, we followed the six steps proposed by Braun and Clarke [4]. For this, one researcher inductively coded all answers using the

[6]https://www.acm.org/code-of-ethics, *last accessed July 31, 2024*
[7]https://childethics.com/ethical-guidance/, *last accessed July 31, 2024*

software MAXQDA[8]. Over three rounds of coding, we combined relatable codes into themes and, where applicable, deductively named them after a matching dark pattern from the ontology by Gray et al. [17]. Afterward, each answer and its respective codes and themes were reviewed and discussed in depth with a second researcher until an agreement was reached. The final coding of all answers was adapted accordingly. In another iteration, we proceeded with content analysis [32] to explore relationships within the data. For this, we quantified the data by counting occurrences of our codes and themes. For the following sections, we translated all participant comments from German into English and placed quotes in quotation marks followed by the anonymized participant ID (e.g., [P42]). For each task, we retrospectively excluded participants if they had not completed the task or clearly had not understood the instructions correctly. We describe exclusion criteria in the respective sections. Reported percentages are always based on the number of included answers.

4.1 Task 1: Initial Impressions

In the first task, we asked the children to spontaneously judge four given designs (see Figure 3) regarding beauty, trustworthiness, and complexity. All 66 children completed this task. Figure 6 displays their responses to these 21-point semantic differential ratings, which spanned the entire range from −10 to 10, indicating clear differences in how children perceived the designs. Friedman tests revealed significant effects of the designs on perceived beauty ($\chi^2(3) = 11.58, p < 0.01$), complexity ($\chi^2(3) = 17.83, p < 0.001$), and trust ($\chi^2(3) = 28.57, p < 0.001$). We used Wilcoxon Signed Rank tests with a Holm correction as post-hoc tests.

4.1.1 Beauty. Children perceived the dark pattern-free *Fair (A)* to be significantly more beautiful compared to *False Hierarchy (C)* ($p < 0.05$) and *Multiple (D)* ($p < 0.05$). On average, *Fair (A)* achieved the highest beauty scores ($M = 1.2$, $SD = 3.8$, $Mdn = 0$, $Mode = 0$). The scores were distributed from −6 to 10, indicating a trend towards higher beauty scores for *Fair (A)*. In contrast, the design with the aggressive *False Hierarchy (C)* was rated the least beautiful ($M = -1.5$, $SD = 4.9$, $Mdn = -1$, $Mode = -7, 0$). For this design, the responses are scattered more towards the negative half of the scale, ranging from −10 to 8. Moreover, they show a bimodal distribution, with peaks at −7 (less beautiful) and 0 (neutral). The *Confirmshaming (B)* design was, on average, rated as the second most beautiful ($M = 0.1$, $SD = 4.7$, $Mdn = 0$, $Mode = 0$), while the *Multiple (D)* design with a combination of both dark patterns was placed third ($M = -1.2$, $SD = 4.1$, $Mdn = -1$, $Mode = 0$).

4.1.2 Trust. Here, children found *Fair (A)* to be significantly more trustworthy than *False Hierarchy (C)* ($p < 0.001$) and *Multiple (D)* ($p < 0.01$). We identified a comparable distribution of judgments for the trustworthiness scale: The highest average scores were achieved for *Fair (A)* ($M = 2.1$, $SD = 4.3$, $Mdn = 2$, $Mode = 0$). Also, the aggressive *False Hierarchy (C)* gained the lowest scores ($M = -2.4$, $SD = 4.8$, $Mdn = -2$, $Mode = 0$), ranging from −10 to 8. Similar to the beauty scale, *Confirmshaming (B)* was perceived as the second-most trustworthy ($M = 0.3$, $SD = 4.9$, $Mdn = 0$, $Mode = 0$), and *Multiple (D)* with two dark patterns as the third ($M = -1.4$, $SD = 4.8$, $Mdn = -0.5$, $Mode = 0$). Notably, all four distributions show a clear peak at 0. However, both *Fair (A)* and *Confirmshaming (B)* are distributed more towards the positive half of the scale, while *False Hierarchy (C)* and *Multiple (D)* lean more towards the negative half.

4.1.3 Complexity. Children rated *Fair (A)* to be significantly less complex than *Confirmshaming (B)* ($p < 0.01$) and *Multiple (D)* ($p < 0.01$). Additionally, *False Hierarchy (C)* was seen as significantly less complex than *Multiple (D)* ($p < 0.05$). We discovered a ranking that deviated from the other two distributions: *Multiple (D)* was perceived as the most complex design. It obtained the highest but slightly negative complexity score ($M = -0.1$, $SD = 5.5$, $Mdn = 0$, $Mode = 0$). On the other hand, *Fair (A)* was perceived as least complex ($M = -3.5$, $SD = 5.4$, $Mdn = -5$, $Mode = -10, 0$). Its bimodal distribution, with peaks at −10 (not complex at all) and 0 (neutral), indicates a stronger scattering on the negative half of the scale. *Confirmshaming (B)* was judged the second-most complex ($M = -0.7$, $SD = 5.5$, $Mdn = 0$, $Mode = 0$). *False Hierarchy (C)* was placed third ($M = -2.1$, $SD = 5.8$, $Mdn = -2$, $Mode = -7, 0$) with a bimodal distribution, peaking at −7 (less complex) and 0 (neutral).

4.1.4 Summary. In all, our results show that children spontaneously judged dark pattern-free, fair designs significantly more favorably. In contrast, designs that contained aggressive *False Hierarchy* dark patterns were perceived most negatively. Only regarding complexity, the mild *Confirmshaming* was rated more poorly than the aggressive *False Hierarchy*. *Multiple (D)* was rated third regarding beauty and trust and was perceived as the most complex design.

4.2 Task 2: Understanding Malicious Intents

In the second task, we asked children first to highlight the differences between two designs (*Fair (A)* and *False Hierarchy (C)* from Figure 3), then choose the one which they thought would lead to more people watching an ad, and then justify their decision. Overall, the designs differed in the position and style of the Close button and the number of large buttons.

