'Surprised, Shocked, Worried': User Reactions to Facebook Data Collection from Third Parties

Patricia Arias-Cabarcos
Paderborn University
Germany
pac@mail.upb.de

Saina Khalili*
KIT
Germany
saina.khalili72@gmail.com

Thorsten Strufe
KASTEL Security Research Labs, KIT
Germany
strufe@kit.edu

ABSTRACT

Data collection and aggregation by online services happens to an extent that is often beyond awareness and comprehension of its users. Transparency tools become crucial to inform people, though it is unclear how well they work. To investigate this matter, we conducted a user study focusing on Facebook, which has recently released the "*Off-Facebook Activity*" transparency dashboard that informs about personal data collection from third parties. We exposed a group of N = 100 participants to the dashboard and surveyed their level of awareness and reactions to understand how transparency impacts users' privacy attitudes and intended behavior. Our participants were surprised about the massive amount of collected data, became significantly less comfortable with data collection, and more likely to take protective measures. Collaterally, we observed that current consent schemes are inadequate. Based on the survey findings, we make recommendations for more usable transparency and highlight the need to raise awareness about transparency tools and to provide easily actionable privacy controls.

KEYWORDS

usable privacy, social networks, transparency, user study

1 INTRODUCTION

Online social networks collect and process a great deal of personal information, mainly to target users with personalized content and advertisements, making profit from their interactions. The abuse of these data can have a direct impact on user autonomy and agency [2, 68], and even disrupt society in unprecedented ways, as revealed by scandals like the Cambridge Analytica interference with the US elections in 2016 [11]. Facebook (FB) in particular, the most used social media platform, has created a vast tracking infrastructure that silently collects user data as they navigate the web and use applications, through ubiquitous[1] social plugins (e.g., "Like" buttons), hidden pixels, and other business tools offered to third party companies[2], which extend to collecting information

about offline activities, like sales in brick and mortar shops[3]. Recent research [9] uncovered that, based on these data, Facebook assigned ad preference labels that correspond to 'special categories' of sensitive data to 73% of EU FB users, even after the GDPR [20] was enacted. Sensitive categories were also attributed to users in countries where disclosure of certain information (e.g., sexual orientation) can be life threatening. This form of aggressive attribution for targeting purposes poses important threats to privacy, and it fosters discrimination based on race, gender, or age, among other sensitive attributes [3, 4, 63].

Amidst this surveillance reality, users do little to protect themselves [61, 72]. The privacy literature points out that social network users seem to care more about *social privacy* (the concern about controlling access to personal data by other people), than about *institutional privacy* (the concern about how providers and third parties will use personal data) [72]. However, user inaction against institutional surveillance likely stems from a poor understanding about how tracking works [3, 16, 51, 61]. Transparency is therefore crucial to inform and empower users, but this type of feature is rarely available within social networks or just provided in a limited manner. It was not until January 2020 that FB released the "*Off-Facebook Activity*" transparency tool, a dashboard for users to see which third party apps and websites have shared their information with Facebook to expand the aggregated digital dossiers that represent the user profiles [23]. Given the knowledge gap on how FB transparency can affect users behavior and explain their (lack of) privacy-protective strategies, we decided to explore the following research questions:

RQ1 [Awareness and Impact on Privacy Attitudes] How does the "*Off-Facebook Activity*" tool affect users' privacy attitudes?

RQ2 [User Reactions and Intended Behavior] What are users' reactions and intended behavioral changes after exposure to the "*Off-Facebook Activity*" tool?

RQ3 [Usability and Usefulness] How usable is the "*Off-Facebook Activity*" dashboard as a transparency enhancing tool as evaluated by users?

For the first two questions, we hypothesize that more knowledge, gained through the transparency tool, increases privacy concerns and leads to privacy-protecting behavior, based on the "bounded rationality" explanation for the privacy-paradox [7]. The third question is motivated by the apparent complexity of the FB transparency tool –plagued with dark patterns and complex menus– since low usability is a recognized impediment to efficiently use security and privacy tools [24, 60]. Therefore, we seek to understand how to

---

*Contributed to this research while doing her Master Thesis at KIT in 2019 (Currently unaffiliated).

[1]In April 2018 around 8.4 million websites were using the Like button, another million the Share button, and 2.2 million websites had Facebook Pixels [44]

[2]https://www.facebook.com/help/331509497253087

[3]https://www.facebook.com/business/learn/facebook-offline-conversions

improve usability for more efficient transparency. To address our research questions, we conducted an online survey with N = 100 participants. Users were first asked about their privacy attitudes and protective strategies, then directed to the "*Off-Facebook Activity*" dashboard to visualize and navigate the data collected by FB about them. After exposure to the tool, we asked them again about privacy attitudes, including questions on their reactions, intended behavior, and how usable the tool was to achieve their privacy goals.

Our findings show that users want transparency tools and that they would like these tools to be not only easy to use, but also easy to access and more actively advertised to the public. Despite most users where active in FB, 85% of the participants did not know about the "*Off-Facebook Activity*" tool more than a year after its release. Once participants are exposed to transparency information, their self-reported likelihood to take protective measures against FB tracking increases. They met the amount of collected data with shock and surprise, which makes it questionable if the consent given to share the data can be considered informed, and hence effective. On the positive side, our participants experienced positive feelings related to the possibility of exercising control, though there are limitations that hinder user understanding and ability to take action. Based on these findings, we contribute to the literature with recommendations on how to improve the design of transparency tools.

2　BACKGROUND AND RELATED WORK

2.1　Off-Facebook Activity

Facebook rolled out the *"Off-Facebook Activity"* tool in August 2019, initially in a few countries, until it became globally available in January 2020[4]. The goal of the tool is to provide transparency and control regarding the data that 3rd parties send to Facebook about their users. The tool's landing page (see Fig. 1) gives access to a summary of the apps and websites that share data with Facebook, where the user can see the number of interactions involved, and a generic list of the type of data that might have been exchanged. To obtain details about the specific information sent, the tool links to a data export page where users can download JSON or HTML files to navigate offline. As mechanisms for control, *"Off Facebook Activity"* lets users clear their previous activity and disconnect their future activity, options that eliminate the link between the data and the user account, only for past exchanges or for any past and upcoming exchange, respectively. Our research on this tool extends the body of knowledge on privacy in online social networks [36, 54, 73], building on and complementing prior work on tracking transparency.