4.2.1 Visual Differences. 51 of the 66 children (77.3%) found all differences between the given designs. Ten children (15.2%) missed the Close option in *False Hierarchy (C)* and stated that there was no option to decline watching an ad, so people had no choice for this design: *"Some people do not like ads and still have to watch them. With the left one [*Fair (A)*], one can decide" [P22]*. Another child mentioned that *False Hierarchy (C) "does not allow a decision and is not trustworthy" [P58]*. Six children (9.1%) only spotted the different positions of the Close buttons, while eight children (12.1%) overlooked them and only realized that one of the larger buttons was missing in *False Hierarchy (C)*.

4.2.2 Picking a Design. When choosing a design that makes people watch an ad, 11 children (16.7%) provided answers that clearly did not match the task description: *"If you made everyone watch an ad, that wouldn't be so good. That's why I'd rather let people choose because otherwise, you're blackmailed" [P14]*. With this, the following percentages are based on **answers from 55 children**.

51 out of these 55 children (92.7%) opted for *False Hierarchy (C)*, and four (7.3%) thought that *Fair (A)* was the better choice. 24 times (43.6%), children argued with concepts related to *False Hierarchy*,

[8]https://www.maxqda.com, *last accessed July 31, 2024*

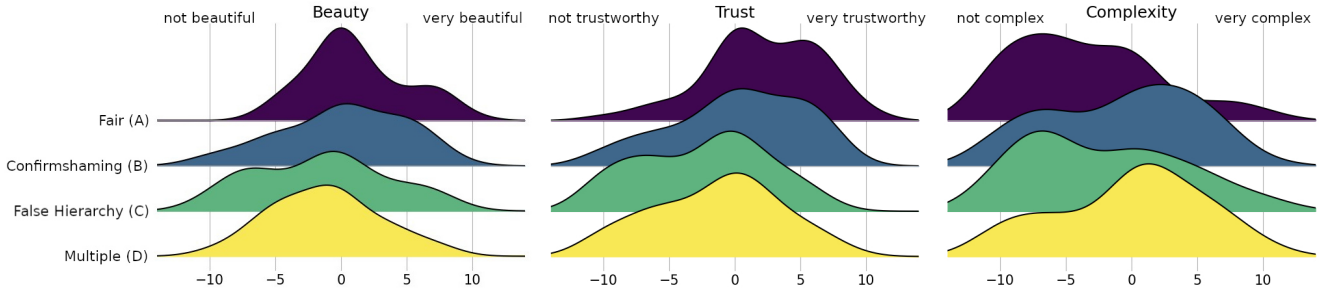

Figure 6: Ridgeline plots showing the distributions of the children's ratings for four screens—*Fair (A)*, *Confirmshaming (B)*, *False Hierarchy (C)*, and *Multiple (D)*—regarding three scales (*beauty*, *trust*, and *complexity*). Overall, *Fair (A)* and *Confirmshaming (B)* were perceived as more beautiful and trustworthy compared to *False Hierarchy (C)* and *Multiple (D)*. Regarding perceived complexity, *Fair (A)* and *False Hierarchy (C)* were rated as being rather simple, while *Confirmshaming (B)* and *Multiple (D)* were perceived more complex.

e.g., they stated that the Close option was barely visible or that the button for watching an ad was rather large: *"Because one can easily overlook the cross and just the one large choice 'Watch Ad' is seen" [P13]*. Overall, children mentioned the barely visible Close button 19 times (34.5%), while the large eye-catching button was only actively mentioned eight times (14.5%). Another reason was that, according to 22 children (40.0%), *False Hierarchy (C)* did not contain any closing option: *"In [*False Hierarchy (C)*], one does not have any other option apart from watching an ad! That is really mean!" [P16]*. However, 15 of those children (27.3%) had spotted and marked the difference in the closing options for the first part of this task. Lastly, four children (7.3%) argued that *False Hierarchy (C)* contained fewer Close option possibilities than *Fair (A)*, and two stated that *False Hierarchy (C)* contained *"less text, and then one decides faster" [P56]*.

Even though most children chose *False Hierarchy (C)* as the design more likely to make people watch an ad, three children (5.5%) argued for *Fair (A)*. Their main reason was that *Fair (A)* appeared fairer, and therefore people would be more willing to watch an ad. Additionally, one child stated that with *False Hierarchy (C)*, people would not want to watch an ad, as the interface tries to force this option onto them: *"I chose picture [*Fair (A)*] because, with picture [*False Hierarchy (C)*], one is basically forced to watch the ad, so I would reject it." [P31]*.

4.2.3 Deducing Manipulative Intents. In the last part of this task, we asked the children what the designer's goal behind choosing *False Hierarchy (F)* over *Fair (E)* might be (see Figure 4). We had to exclude ten answers (15.2%) from children who apparently had not understood the task why someone would choose *False Hierarchy (F)* over *Fair (E)*: *"I would choose [*Fair (E)*] because it looks more secure" [P58]*. With this, the following percentages refer to a **total of 56 answers**. 44 children (78.6%) were able to correctly deduce the goal to make people accept more cookies. 20 children (35.7%) justified this by stating that the visibility of the two buttons differed. Also, 20 children (35.7%) specified that for *False Hierarchy (F)*, the Accept button was bigger than the Reject button. Another five comments explicitly mentioned button colors: *"With the chosen picture [*False Hierarchy (F)*], it can be that people click on "Accept all" because the button is larger and the other button is gray, making it easier to be overlooked" [P11]*. Three children (5.4%) also stated that the Reject button looked inactive: *"Maybe one thought that since 'Reject all' is gray, it cannot be clicked [...]" [P15]*.

Some children did not state a goal but still mentioned differences in the designs between *Fair (E)* and *False Hierarchy (F)*. Eight (14.3%) claimed that the button sizes differed, while three stated that people could not click Reject in *False Hierarchy (F)*: *"[...] 'Reject all' is not easy to see. 'Accept all' is big and easy to see" [P32]. "[...] In [False Hierarchy (F)], one cannot [click] Reject" [P41]*.