2.2　Privacy in Social Networks

With the development and continued increase of social media usage, privacy concerns have been on the rise and a highly debated and investigated topic. Many researchers posit that there is a "privacy paradox" [6] in that users expose their personal information in these platforms even if they claim to worry about privacy.

Prior work in this area explored different factors shaping the apparent paradox. A survey study on the privacy strategies of Facebook users with 384 participants [24] showed that behavior is biased by the lack of information, as people do not know enough to grasp privacy risks and so they have to make "bounded rational decisions". Another key factor observed was that the usability of privacy controls is a deterrent for FB users to apply privacy measures.

A lion share of this literature has focused on the social dimension of privacy (control of personal information during social interactions) instead of on the institutional side of it (control of information usage by institutions/3rd parties). Indeed, despite the large body of research and development projects on protecting users from the providers [14, 28, 29, 33], privacy-enhanced alternatives have not managed to gain much traction so far. In this regard, Young et al. [72] observed that social site users seem to be less concerned with institutional privacy because they report little protective actions against surveillance. But this work also suggests that users might not be aware of the applied surveillance practices, given the limited institutional transparency in place. A later study on user privacy strategies [12] confirms that social network users have but vague knowledge of institutional surveillance practices. We aim at extending this body of research by investigating users' awareness about institutional transparency, and the effect of transparency on privacy attitudes and behavior. We selected Facebook because of its popularity as the most used social network (~3 billion users[5]) and the consequential strong impact that transparency or the lack of it can have. The "*Off-Facebook Activity*" tool is unique in that it shows data sent to the social network by 3rd parties, while other transparency interfaces only agglutinate user activity within a single application or domain of applications (e.g., Google My Activity shows actions in Youtube, Gmail, etc.). We take a first step towards understanding reactions to this less known data sharing practices by online social network providers.

2.3　Transparency Tools

Transparency is an important privacy principle at the heart of informational self-determination. Users have the right to know how their data are collected and used, which is especially relevant given that the business model ingrained in the Web is based on user profiling. Transparency Enhancing Tools (TETs) provide technological means to this end [49]. With regard to web tracking, the most common transparency tools provide information about: 1) which trackers are present in a website, (e.g., ad blockers like Ghostery [25]), and 2) why specific ads are shown to a user (e.g., the AdChoices icon [37]). Furthermore, there is extensive work on privacy or transparency dashboards to allow online service users exploring and managing data collected about them [34, 49, 57, 62, 65]. Research so far has shown that these tools are limited, often incomplete or imprecise, and sometimes misleading [3, 70]. Weinshel et al. [70] conducted a longitudinal field study with 425 users concluding that web tracking transparency tools with more detailed information (potential data inferences instead of just tracker presence) lead to increased intention to take privacy-protecting actions. We build on this study

---

[4]https://about.fb.com/news/2020/01/data-privacy-day-2020/

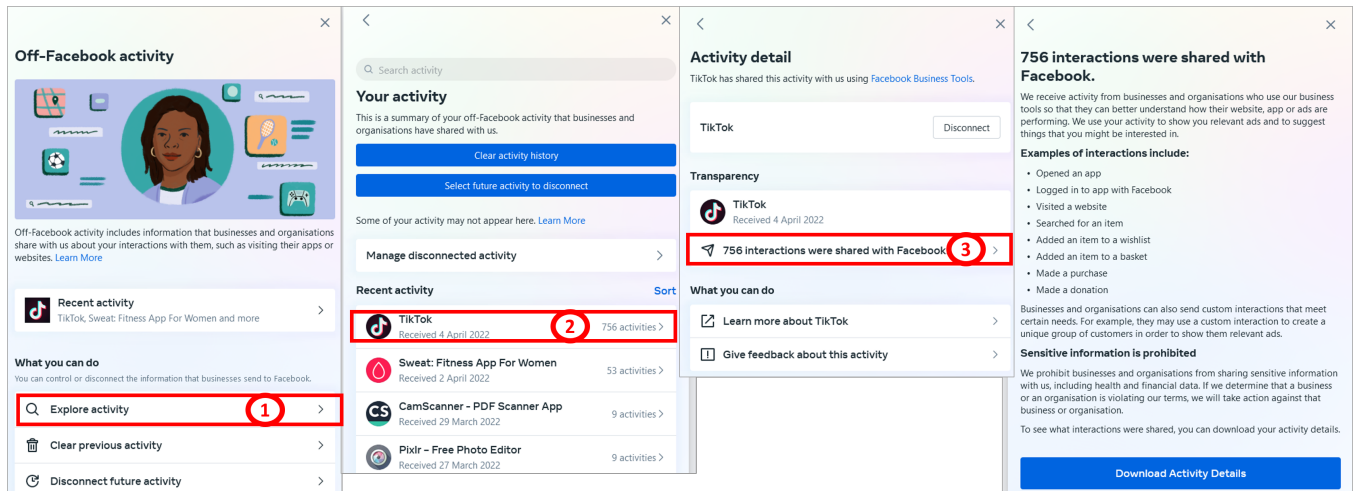[5]https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/

Figure 1: Interface of the "*Off-Facebook Activity*" transparency tools. It shows screenshots of the sequence of steps to visualize the list of 3rd. parties that share user data, and access detailed information about the activity with a specific app.

with the aim to extend and complement research knowledge in a different domain: the domain of social network transparency, specially focusing on 3rd party data collection.

In the social network domain, Wei et al. [69] explored user's perceptions about Twitter ad targeting mechanisms, finding out that they consider these practices invasive and prefer detailed explanations about how their data are used. Similar to this work, we see users of the Off Facebook Activity tool demand increased detail about 3rd party data collection from FB.Farke et. al [21] conducted a user study with Google My Activity, providing evidence that users concerns about Google's data collection significantly decreased, and perceived benefit increased post exposure to My Activity. They also observed that most users would not use the features of the dashboard to modify their privacy settings nor change their behavior. In contrast to these findings, we observe stronger discomfort with data collection after using FB's transparency tool, as well as increased intention to change behavior, which might be explained by the type of information shown (3rd parties outside FB). In fact, users are not only surprised by the amount and detail of the collected data, as in the My Activity study [21], but also about the number and type of sharers (e.g., well-reputed banks).

Furthermore, transparency tools and studies related to FB are scarce. As external tool, González et al. [26] proposed a browser plugin to inform users about the real-time revenue they generate for Facebook while using it. Their Facebook Data Valuation Tool (FDTV) was efficient in raising awareness about the monetary value associated to user data. Subsequent extensions and studies with the same tool, uncovered that FB assigns sensitive ad preference labels (e.g., related to sexual orientation) to 67% of users worldwide in 2019 [9] and estimated the cost to identify user labels in only 0.015€ [10]. These tools could be integrated and used together with FB native tools to provide a broader transparency picture.

We complement and go beyond the current literature that provides insights on user reactions to online social network's transparency by: 1) analyzing the impact of a different type of transparency (data collection from 3rd parties) on user attitudes and behavior, and 2) studying the role of transparency usability, which has not been yet explored in this specific domain.

3 METHODOLOGY

This section explains our methodology to explore the impact of transparency about Facebook data collection on users' privacy attitudes and behavior, as well as to analyze the usability of the tool to satisfy users' transparency goals. We include details on the design rationale, implementation, and methods for data analysis.

3.1 General Overview

Structure. We designed an intervention study to be executed in three steps, summarized in Fig. 2. First, participants fill a *Pre-Usage Questionnaire* that registers their initial knowledge, privacy attitudes, and other demographic and background information. Second, they are instructed to use the "*Off-Facebook Activity*" transparency tool, from now on abbreviated as OFA. Third, participants answer a *Post-Usage Questionnaire* to gauge the impact of the exposure to Facebook data collection practices after using OFA and evaluate the usability of the tool. Besides the survey, we complement and contextualize the participants' usability evaluation with an expert review of the tool conducted by the authors using a heuristics-based approach [40] following Nielsen's 10 usability rules and usable privacy best practices [41, 50].

Recruitment and Ethical Aspects. We recruited participants through Mechanical Turk (MTurk)[6]. We decided to focus on English-speaking respondents[7] in countries where GDPR was in place[8]

---

[6]https://www.mturk.com/

[7]English is the common language of the authors and used in the survey.

[8]To guarantee that users' data rights regarding transparency are governed by this law, whose principles we use to discuss functionality and usability of the '*Off-Facebook Activity*" tool
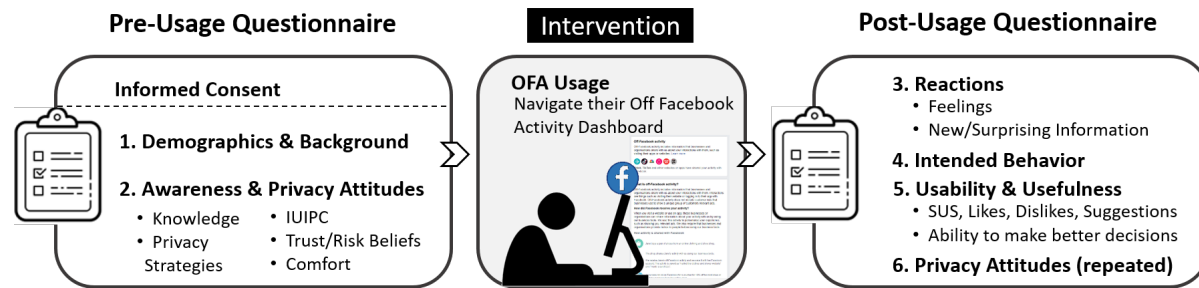
Figure 2: Structure of the user study. The study is divided into three parts: it starts with a questionnaire about initial awareness and attitudes, then users navigate their "*Off Facebook Activity*" (OFA) dashboard, and lastly they are questioned about their usage experience.

and with strong presence in MTurk [15], hence we selected UK-based participants. The study advertisement did not mention or even allude to privacy; it invited participants to a "*Facebook usage study*" to "*understand how people use social networks these days*". Participants were informed that the survey is anonymous and that all collected data is processed according to the EU General Data Protection Regulation (GDPR) [20], before asking for consent and confirmation of being over 18 years old. The questionnaires were administered via the LimeSurvey web-based survey tool[9], whose servers are located in Germany and comply with the European privacy regulations. The survey was approved by our university's Institutional Review Board.

3.2 Survey Design

Our survey includes a mix of quantitative questions, to measure privacy attitudes, behavior and usability, and open-ended qualitative questions, to get further insights on participants privacy mentalities. In the following we detail the most important question categories to address our research questions and refer the reader to the Appendix A for the full questionnaire.

- **Demographics**. We seek to understand the background of our participants, more specifically: age, gender, education, technical knowledge, and frequency of FB usage. Questions: **Q1-Q6**.
- **Awareness & Privacy Attitudes**. To evaluate the level of initial awareness, we ask participants about their knowledge on FB data collection and what strategies they use to protect their privacy. Questions: **Q26-Q44**.
  We evaluate user attitudes through the Internet Users' Information Privacy Concerns (IUIPC) scale by Malhotra et al. [43]. This construct allows us to measure multidimensional facets of privacy, including attitudes towards the *collection* of personal information, *control* over personal information, and *awareness* of privacy practices by online companies. We also borrow questions regarding trust and risk beliefs from Malhotra & Agarwal [43], as these factors usually mediate privacy-decision making processes. Additionally, we ask participants about their level of comfort with data collection. To measure the effect of using the OFA

transparency tool on privacy attitudes, we repeat this set of questions in the *Pre-Usage* and *Post-Usage* Questionnaires. Questions: **Q7-Q25**, and **Q72-Q86**.
- **Reactions & Intended Behavior.** We use qualitative questions to complement the quantitative information on privacy attitude change, in order to capture additional details about how users react to FB data collection practices. Inspired by Weinshel et al. [70], we ask participants to report what new information and surprising information they discovered when using the OFA transparency tool. We also ask about their feelings post-usage. To explore intended behavioral changes, we ask participants to report how likely they are to take specific protective measures after using the dashboard. Questions: **Q45-Q54**.
- **Usability & Usefulness** To investigate the role of usability in transparency, we ask participants to quantitatively evaluate the OFA dashboard through the standard System Usability Scale (SUS) [8], which has been widely used[10] and proved valid and reliable to measure usability [56] [38]. To find out what features were most relevant or missing, we include open ended questions on likes, dislikes, and suggestions for improvement. Additionally, we ask participants if the tool was useful for them to make better privacy decisions and how it shifted their risk-benefit perception on using Facebook. Questions: **Q55-Q71**.