4.2.4 Summary. Overall, we identified two groups of children within this task: 17 of all 66 children (25.8%) were able to understand the connection between using specific design elements and user decisions, as they had both mentioned the manipulative power of *False Hierarchy (C)* and had correctly identified the malicious intent behind the dark pattern contained in *False Hierarchy (F)*. On the other hand, 15 of all 66 children (22.7%) did not demonstrate this understanding in any of the parts of **Task 2**. The understanding of the remaining participants varied between these two extremes, with a noticeable proportion of children seemingly lacking comprehension of the entire design. For instance, children overlooked the Close button in *False Hierarchy (C)* or did not grasp the semantic meaning of the icon.

4.3 Task 3: Drawing Cookie Banners

For the drawing task, children redesigned cookie banners to increase the likelihood of users accepting cookies.

Of the 66 participating children, 52 (78.8%) children drew an interface. We excluded eight images for which neither the image nor the explanation given was clear to us. With this, the following percentages are based on a total of **44 drawings**. A selection of drawings that children created for this task is depicted in Figure 1 (English translations are available in Appendix A). In total, 38 children (86.4%) applied at least one established dark pattern to their redesigns. The following sections summarize our main codes. Frequencies of codes we assigned to answers during our coding are indicated with parentheses containing the number followed by an "×". Table 1 shows the codebook for this task. Overall, children used

Theme	Code
False Hierarchy (30)	Accept bigger (14) Accept more colorful (14) Alternative hidden (14) Decline greyed out (7) Decline position unexpected (3)
Emotional and Sensory Manipulation (13)	Influential formulation (7) Emojis/Symbols (6) Colors (3)
Forced Action (8)	No alternative (8)
Undesirable Alternative (4)	Compromise alternative (2) Light forced action (2)
Other (5)	*Bait and Switch* (2) Bribery (2) Trustworthiness (1)
Unclear Answer (8)	Unclear (6) Task not understood (2)

Table 1: Codebook for Task 3 where children created drawings containing dark patterns. Names in italics represent Dark patterns from the ontology of Gray et al. [17]. Numbers in parentheses resemble the number of occurrences. A drawing could receive multiple codes if the child used multiple dark patterns.

many different approaches to make users accept cookies, with *False Hierarchy* being the most frequent choice, followed by *Emotional and Sensory Manipulation* and *Undesirable Alternative*. In the following, an asterisk (*) behind a participant number indicates that their drawing is shown in Figure 1.

4.3.1 False Hierarchy. This was the most commonly used pattern, occurring in 30 drawings (68.2%). In particular, children included designs in which the Accept button was bigger (14×) or more colorful (14×), and in which the Reject button was hidden (14×), grayed out (7×), or placed in uncommon positions (3×). Examples for *False Hierarchy* in Figure 1 are P13*, P38*, P57*, P65*, and P66*. The design of P13* is particularly tricky, as it contains two visually identical buttons stating "Accept all" and "Partially use". This creates the impression that it is a fair choice since no obvious manipulation is being used. However, this design still contains a reject button in the shape of a small "x" in the top right corner.

4.3.2 Emotional or Sensory Manipulation. Another approach for designing manipulative cookie banners was to toy with the emotions of the user, which we identified in 13 designs (29.5%). These included using influential formulations (7×) or emojis (6×). For example, P52* used emojis in the buttons to influence users, and P10* claimed that the cookies were used *"for your security and your data privacy"* while adding that they *"[...] lied a bit so that the person thinks that everything is safe, which is not true" [P10]*. Three children also relied on specific colors on the buttons or the background to toy with emotions and, thereby, influence people's decision-making.

4.3.3 Forced Action. Eight children (18.1%) decided to remove the "Decline" button completely, thus forcing users to accept cookies to be able to continue or leave the application entirely: *"Nobody wants to leave the app so they have no other choice but to accept all [cookies]" [P66]*. Examples are the drawings from P07*, P10*, and P66* in Figure 1.

4.3.4 Undesirable Alternative. A fourth approach we saw on 4 drawings (9.1%) was presenting undesirable alternatives to the user. Two children implemented a compromise alternative where users were given the option to not accept cookies, which resulted in other consequences. P65* introduced a waiting time, preventing users from proceeding with using the app if they did not want to accept cookies, and P42* altered the buttons to say "Accept without ads" and "Reject but with ads" to create said compromise (Figure 1). The other two designs replaced the reject button with a button stating "partially accept" (P02 and P26). With this, users would always need to accept at least a certain amount of cookies, creating a *weaker version* of a *Forced Action*.

4.3.5 Other. Less frequent design choices used *Bribery* by, e.g., promising an Amazon voucher when the user accepted the cookies (P07*) or adding unpleasant terms in the fine print (P49 and P58).

4.3.6 Summary. Overall, the most frequently used manipulation was *False Hierarchy* (68.2%), followed by *Emotional and Sensory Manipulation* (29.5%) and *Undesireable Alternatives* (25.0%). While the children transferred dark patterns from the first two tasks to their drawings (e.g., P38* and P57*), they also used manipulations they had not seen in the study before (e.g., P13* and P52*).

4.4 Task 4: Finding Manipulations

In the last task, children searched for manipulative elements in the screen shown in Figure 5. We embedded four different dark patterns into this design (see Section 3.1). We excluded all answers where no justification was given or where the children did not seem to have understood the task and rather described what they would do on the given screen. With this, we excluded 15 answers (22.7%), leaving **51 answers** for our analysis. All following percentages refer to these included 51 answers.

4.4.1 Our Intended Manipulations. The screen contained four dark patterns as specified in Section 3.1: *Emotional and Sensory Manipulation*, *Confirmshaming*, *Trick Question*, and *Bad Defaults*. Figure 5 shows how often children were able to find each of the dark patterns. They spotted both the reversed colors of the two buttons (*Emotional and Sensory Manipulation*) and the *Trick Question* in the cookie banner 24 times (47.1%): *"The buttons have switched colors, which nudges you to click on Abort" [P18]*. Regarding the cookie banner, P29 also stated that *"one should re-formulate it because people won't get it otherwise"*. *Confirmshaming* was found by 13 children (25.5%): *"Nobody is going to confirm to others that he is bad" [P23]*. Only three children noticed the *Bad Default* in the form of a *preselection* (5.9%). Altogether, 64 of all 204 manipulations[9] (31.4%) were spotted correctly in this task.