3.3 Analysis

Methods and Metrics. Closed-ended responses were analyzed using targeted hypothesis testing with α =.05 , selecting the appropriate test based on the data type. We performed Wilcoxon signed-rank tests for repeated measurements (pre and post-OFA usage) on the Likert responses to IUIPC (Q7-Q16), trust and risk beliefs (Q21-Q25), and comfort (Q18)). In all cases, we first averaged responses across sub-scale items, treating the data as continuous.

Open-ended responses were analyzed following an iterative, inductive coding approach [46]. One member of the research team read responses and created the codebook with thematic codes, and a second researcher independently coded the full set of data. The inter-coder reliability for the final codes measured using Cohen's Kappa [13] was satisfactory for all questions (κ > 0.7). The cases

---

[9]https://www.limesurvey.org/

[10]Since SUS was developed by Brooke [8] in 1996, more than 2300 individual surveys were conducted using SUS in over 200 studies by 2008 [5]

where the coders differed in their final codes were discussed and reconciled.

Pilot Testing. Before publishing the study, we ran a pre-test with 9 subjects, asking them to provide feedback about the questionnaire, and a pilot with 28 users. Based on these tests, we verified that the collected data was correct and in the expected format. The experience also helped us in understanding the expected completion time (~30 min), and the feedback was useful to make minor modifications to the questionnaires regarding phrasing and order. Given the length of the study, we decided to include two attention check questions to filter bad quality responses from users that get fatigued and want to get away with the survey clicking trough it carelessly (Q24, Q71). Furthermore, to validate that the users actually opened the OFA dashboard, we asked them to upload an screenshot of the main interface, optionally showing the aggregated number of companies in their list, but including no personal information. Participants were given the possibility to donate their anonymous OFA data for enabling further research on data collection practices, and 6 of them did submit their information. The final version of the survey, which includes 87 questions, can be found in Appendix A.

Sample Size. We determined the sample size with an *a priori* power analysis using G*Power [22]. Accordingly, we observed the need for a sample size of N'= 57 for repeated measurement analysis with the Wilcoxon signed-rank tests to reach 95% power in detecting medium size effects (0.5), with alpha=0.05, two tailed. Hence, we set 57 users as a minimum requirement for our study for the quantitative analysis. Regarding the adequacy of this minimum for the qualitative analysis, research best practices recommend sample sizes of 10-20 participants [27] and usually below 50 [66] to achieve saturation, i.e., the point at which no new themes emerge. As there can be variations depending on the domain of study, we targeted a larger sample, N = 100. We finally reached saturation after 35-40 responses.

4 RESULTS

The study was conducted in April 2021. Users were recruited through MTurk with the conditions of having a HIT approval rate >95% and being located in the UK. We used MTurk premium filters to recruit verified FB account holders and have additional guarantee that the participants could access the tool.

We received 126 complete responses out of 172 submitted. The final sample, after filtering duplicates and answers that failed the check questions is N=100. It took participants, on average, around 32 minutes to complete the survey, and we paid them 10$. This compensation scheme, slightly above average[11], was intended to factor in the additional effort of using a web tool and complete a long survey. A more detailed discussion on participants effort and mechanisms to ensure data quality is provided in the Limitations (section 5.3).

Demographics. The user sample is composed of a 38% of women, a 61% of men, and 1 participant that preferred not to disclose their gender. Regarding age and education, participants are mostly young adults (72% <35 years old) with a completed bachelor, professional,

Table 1: Participant demographics

		Study Sample (n = 100)	UK FB Users (n = 49.69Million)
Gender	Female	38 %	52.9 %
	Male	61%	47.1%
	No Answer	1%	-
Age	13-17	-	4.2%
	18-24	29%	17.1%
	25-34	43%	24.6%
	35-44	16%	18.7%
	45-54	8%	14.9
	55-64	2%	11.1%
	>= 65	1%	9.5%
	No Answer	1%	-

master, or doctorate degree (60%). Other relevant factors are that we have a sample of active FB users (60% use FB daily), and that 48% of the participants have a background on IT. Table 1 provides an overview of the age and gender distribution in our sample as compared to the UK Facebook population[12]. A more granular overview of the sample characteristics can be found in Appendix B.

4.1 RQ1: Awareness & Privacy Attitudes

Initial awareness. Participants reported using FB mainly for socialization reasons (81%), like staying in touch and sharing with friends and family. When it comes to privacy protection, the majority of users (99%) do not have people they never met before as contacts, and 72% of the participants configure their privacy settings. They generally adjust settings only once (40%) and do not revisit the configuration frequently, but 11% of respondents report changing privacy settings under specific events like *"changes in the privacy policy"* or after a change in personal life circumstances, such as *"after a breakup"*. When we asked them about how they configure their *privacy settings* and what other *strategies* they use to protect their privacy (Q37, Q39), the general goal revealed in the answers is restricting access to their information by other persons. Consistently, 55% of the participants reported using restricted lists to limit access to their FB profiles. Other commonly reported strategies include limiting the amount of information they post or have in their profiles (e.g., no real name or personal details) and the actions that others can take about their data, such as tagging and posting to their timeline. Just a minority of the mentioned protections (9% of the answers) were oriented at limiting tracking or its effects, including using tracker blockers, ad blockers, cookie managers, and not linking their FB account with other services. Regarding *knowledge* about the reasons why FB collects user data, even if the vast majority of users (92%) reported being aware that FB gathers users' data for advertising purposes or marketing (Q33), only 24% of the participants had configured privacy settings to avoid being targeted by ads based on their activity within and/or outside the social network (Q41). It is specially notable how accurately participants described how they think FB uses their data:

---

[11]Minimum national wage in UK April 2021 was 8.91£/hour [64], i.e., ~11,13$/hour, and we paid 20$/hour

[12]https://napoleoncat.com/stats/facebook-users-in-united_kingdom/2021/08/

Table 2: Wilcoxon signed rank tests for IUIPC-based privacy concerns, trust and risk beliefs before and after using the "*Off-Facebook Activity*" (OFA) transparency tool. Responses were given on a 7-point Likert scale from "strongly disagree" (1) to "strongly Agree" (7).

		Pre-OFA		Post-OFA			
		Avg.	SD	Avg.	SD	Z	*p*
IUIPC	Awareness	6.29	0.69	6.35	0.77	-0.99	0.318
	Control	5.70	0.93	5.81	0.79	-0.69	0.238
	Collection	5.92	0.74	6.01	0.79	-1.184	0.236
	Total	5.97	0.55	6.06	0.61	-1.88	0.059
Trust	Honesty	3.31	1.59	3.37	1.77	-2.89	0.004
	Predictability	4.29	1.56	3.81	1.82	4.75	<.001
Risk	Privacy Loss	5.57	1.04	5.72	1.02	-1.02	0.309
	General Risk	5.08	1.33	5.37	1.09	-1.89	0.058

"To create a profile of me that allows them to target me for advertising that I am likely to engage in. And also to create categories to fit each person into that can then be used for any number of future specific targeting." (P12)

From participants' answers about their privacy awareness and behavior in FB, they seem to be well equipped to deal with social privacy but pay less attention to institutional privacy issues. While the reasons for inaction are unclear, it is plausible to think that restriction mechanisms to veto peer access are easier to understand and apply, as they resemble real life situations, but institutional tracking happens in the background, does not mirror real life situations, and therefore is not as straightforward to comprehend. Actually, from the 77 users who uploaded an OFA screenshot with a visible number of companies sharing their information with FB, 40% underestimated the extent of tracking (Q31).

Before the study, the vast majority (85%) never heard about OFA, and just 13 people had used it. Almost all the participants (97) stated that they "*would like to use a system that shows you what information has been collected about you online.*" In the following we analyse the impact of using such a tool, the OFA dashboard, on their attitudes.

Impact on attitudes. We measured the *privacy attitudes* of participants through the IUIPC scale (Q7-Q16) before and after using the OFA tool. We also measured the *trust and risk beliefs* (Q21-Q25), as well as the *comfort* with data collection (Q18) on Likert scales pre and post-exposure. To determine if there are significant changes, we performed Wilcoxon signed-rank tests. Results are summarized in Table 2 and Fig. 3.

We find that there is no effect on IUIPC. Subjects reported high levels of privacy concerns even prior to being exposed to transparency information. When it comes to *trust beliefs*, on the one hand we see a increase on trust post-usage regarding the perceived honesty of the social network ($Z = -2.89$, $p = 0.004$), with an effect size of $r = 0.35$, suggesting that this increase is moderate in size. On the other hand, trust regarding the predictable usage of users' information decreases $Z = 4.75$, $p < 0.001$), with a large effect size ($r = 0.56$). With regard to *risk beliefs*, we observe no significant changes in risk perceptions. Users consider it risky to give their information to social networks, and this remains the same after interacting with the OFA tool. Additionally, participants' comfort with data collection decreased significantly ($Z = 3.023$, $p = 0.0024$)
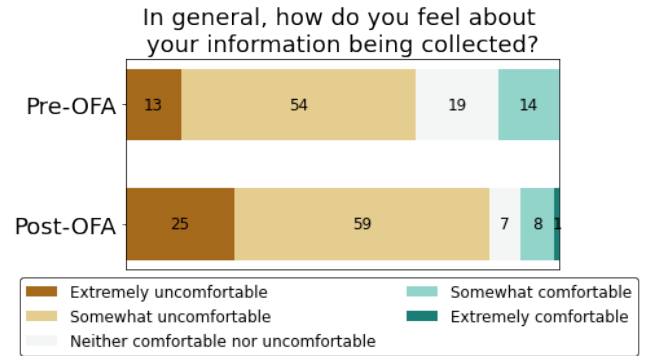

Figure 3: Proportions of the participants' level of comfort about data collection before and after using the "*Off-Facebook Activity*" (OFA) transparency tool.

after exposure to transparency information with a moderate effect size ($r = 0.45$).

While the results on privacy concerns and risk do not show significant increase post-usage, this could relate to the fact that the scores were already very high initially. The qualitative open reactions shed light on user changes in knowledge and sentiment and give further insights on why users' trust beliefs vary and why they are more uncomfortable after discovering the transparency information.

4.2 RQ2: Reactions & Intended Behavior

New and Surprising Information. After using the assigned transparency tool, users reported what *new knowledge* they gained (Q45), as well as what information they found *surprising* (Q49). Our qualitative analysis is summarized in Tables 3 and 4.

In most cases, the **new information** gained is related to the fact that there is *tracking outside FB* (49.5% of mentions) and to its extent (30.04% of mentions). More specifically, users learn about the data sources, i.e., what are the apps and webs sharing their data, but not so much about the methods used for tracking and its purpose (only 7 participants report gaining this knowledge in both cases). Regarding the *extent of tracking*, participants report learning about the amount of involved third parties and frequency of data collection. However, a lower amount of participants (n=7) report learning about the *type of data* that is being collected. In fact, this is one of the main aspects where OFA should be improved according to user suggestions as we will detail in Section 4.3.

Additionally, a 14.9% of the coded responses reveal users learning new insights about *privacy controls*. Some participants find out that there are settings to delete Off-Facebook activity and configure related ad preferences, or they realize they can download a copy of their data. A minority of the users (5.15% of the coded answers) learned about their own habits or nothing new. In this latter group, the majority were already OFA users.

The percentage of users that found **surprising information** after using the OFA tool is 67%. Table 4 summarizes the qualitative content analysis of participants' responses. Accordingly, surprising information falls into three categories. The most surprising

Table 3: Categories and codes used to code free text answers on new information gained after using the "*Off-Facebook Activity*" transparency tool. The frequency for each category appears within parentheses. The last two columns indicate the number of times and percentage a code appears in the dataset.

Category	Codes	Representative Quote	N	%
Tracking Off Facebook (49.5%)	Tracking Sources	"I found out what apps and websites are sharing my activities"	83	42.8%
	Tracking Methods	"How your information is shared when you buy something: they use business tools."	7	3.6%
	Tracking Purpose	"That the information is used to learn about your interests and offer you personalised adverts."	6	3.1%
Extent of Tracking (30.04%)	Amount of Data	"I learned there is a vast amount of data being collected and transmitted about me."	15	7.7%
	Amount of Webs/Apps	"I realised that over 1,000 applications use and have access to my data."	36	18.6%
	Data Persistence	"My data is stored for ever seemingly, which is disturbing."	1	0.5%
	Type of Tracked Data	"I now know that basically every app I use will send information of each time I have logged in and out of the app, or added an item to a wishlist or basket to Facebook."	7	3.6%
Privacy Control (14.09%)	Controls Available	'I learned that it's possible to clear this history and presumably no longer get relevant ads from those companies, until I visit their site again. I also learned that it's possible to manage which off-Facebook activity is saved."	26	13.4%
	Data Available	"That I can download my recorded activity"	3	1.5%
Other (5.15%)	Nothing	"Didn't gain any new information."	6	3.1%
	Own Habits	"Knowing more what I use."	4	2.1%

Table 4: Categories and codes used to code free text answers on surprising information gained after using the "*Off-Facebook Activity*" transparency tool. The frequency for each category appears within parentheses. The last two columns indicate the number of times and percentage a code appears in the dataset.