4.4.2 Other Perceived Manipulations. Apart from the four dark patterns that we embedded into the image, several children reported other aspects that they found manipulative. 13 children (25.5%) stated that the appearance of the whole window itself was

[9] 4 dark patterns × 51 children that provided answers resulted in 204 dark patterns that could have been found.

already manipulative as it tried to keep the user in the game. For example, P51 argued that *"with the first sentence, you think again [whether to quit the app]"*. Eight children (15.7%) reinterpreted the *Confirmshaming* formulation in a way that should motivate a person to keep playing until they received a better score instead of sharing their current achievement: *"It is sort of a small insult that motivates you to keep playing" [P29]*. Four children noticed that the user still had one life left (indicated by the heart over the dialog), which could make them want to continue playing: *"If you still have a heart and some time left, that encourages you to keep playing" [P62]*. Five children stated that the cookie banner formulation "for the best experience on our page" was manipulative and was trying to make people accept cookies: *"For the cookies, it states that it promises the 'best experience'" [P25]*. Finally, *False Hierarchy* was reported by six children (11.8%) regarding the slightly different sizes of the two buttons, and two children mentioned the order of the buttons as well: *"[...] 'Cancel' is larger [...] while 'Ok' is smaller" [P26]*.

4.4.3 Summary. In summary, about half of the children were able to spot the *Trick Question* and the *Emotional and Sensory Manipulation* (switched button colors) as being manipulative. *Confirmshaming* was spotted less frequently, with about every fourth child noticing it, and *Bad Defaults* were only recognized by three children in total. However, some children also argued that other aspects of the interface were manipulative, such as the sheer appearance of the dialog box. Overall, 40 out of 51 children (78.4%) found at least one dark pattern: 23×1 pattern (45.1%), 10×2 patterns (19.6%), 7×3 patterns (13.7%), and 0×4 patterns (0.0%). 11 children (21.6%) did not find any dark pattern in the screenshot.

5 DISCUSSION

Our study provided valuable insights into how children perceive and understand dark patterns. In the following, we highlight and discuss our most important findings and relate them to each of our four research questions.

5.1 Regarding RQ1: Children Perceive Manipulative Designs More Negatively

Our results for **Task 1** revealed that children perceived designs that contained dark patterns more negatively compared to a fair design. This suggests a similarity to adults in the perception of dark patterns as, for example, Conti and Sobiesk [8] observed that users would develop overall negative attitudes towards websites when they realized they entailed manipulative intents. This notion is also showing in our study as the children rated the dark pattern-free design *Fair (A)* and the mild *Confirmshaming (B)* as the most beautiful, while the aggressive *False Hierarchy (C)* was rated the least beautiful. However, this result diverges from existing literature on design aesthetics, which typically favors clean designs, which is given in *False Hierarchy (C)* [6]. This discrepancy raises the question of whether the children, maybe driven by their curiosity and their nascent logical and inferential thinking regarding beauty [49], not only evaluated the designs objectively but already recognized elements of manipulative designs. Interestingly, *Confirmshaming (B)* did not seem to negatively affect children's trust perceptions, which differs from prior findings about adults [34]. Potentially, the manipulation of *Confirmshaming* might have been too subtle for children to notice but instead increased the overall perceived complexity of the screen. This would be supported by the explanations from Pechmann et al. that the areas in the human brain that are responsible for risk assessment are only fully developed at the end of childhood [37], potentially making children especially vulnerable to latent dark patterns. Still, it supports the findings of Luguri and Strahilevitz [27] that mild dark patterns like *Confirmshaming* generally receive less backlash from users than aggressive dark patterns such as *False Hierarchy*. Likewise, children judged *False Hierarchy (C)* as the least trustworthy, which is in accordance with prior findings about adults tending to have less trust when encountering dark patterns [50].

5.2 Regarding RQ2: Some Children Understand Manipulative Intents

In **Task 2**, children had to first spot differences between *Fair (A)* and *False Hierarchy (C)* and then think about which design would make more people watch an ad. Here, most children (77.3%) found all differences. Interestingly, ten children stated that the user had only one choice in *False Hierarchy (C)*. Since three of those children had spotted the Close button, this could mean that some children were not able to associate the × icon with the function of closing a window.

In the second part of this task, children were given *Fair (E)* and *False Hierarchy (F)* and had to deduce why a designer chose the latter design in their app. 44 of 56 children (78.6%) correctly stated that the reason was to make more people accept cookies. The two most frequent justifications were the different visibility of the buttons and the different button sizes. With this, most children realized that this difference might cause people to accept more cookies. However, it is unclear whether those children thoroughly understood the reasons behind this or simply concluded the intended manipulation solely based on the task focusing on the existence of visual differences, in turn, potentially priming the children. In fact, Di Geronimo et al. [10] found that adult users could only scarcely detect dark patterns but performed notably better when informed about the topic of manipulations. Accordingly, the potential priming induced by the task design might have contributed to the high number of children detecting the manipulations behind *False Hierarchy*. Nevertheless, three children explicitly mentioned that the reject button in *False Hierarchy (F)* looked inactive, indicating an understanding of the applied manipulation.

5.3 Regarding RQ3: Children Can Use Dark Patterns Themselves

In their drawings for **Task 3**, the most frequently applied pattern was *False Hierarchy* (e.g., P38 and P57 in Figure 1). This might stem from children's frequent prior exposure to this pattern due to its strong prevalence on popular websites [29] and in apps [10]. Another explanation could be that children were primed for *False Hierarchy* since this specific dark pattern had occurred in the previous tasks of our study. Because of this, we expected many children to use *False Hierarchy*. However, since children were able to use this pattern correctly in their drawings, they may also have a basic understanding of how it manipulates users. Additionally, children also used manipulations that had not been part of previous tasks,

such as *Emotional and Sensory Manipulation* (P52) or *Bribery* (P07) (see Figure 1). These approaches to nudge users also contain similarities to the results obtained by Sánchez Chamorro et al. [42], who worked with adult UX designers to discuss approaches to influence users to provide their email addresses. For example, the designers stated that providing incentives or rationales to users as to why they should choose a certain decision can influence users in their decision-making. Children in our study tried to convince users to accept cookies by stating that the cookies were used for the user's security (P10) or for a good time (P52). Others, like P07, even offered shopping discounts as an incentive. Additionally, the designers from the study of Sánchez Chamorro et al. [42] stated that they themselves try to avoid taking away autonomy from users as it is unethical. In our study, several children (e.g., P10) undermined decision autonomy by removing the option to reject cookies or by adapting the text to point out that these cookies were used for the user's security and privacy. Adding value for the user is another way of convincing them to make a certain choice. In our study, P42 made it clear to the user that they could use the application without ads if they accepted the cookies. While these nudges might not always be instances of dark patterns, these techniques can still be misused to undermine the user's decision autonomy by, e.g., providing irresistible incentives [42].