Category	Codes	Representative Quote	N	%
Tracking Off Facebook (36.73%)	Unexpected Webs/Apps	"They have access to apps that are supposed to be completely secure, like online banking ones."	9	9.18%
	Unknown Webs/Apps	"The most concerning element is that there are sites I don't even recognise which is worrying."	6	6.12%
	FB has Access to Data	"I did not know that Facebook collected all this information!"	12	12.24%
	Apps are tracked	"This is surprising because I didn't expect all these apps installed on my phone to be sharing this type of information with Facebook."	3	3.06%
	Tracking Methods	"It is a bit surprising that it shows sites that I have used in google without signing in or login."	6	6.12%
Extent of Tracking (55.1%)	Amount of Data	"I feel surprised that so much information is being sent to one place."	17	17.35%
	Amount of Webs/Apps	"I found it surprising how many websites had transmitted my information and stored it."	27	27.55%
	Data Persistence	"How long the data is kept for."	2	2.04%
	Type of Tracked Data	"I was surprised that they have all this information including date and time of visit"	3	3.06%
	Unconsented Exchange	"These other apps are sharing my information with Facebook when I haven't expressly consented to sharing this information."	5	5.1%
Privacy Control (7.14%)	Data Available	"Open availability of the information."	3	3.06%
	Controls Available	"I learnt how I can manage my off Facebook activities and how to disconnect it."	1	1.02%
	Public Unawareness	"How this is not at all public (or publicly known)"	3	3.06%
Other (1.22%)	Own Habits	"I didn't know that I was sharing my activity with that many websites."	1	1.02%

information is by far the *extent of tracking* (55.1% of the codes), especially the amount of information collected and the number of third parties involved. This reaction suggests that users might not be aware of how consent to share these data is managed – and several participants (n=5) alluded to this issue, for example:

> "I genuinely had no idea that so many websites and apps fed information back (sold my information?) to FB. For instance, one of my banks that I use an app for has a marketing app that sends info to FB. I mean really... FFS. If I can trust anyone, it should be my bank" (P175)

> "I didn't realise that if I visit a site, they can pass info to facebook. I assume it is because I have used the same browser to use both sites. I thought the list would be low, as I only use apps that ask "link with facebook" or similar, and I have said No to like 99% of these. But it looks like websites don't do this. Well, they may have

> those privacy pop ups, but does any one read what they allow? I don't!" (P85)

> If I don't explicitly realise or remember that I'm giving my permission to a company to share my business with them, I should be able expect confidentiality in the same way I would from a public service such as a Library or Doctors Surgery. (P12)

With regard to *tracking outside FB* (36.73% of the mentions), users are shocked to find unexpected or unknown applications and websites sharing their data. Some users are also surprised that FB has access to this information and that even offline activity and smartphone apps are tracked.

A minority of responses showed surprise about the availability of privacy controls (7.14% of mentions) and one user was surprised in relation to discovering their own habits.
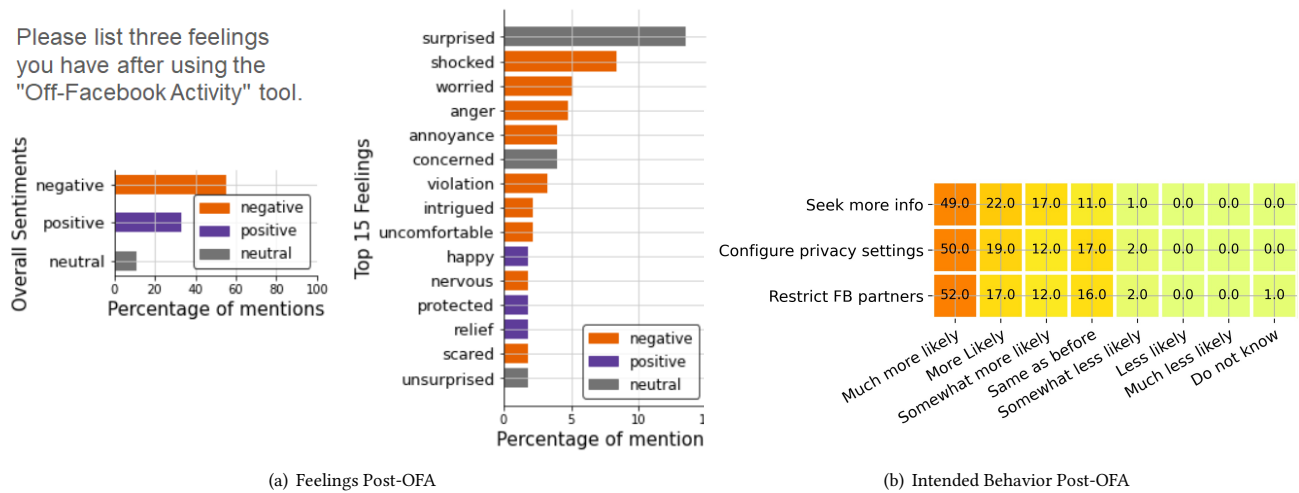
(a) Feelings Post-OFA

(b) Intended Behavior Post-OFA

Figure 4: (a) Sentiment analysis (using NRC EmoLex [48]) of reported feelings after using the "*Off-Facebook Activity*" (OFA) transparency tool. (b) Participants' reported likelihood to take specific privacy protecting actions after using the OFA transparency tool. The heatmap shows the percentage of answers to a likelihood category (x-axis) for three privacy behaviors (y-axis).

Feelings. We conducted a sentiment analysis of users' feelings after using the OFA transparency tool. We specifically asked participants to list three feelings they had (Q46). We curated this list to remove invalid answers (e.g., long sentences instead of a word), corrected misspellings, and converted verbs and nouns to adjectives where necessary. The final list contained 275 words and 87 unique feelings. We processed this list using the NRC EmoLex lexicon [48] to cluster the feelings as positive, negative, or neutral. The sentiment that dominates across all responses is negative (57,82%), while the frequency of positive feelings in the whole set of responses is only 20%. Fig. 4 shows the top 15 feelings, which cover ~60% of all user answers. As we can observe, after being exposed to transparency information, the most common reaction was *"surprise"* (13.45%), a neutral feeling. Surprise is mostly triggered, as participants explained in a follow-up *"Why do you feel like this?"* question (Q47), because they underestimated the extent of FB tracking practices. This is supported by the quantitative observation in section 4.1 about the difference in the number of companies listed by OFA and the user's preliminary estimation.

The dominant negative feelings in the top 5 are *"shocked"*, *"worried"*, and *"anger"*. These feelings come from the realization of the sheer amount of companies tracking them and the concerns about being invaded and violated, as well as preoccupation about how their data are used and shared.

In turn, the small share of positive feelings (20%) are related mainly to the gain of *knowledge* and *awareness* after using the transparency tool, *curiosity* to explore further, or *relief* to find out that they have options to control their privacy. It is to note that the majority (62%) of users that reported positive feelings already used the OFA dashboard before, which suggests that the transparency and control that comes with a continued use of the dashboard are valuable. Participant P173, for example, felt horrified, shocked, and angry, but they clarify:

'I don't currently feel this way but did the first time I accessed it. I think it is done in such a way that most normal people probably have no awareness about it. That is close to criminal in my estimation."(P173)

The prevalent negative feelings highlight the importance of raising awareness and designing transparency tools that are easy to navigate and that give clear actionable choices for users to control their data. We further elaborate on this topic during the discussion in Section 5. Now, we explore how exposure to OFA influences users' intentions to protect themselves.

Intended Behavioral Change. In the post-usage questionnaires, we asked participants to rate how using the transparency tools had changed their likelihood to take three protective actions: *i)* seek out more information about privacy settings in FB (Q51); *ii)* configure more private settings in FB (Q52); and *iii)* restrict FB partners sharing their information (Q53). Figure 4 presents the results.

Participants overwhelmingly reported increased intention to take all three privacy protective actions. Specifically, the percentage of OFA users that reported being between *"somewhat more likely"* and *"much more likely"* to take actions is: 88% for seeking information, 81% for configuring privacy, and 81% for restricting FB partners.

We also explored additional intended behavior with a follow-up question asking participants what measures would they take other than the three protective actions we specified. The majority of users reported they will explore and apply the settings offered by OFA to clear their history and manage their activity. Another common response was that they will modify their online behavior, for example, paying more attention to privacy policies, sharing less information with other apps, avoiding visiting certain webs or not using FB login. Some of the participants stated they would
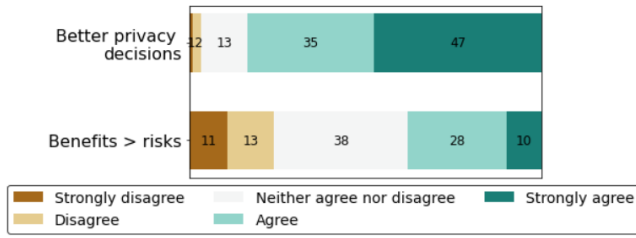
Figure 5: Participants' agreement levels to: *"feeling able to make better privacy decisions"* and considering that *"benefits of using Facebook outweigh the risks"*, after using the *"Off-Facebook Activity"* (OFA) transparency tool.

directly delete FB. Another significant observation is that several users reported they would search for similar transparency tools in other online social networks.

Overall, these results signal people interests on transparency and, similar to Weinshel et al.'s research on web tracking [70], we observe the same trend that comprehensive transparency information increases the reported privacy-preserving intentions, which underscores the value of TETs.

4.3 RQ3: Usability & Usefulness

Usability. According to ISO 9241 [32], usability is the extent to which a software tool "*can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use*". We understand the goals for a transparency tool as in the GDPR description that individuals should know "*[what] personal data concerning them are collected, used, consulted or otherwise processed and to what extent the personal data are or will be processed*" [20]. Considering these definitions, we measure and analyze the satisfaction usability dimension of the OFA tool through the SUS scale (Q58-Q67) and use open-ended questions on likes, dislikes, and suggested improvements (Q68-Q70) to capture richer information on how the tool supports the transparency goal.

Participants rated the OFA tool with a *SUS score* of 69.65 (±16.47). This value is close to, but not yet at the "acceptable" usability threshold of 70 points based on previous research [5], which indicates that there is a need for improvement.

Previous to analyzing the answers to participants' open questions, we conducted an expert heuristic review [40, 50] of the OFA user interface, observing three main areas where the tool is limited:

(1) **Transparency**. The level of transparency detail is vague. The OFA interface displays a list of third parties that sent data to Facebook (see Fig. 1). However, when clicking on a specific entry to get further details it only shows generic information, the same for every entry, providing examples of possible interactions, such as searches or purchases. But these examples constitute a limited subset of the data types captured by FB. For example, the set of events that can be shared through the Facebook Pixel[13] specifies 10 additional categories, including information of when a person books an appointment to visit a concrete location (e.g., a shop).

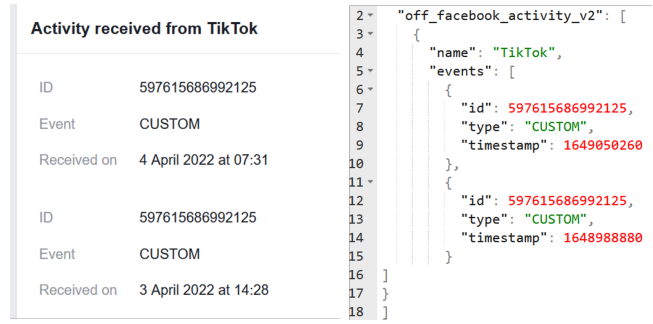[13]https://developers.facebook.com/docs/facebook-pixel/reference

Figure 6: Example excerpt of detailed activity found in an HTML (left) and JSON (file) downloaded using the *"Off-Facebook Activity"* tool.

Besides, companies can use the more opaque "CUSTOM" event to send anything they consider relevant that it is not captured in a standard event[14]. If the user wants to know the specific data exchanged, they are referred to a download page, where they can obtain a copy of their Off-Facebook activity, but it is not straightforward to discover and navigate the Off-Facebook data export. Furthermore, the download takes time and users may lose motivation due to the lack of immediacy. Even for users that have the time and ability to go through the data export, they will find scattered raw data as lists of timestamped event types per application (see Fig.6), from which is difficult to extract meaningful conclusions and, in many cases, CUSTOM events with no indication about the exact exchanged data.