Several children combined multiple dark patterns, which is also common in real user interfaces [10, 29], and some designs were particularly perfidious, such as the drawing by P13 (see Figure 1): The child created a seemingly fair choice for the user with the two equally designed buttons "Accept all" and "Partially accept". While cookie banner designs often entail *False Hierarchy* [45], its absence in this drawing suggests trustworthiness and makes people less suspicious that the banner contains any other options. As a result, the user might deem the "Partially accept" button sufficient and overlook the small Close button in the top right corner.

Overall, we are uncertain whether our results should be attributed to the children's previous exposure and their reproduction of known manipulative designs for their drawings or whether they stem from the children's own ideas on how to manipulate users. Either way, letting children draw, as proposed by Doyle et al. [11], showed us that children are capable of using manipulative designs themselves. More research is needed to understand this phenomenon fully.

5.4 Regarding RQ4: Spotting Manipulative Elements Is Challenging

Our results strongly indicate that some children will detect "fishy" designs even when they cannot discern the manipulative elements. One example is the complex *Trick Question* in our cookie banner (see Figure 5). Here, P23 stated that the text *"sounds kind of weird"*. P29 even suggested that this text should be re-formulated since *"some people will not understand it"*. Another example is the *False Hierarchy* in the cookie banner in Figure 4 (F) that nudges people towards accepting all cookies. P23 explained that *"the 'Reject all' button looks different, and this can lead to a different result that I do not fully know"*. This observation aligns with the finding of Zhao et al. [53] that most children have a basic level of understanding of online privacy and security.

Bad Defaults in the form of a preselection, however, were barely identified by any child in the last task. While we also regard this dark pattern as the hardest to spot among the patterns we chose, it is still surprising that only 3 of 51 children (5.9%) mentioned it at all. This raises the question whether children generally question given default settings and whether they can understand the data-sharing implications of this element. Either way, children might be especially vulnerable to dark patterns like *Bad Defaults* that use *preselections* to make users share more data by default than necessary. As *preselections* are commonly used in mobile apps [10] and are also harder to detect than other dark patterns by adults [3], increasing children's awareness of such nuanced manipulations becomes especially important.

The second least frequently found dark pattern was *Confirmshaming*, with only 13 of 51 children (25.5%) spotting it. This appears to diverge from results about adults. For example, in the study by Bongard-Blanchy et al. [3], *Confirmshaming* was one of the patterns that the majority of participants detected correctly. This strengthens our hunch regarding RQ1 that the manipulation of *Confirmshaming* might be too subtle for some children to notice, so that it does not feel manipulative to them. While this could be traced back to the particular example we used in this task, another explanation for children's poor ability to detect *Confirmshaming* patterns might be their developing cognitive abilities [37]. However, more empirical data is needed on this topic.

5.5 Differences Between Children

We know from the literature that there are significant individual differences between adults regarding their susceptibility to dark patterns [29]. Our results show that children also exhibit quite different levels of susceptibility. During our study, we observed two groups of children:

- children who understood how certain design elements can influence user decisions (17/66), and
- children who could not make a connection between design elements and user decisions (15/66).

With this, our results also suggest the existence of a group of children who would have likely fallen prey to the dark patterns in our study in a comparable real-world scenario: One-third of the children apparently misunderstood *False Hierarchy (C)*, as they claimed the design would not provide an option to reject watching an ad. We presume that this group of children requires special consideration in the fight against dark patterns, as they first need to develop a fundamental understanding. For instance, simple tools that detect and highlight dark patterns would not suffice if a child did not understand what they were being warned about. Another step to reduce the effect of dark patterns on children would be to also educate parents, as already suggested in the context of online privacy [25] and mobile games [46]. We consider this desirable for dark patterns as well, to enable children to practice handling dark patterns correctly under parental guidance.

6 LIMITATIONS

We conducted our study with three different classes from the same local school. While the social background of these children was reasonably diverse, our results are not necessarily generalizable

to children from other countries or other educational tracks. We also only covered a small age range (mainly 10- to 11-year olds), so our results do not simply generalize to other age groups. Since *Emotional and Sensory Manipulation* also uses colors for its manipulation, color-blind children might not spot this pattern for Task 4. The screenshots in our study often used *False Hierarchy*, which likely primed our participants for this specific dark pattern so that most chose a variant of it in their drawing for **Task 3**. We were aware of this potential bias but chose this order of tasks to guarantee unbiased first impressions of our fictional screenshots in **Tasks 1 and 2**. Letting children draw manipulative designs up front would have primed them for both tasks. In the drawings of **Task 3**, we also observed other manipulative techniques not previously contained in the designs for **Tasks 1 and 2**, showing that children did not only re-use patterns they had seen during the study. Finally, our study only covered a relatively small set of dark patterns, and future work could build on our findings by investigating a broader range of dark patterns with children.

7 CONCLUSION AND FUTURE WORK

In this work, we investigated how children perceive and understand dark patterns and their ability to deduce the deceptive intents behind them. For this, we conducted a study at a school in Germany with 66 fifth-graders aged 10 to 11. We used several established elicitation techniques with consideration of our under-age sample, such as *free-recall* and *drawings*, using four tasks. We found that children rated designs containing dark patterns more negatively than those without (**Task 1**), which is in line with findings for adults [8, 27]. As with adults [27], this trend was amplified when aggressive dark patterns were used. When choosing a design that would likely influence users as demanded (**Task 2**), 92.7% picked the version containing the dark pattern *False Hierarchy* over a fair design. Drawing manipulative elements (**Task 3**) showed that several children were capable of applying manipulative design strategies themselves. However, spotting manipulative elements in a screenshot (**Task 4**) was challenging. Overall, children only found around 32% of all manipulations. Here, about half of the children were able to identify *Trick Question* and *Emotional and Sensory Manipulation*. *Confirmshaming* was only spotted by every fourth child, and only 3 of 51 children noticed malicious *Bad Defaults* in the form of a preselection. This indicates a particular vulnerability to subtle privacy-related dark patterns that should be investigated further.

Currently, most research on dark patterns is focused on adults. This leaves numerous questions unexplored about how children understand and interact with such malicious designs and how they could be protected. Even though we analyzed answers from 66 children, more research in this field is needed, as youth internet access and thus exposure to these manipulations is growing rapidly. Future work could focus on a wider span of dark pattern types and other underage groups. Furthermore, new approaches to countermeasures that particularly support children need to be studied and evaluated. Additionally, educators and lawmakers should consider children's early exposure to such malicious designs, and implement regulations to educate children as early as feasible to minimize their vulnerability.

With our work, we hope to motivate researchers to investigate other user groups in the context of dark patterns and help protect upcoming generations from the influence of such manipulative UI designs.

ACKNOWLEDGMENTS

This work was funded in part by the German B-IT Foundation. We also want to thank the teachers and the children involved in our study. Without you, this work would not have been possible.

REFERENCES

[1] Fadi A. Aloul. 2012. The Need for Effective Information Security Awareness. *Journal of Advances in Information Technology* 3, 3 (2012), 176–183. https://doi.org/10.4304/jait.3.3.176-183

[2] J. Craig Andrews, Kristen L. Walker, and Jeremy Kees. 2020. Children and Online Privacy Protection: Empowerment from Cognitive Defense Strategies. *Journal of Public Policy & Marketing* 39, 2 (2020), 205–219. https://doi.org/10.1177/0743915619883638

[3] Kerstin Bongard-Blanchy, Arianna Rossi, Salvador Rivas, Sophie Doublet, Vincent Koenig, and Gabriele Lenzini. 2021. "I Am Definitely Manipulated, Even When I Am Aware of It. It's Ridiculous!" - Dark Patterns from the End-User Perspective. In *Proceedings of the 2021 ACM Designing Interactive Systems Conference* (Virtual Event, USA) *(DIS '21)*. Association for Computing Machinery, New York, NY, USA, 763–776. https://doi.org/10.1145/3461778.3462086

[4] Virginia Braun and Victoria Clarke. 2006. Using Thematic Analysis in Psychology. *Qualitative Research in Psychology* 3, 2 (2006), 77–101. https://doi.org/10.1191/1478088706qp063oa

[5] Rubens Cantuni. 2020. Concept. In *Designing Digital Products for Kids: Deliver User Experiences That Delight Kids, Parents, and Teachers*. Apress, Berkeley, CA, 41–74. https://doi.org/10.1007/978-1-4842-6287-0_4

[6] Rubens Cantuni. 2020. UI Design. In *Designing Digital Products for Kids: Deliver User Experiences That Delight Kids, Parents, and Teachers*. Apress, Berkeley, CA, 147–204. https://doi.org/10.1007/978-1-4842-6287-0_8

[7] Seung Youn Chyung, Ieva Swanson, Katherine Roberts, and Andrea Hankinson. 2018. Evidence-Based Survey Design: The Use of Continuous Rating Scales in Surveys. *Performance Improvement* 57, 5 (2018), 38–48. https://doi.org/10.1002/pfi.21763

[8] Gregory Conti and Edward Sobiesk. 2010. *Malicious Interface Design: Exploiting the User*. Association for Computing Machinery, New York, NY, USA. 271–280 pages. https://doi.org/10.1145/1772690.1772719

[9] Pearl Denham. 1993. Nine- to Fourteen-Year-Old Children's Conception of Computers Using Drawings. *Behaviour & Information Technology* 12, 6 (1993), 346–358. https://doi.org/10.1080/01449299308924399

[10] Linda Di Geronimo, Larissa Braz, Enrico Fregnan, Fabio Palomba, and Alberto Bacchelli. 2020. UI Dark Patterns and Where to Find Them: A Study on Mobile Applications and User Perception. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3313831.3376600

[11] Emma E.H. Doyle, Sara E. Harrison, Stephen R. Hill, Matt Williams, Douglas Paton, and Ann Bostrom. 2022. Eliciting Mental Models of Science and Risk for Disaster Communication: A Scoping Review Of Methodologies. *International Journal of Disaster Risk Reduction* 77 (2022). https://doi.org/10.1016/j.ijdrr.2022.103084

[12] Dan Fitton, Beth T Bell, and Janet C Read. 2021. Integrating Dark Patterns into the 4Cs of Online Risk in the Context of Young People and Mobile Gaming Apps. In *IFIP Conference on Human-Computer Interaction*. Springer International Publishing, Cham, 701–711. https://doi.org/10.1007/978-3-030-85610-6_40

[13] Dan Fitton and Janet C. Read. 2019. Creating a Framework to Support the Critical Consideration of Dark Design Aspects in Free-to-Play Apps. In *Proceedings of the 18th ACM International Conference on Interaction Design and Children* (Boise, ID, USA) *(IDC '19)*. Association for Computing Machinery, New York, NY, USA, 407–418. https://doi.org/10.1145/3311927.3323136

[14] Paul Graßl, Hanna Schraffenberger, Frederik Zuiderveen Borgesius, and Moniek Buijzen. 2021. Dark and Bright Patterns in Cookie Consent Requests. https://doi.org/10.31234/osf.io/gqs5h

[15] Colin M. Gray, Jingle Chen, Shruthi Sai Chivukula, and Liyang Qu. 2021. End User Accounts of Dark Patterns as Felt Manipulation. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW2, Article 372 (Oct 2021), 25 pages. https://doi.org/10.1145/3479516

[16] Colin M. Gray, Yubo Kou, Bryan Battles, Joseph Hoggatt, and Austin L. Toombs. 2018. The Dark (Patterns) Side of UX Design. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3173574.3174108

[17] Colin M. Gray, Cristiana Teixeira Santos, Nataliia Bielova, and Thomas Mildner. 2024. An Ontology of Dark Patterns Knowledge: Foundations, Definitions, and a Pathway for Shared Knowledge-Building. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 289, 22 pages. https://doi.org/10.1145/3613904.3642436

[18] Robin S. Grenier and Dana Dudzinska-Przesmitzki. 2015. A Conceptual Model for Eliciting Mental Models Using a Composite Methodology. *Human Resource Development Review* 14, 2 (2015), 163–184. https://doi.org/10.1177/1534484315575966

[19] Jens Grossklags and Alessandro Acquisti. 2007. When 25 Cents is Too Much: An Experiment on Willingness-To-Sell and Willingness-To-Protect Personal Information. https://www.cs.cit.tum.de/fileadmin/w00cfj/ct/papers/2007-WEIS-Grossklags-Acquisti.pdf

[20] Johanna Gunawan, Amogh Pradeep, David Choffnes, Woodrow Hartzog, and Christo Wilson. 2021. A Comparative Study of Dark Patterns Across Web and Mobile Modalities. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW2, Article 377 (Oct 2021), 29 pages. https://doi.org/10.1145/3479521

[21] Ralph Hertwig and Ido Erev. 2009. The description–experience gap in risky choice. *Trends in Cognitive Sciences* 13 (2009), 517–523. https://api.semanticscholar.org/CorpusID:2190783

[22] Maxwell Keleher, Fiona Westin, Preethi Nagabandi, and Sonia Chiasson. 2022. How Well Do Experts Understand End-Users' Perceptions of Manipulative Patterns?. In *Nordic Human-Computer Interaction Conference* (Aarhus, Denmark) *(NordiCHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 52, 21 pages. https://doi.org/10.1145/3546155.3546656

[23] Bran Knowles, Joe Finney, Sophie Beck, and James Devine. 2018. What Children's Imagined Uses of the BBC micro:bit Tells Us About Designing for their IoT Privacy, Security and Safety. In *Living in the Internet of Things: Cybersecurity of the IoT - 2018*. Institute of Engineering and Technology, London, UK, 6 pages. https://doi.org/10.1049/cp.2018.0015

[24] Christie Kodama, Beth St. Jean, Mega Subramaniam, and Natalie Greene Taylor. 2017. There's a Creepy Guy on the Other End at Google!: Engaging Middle School Students in a Drawing Activity To Elicit Their Mental Models of Google. *Information Retrieval Journal* 20, 5 (2017), 403–432. https://doi.org/10.1007/s10791-017-9306-x

[25] Priya Kumar, Shalmali Milind Naik, Utkarsha Ramesh Devkar, Marshini Chetty, Tamara L. Clegg, and Jessica Vitak. 2017. No Telling Passcodes Out Because They're Private: Understanding Children's Mental Models of Privacy and Security Online. *Proc. ACM Hum.-Comput. Interact.* 1, CSCW, Article 64 (Dec 2017), 21 pages. https://doi.org/10.1145/3134699

[26] Gitte Lindgaard, Cathy Dudek, Devjani Sen, Livia Sumegi, and Patrick Noonan. 2011. An Exploration of Relations Between Visual Appeal, Trustworthiness and Perceived Usability of Homepages. *ACM Trans. Comput.-Hum. Interact.* 18, 1, Article 1 (May 2011), 30 pages. https://doi.org/10.1145/1959022.1959023

[27] Jamie Luguri and Lior Jacob Strahilevitz. 2021. Shining a Light on Dark Patterns. *Journal of Legal Analysis* 13, 1 (Mar 2021), 43–109. https://doi.org/10.1093/jla/laaa006

[28] Kai Lukoff, Alexis Hiniker, Colin M. Gray, Arunesh Mathur, and Shruthi Sai Chivukula. 2021. What Can CHI Do About Dark Patterns?. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI EA '21)*. Association for Computing Machinery, New York, NY, USA, Article 102, 6 pages. https://doi.org/10.1145/3411763.3441360

[29] Francisco Lupiáñez-Villanueva, Alba Boluda, Francesco Bogliacino, Giovanni Liva, Lucie Lechardoy, and Teresa Rodríguez de las Heras Ballell. 2022. Behavioural Study on Unfair Commercial Practices in the Digital Environment: Dark Patterns and Manipulative Personalisation. https://doi.org/10.2838/859030

[30] Ana Maria Marhan, Mihai Ioan Micle, Camelia Popa, and Georgeta Preda. 2012. A Review of Mental Models Research in Child-Computer Interaction. *Procedia-Social and Behavioral Sciences* 33 (2012), 368–372. https://doi.org/10.1016/j.sbspro.2012.01.145

[31] Arunesh Mathur, Gunes Acar, Michael J. Friedman, Eli Lucherini, Jonathan Mayer, Marshini Chetty, and Arvind Narayanan. 2019. Dark Patterns at Scale: Findings from a Crawl of 11K Shopping Websites. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 81 (Nov 2019), 32 pages. https://doi.org/10.1145/3359183

[32] Philipp Mayring. 2014. Qualitative Content Analysis: Theoretical Foundation, Basic Procedures and Software Solution. https://nbn-resolving.org/urn:nbn:de:0168-ssoar-395173

[33] Pekka Mertala. 2021. *Using Playful Methods To Understand Children's Digital Literacies*. Sage Publications Ltd, London, UK, 179–191.

[34] Tasneem Naheyan and Kiemute Oyibo. 2024. The Effect of Dark Patterns and User Knowledge on User Experience and Decision-Making. In *Persuasive Technology*, Nilufar Baghaei, Raian Ali, Khin Win, and Kiemute Oyibo (Eds.). Springer Nature Switzerland, Cham, 190–206.

[35] Maggie Oates, Yama Ahmadullah, Abigail Marsh, Chelse Swoopes, Shikun Zhang, Rebecca Balebako, and Lorrie Faith Cranor. 2018. Turtles, Locks, and Bathrooms: Understanding Mental Models of Privacy Through Illustration. *Proceedings on Privacy Enhancing Technologies* 2018, 4 (2018), 5–32. https://doi.org/10.1515/popets-2018-0029

[36] Georgia Panagiotaki, Gavin Nobes, and Robin Banerjee. 2006. Children's Representations of the Earth: A Methodological Comparison. *British Journal of Developmental Psychology* 24, 2 (2006), 353–372. https://doi.org/10.1348/026151005X39116

[37] Cornelia (Connie) Pechmann, Linda J. Levine, Sandra E. Loughlin, and Frances M. Leslie. 2005. Impulsive and Self-Conscious: Adolescents' Vulnerability to Advertising and Promotion. *Journal of Public Policy & Marketing* 24 (2005), 202 – 221. https://api.semanticscholar.org/CorpusID:9950086

[38] P. J. Pridmore and R. G. Lansdown. 1997. Exploring Children's Perceptions of Health: Does Drawing Really Break Down Barriers? *Health Education Journal* 56, 3 (1997), 219–230. https://doi.org/10.1177/001789699705600302

[39] Samantha Punch. 2002. Research With Children: The Same or Different From Research With Adults? *Childhood* 9, 3 (2002), 321–341. https://doi.org/10.1177/0907568202009003005

[40] Rohani Rohan, Suree Funilkul, Wichian Chutimaskul, Prasert Kanthamanon, Borworn Papasratorn, and Debajyoti Pal. 2023. Information Security Awareness in Higher Education Institutes: A Work in Progress. https://doi.org/10.1109/KST57286.2023.10086884

[41] Bernhard Rohleder. 2022. Bitcom Kinder- und Jugendstudie 2022. https://www.bitkom.org/sites/main/files/2022-06/Bitkom-Charts_Kinder_Jugendliche_09.06.2022_0.pdf

[42] Lorena Sánchez Chamorro, Kerstin Bongard-Blanchy, and Vincent Koenig. 2023. Ethical Tensions in UX Design Practice: Exploring the Fine Line Between Persuasion and Manipulation in Online Interfaces. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference* (Pittsburgh, PA, USA) *(DIS '23)*. Association for Computing Machinery, New York, NY, USA, 2408–2422. https://doi.org/10.1145/3563657.3596013

[43] René Schäfer, Paul Miles Preuschoff, and Jan Borchers. 2023. Investigating Visual Countermeasures Against Dark Patterns in User Interfaces. In *Mensch Und Computer 2023* (Rapperswil, Switzerland) *(MuC '23)*. Association for Computing Machinery, New York, NY, USA, 161–172. https://doi.org/10.1145/3603555.3603563

[44] David Smahel, Hana Machackova, Giovanna Mascheroni, Lenka Dedkova, Elisabeth Staksrud, Kjartan Ólafsson, Sonia Livingstone, and Uwe Hasebrink. 2020. *EU Kids Online 2020: Survey Results From 19 Countries*. Technical Report. EU Kids Online. https://orfee.hepl.ch/handle/20.500.12162/5299

[45] Than Htut Soe, Oda Elise Nordberg, Frode Guribye, and Marija Slavkovik. 2020. Circumvention by Design - Dark Patterns in Cookie Consent for Online News Outlets. In *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society* (Tallinn, Estonia) *(NordiCHI '20)*. Association for Computing Machinery, New York, NY, USA, Article 19, 12 pages. https://doi.org/10.1145/3419249.3420132

[46] Carla Sousa and Ana Oliveira. 2023. The Dark Side of Fun: Understanding Dark Patterns and Literacy Needs in Early Childhood Mobile Gaming. *European Conference on Games Based Learning* 17, 1 (2023), 599–610. https://doi.org/10.34190/ecgbl.17.1.1656

[47] Andrew Thatcher and Mike Greyling. 1998. Mental Models of the Internet. *International Journal of Industrial Ergonomics* 22, 4 (1998), 299–305. https://doi.org/10.1016/S0169-8141(97)00081-4

[48] Patti M. Valkenburg and Jessica T Piotrowski. 2017. Plugged In: How Media Attract and Affect Youth. https://doi.org/10.12987/yale/9780300218879.001.0001

[49] Susanne Vogl. 2021. *Mit Kindern Interviews führen: Ein Praxisorientierter Überblick* (2nd revised ed.). Verlag Julius Klinkhardt, 142–157. https://doi.org/10.25656/01:22252

[50] Christian Voigt, Stephan Schlögl, and Aleksander Groth. 2021. Dark Patterns in Online Shopping: of Sneaky Tricks, Perceived Annoyance and Respective Brand Trust. In *HCI in Business, Government and Organizations*, Fiona Fui-Hoon Nah and Keng Siau (Eds.). Springer International Publishing, Cham, 143–155. https://doi.org/10.1007/978-3-030-77750-0_10

[51] Claire M. White, Michaela Gummerum, and Yaniv Hanoch. 2015. Adolescents' and Young Adults' Online Risk Taking: The Role of Gist and Verbatim Representations. *Risk Analysis* 35 (2015), 1407–1422. Issue 8. https://doi.org/10.1111/risa.12369

[52] José P Zagal, Staffan Björk, and Chris Lewis. 2013. Dark Patterns in the Design of Games. In *Foundations of Digital Games 2013*.

[53] Jun Zhao, Ge Wang, Carys Dally, Petr Slovak, Julian Edbrooke-Childs, Max Van Kleek, and Nigel Shadbolt. 2019. 'I Make up a Silly Name': Understanding Children's Perception of Privacy Risks Online. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3290605.3300336

A DRAWING TRANSLATIONS

Figure 7 contains translated versions of the screens we used in our study and translations of the children's drawings from **Task 3**.

Figure 7: Translated drawings from Figure 1. To visualize all translations, we blurred hand-written text and added the respective English translation on top such that the overall appearance of each image did not change.

B ORIGINAL SCREENSHOTS

Figures 8–10 contain the original versions of the screens that we gave the children in our study.

Figure 8: Original version from Figure 3 that the children received during the study.

Figure 9: Original version from Figure 4 that the children received during the study.

Figure 10: Original version from Figure 5 that the children received during the study.