(2) **Control**. OFA provides limited functionality for users to control their data, considering the intervenavility rights covered in the GDPR. It allows users to detach their Off-Facebook activity information from their Facebook account, but not to delete it (right to erasure). Users cannot withdraw consent to stop the data sharing flow (right to withdraw consent). On the positive side, OFA does allow users to get a copy of their data (right to access), and to provide feedback on their data, though not to rectify it (right to rectification).

(3) **Presentation**. The presentation of both transparency information and control options to users falls short in providing usable and clear visualizations, instructions, and feedback. Users must scroll and wait for the web page to be refreshed in order to explore the full list of companies that shared their information, which is tedious and not practical, specially when the list is long. There is no option to quickly visualize all companies at once and see how they compare to each other in terms of tracking. Though sorting options are available, the results are always shown in a *scroll to refresh* list. Furthermore, the naming of control options to detach past and future activity do not convey the fact that the information is not erased; there is a clear button accompanied by a standard trash can icon, usually representing the act

---
[14]CUSTOM events constituted 20% of the events and one of the three most common types shared for the 6 participants that donated their anonymized data in our pilot study

of deletion. These different terminology and inconsistent iconography go against usability best practices as they can lead to confusion and mislead users. Similarly, instructions to download and navigate the Off-Facebook activity are not explicit. While the tool allows to give feedback about the collected activity (unrecognized, misuse, inappropriate, other), there is no information about how this labeling is used by Facebook or how it is useful for the user, and the labeling status is not shown anywhere in the interface.

We used the above three categories of usability issues as main themes to group and contextualize users' responses about the tool usability in terms of likes, dislikes, and suggestions. The results of the qualitative analysis are summarized in Table 5.

With regard to users' **likes**, participants positively evaluated that FB *provides transparency* and shows information about *companies and interactions* involved in user data sharing (27.9% of mentions combined). They also liked that there are options to *clear their history*, *download* their data, and *manage future activity* (21.8% of mentions combined). Furthermore, some users praised the simplicity of the *interface* and how *easy to use and understand* it is (31,3% mentions combined). On the negative side, the **dislikes** highlighted important limitations in the transparency and control functionalities. The main transparency problems are that the *level of detail* of the information is not enough (10.1% of mentions), for example regarding how it was collected, in which user device, and what was the exact exchanged data. With regard to privacy controls, participants criticized the excessive effort required to remove *consent* (5.3% of mentions) . Several users mentioned that the process should be opt-in rather than opt-out, and other participants noticed that the offered controls just detach the collected data from the user account but do not actually stop the collection and processing, which is undesirable. This is an interesting point to further investigate as it is not clear if all users realize that the OFA controls do not end the information flow and what are the implications. Finally, the most frequently highlighted concern (18.1% of mentions) is the lack of *awareness* about the tool existence and how difficult it is to locate, which might impact users' trust perceptions:

> " *It should be more obvious to find this tool as I have never heard of it before today" (P64)*

> *"I find it dishonest that FB separates its privacy settings into different places, with too many screens to click through." (P45)*

These aspects related to consent and awareness were not discovered during the heuristic review previous to the user study, which highlights the value of combining both methodologies. Additionally, while many users did not dislike specific aspects of the OFA tool, they remarked their discomfort with their data being shared:

> *"I didn't dislike anything about the tool, but I disliked that all this information was being shared in the first place." (P114)*

> *"I disliked realising how much my privacy has been compromised!" (P160)*

The most relevant participants' **suggestions** to improve the tool are aligned with the disliked aspects (see Table 5) and we will summarize the main action points that can be derived from them in our discussion in Section 5.

Usefulness. To further investigate if users feel actually more capable of making better privacy decisions after using the OFA dashboard, we directly asked them. The results, together with participants' evaluation if the benefits of FB outweigh the risks, are shown in Fig. 5. An 83.3% of the participants report feeling capable of making better privacy decisions after using OFA.

As for the risk-benefits tradeoff, the results show that only a minority of users agree that risks outweigh the benefits. However we do not have additional information on how they actually made this evaluation. In this regard, it would be interesting to analyze in future studies to what extent transparency helps in understanding and judging risks.

Despite the majority of users feel more capable of making privacy decisions post-usage, free answers about intended behavior still convey the complexity involved in this type of decisions (it is not only Facebook, but an ecosystem of applications involved in tracking), the lack of clear privacy controls, and the feeling of resignation that privacy is a price to pay:

> *"I won't take any other action because I don't know how to, I feel hopeless"(P67)*

> *"Over 500 listed, I am in shock, how do I remove all of these without having to go through one by one?!" (P76)*

> *"I have turned off the feature, and cleared the history, but I am aware that FB will still use such data for ad targeting if it can collect it." (p45)*

In the light of these result, we argue that, to further empower users making privacy decisions, transparency tools should be better designed to communicate risk and provide easy, actionable, and meaningful privacy-protective options.

5 DISCUSSION

Our study contributes to a deeper understanding on the importance of transparency about online mass surveillance. We showed that when users are enabled to learn how their data are collected, they are eager to do so and they discover new insights that shift their intention to take protective actions. Even if we measured intended behavior, recent events, like the massive migration of Whatsapp users to Signal after the former published a privacy policy update implying unnecessary data sharing with FB [53], highlight the value of transparency to change behavior. Transparency tools also bring value to well-informed, privacy-active users, as these tools can help verifying that data collection and processing is done according to their expectations, i.e., their controls are working. Our findings strengthen previous work on web tracking and ad explanations transparency, where users were also moved to apply protective measures [70] after navigating their data. Contextualizing our results within the state of the art, we identify two critical issues to be addressed in the OFA transparency tool, and potentially applicable to other transparency dashboards: 1) current levels of transparency are not enough, and 2) users' real choices to control their data are limited. Additionally, usability of transparency interfaces should be improved to better support users. In the following, we discuss concrete problems and suggestions for future research and TETs implementations based on our user study and related work.

Table 5: Categories and codes used to code free text answers on likes, dislikes, and suggestions given by participants after using the "*Off-Facebook Activity*" transparency tool. The columns N and % indicate the number of times and percentage a code appears in the respective dataset. Related codes in the likes/dislikes/suggestions are shown in the same row to facilitate interpretation.

Category	Likes Code	N	%	Dislikes Code	N	%	Suggestions Code	N	%
Transparency	shows companies & interactions	22	15%						
	level of detail	1	0.7%	level of detail	9	9.6%	increase detail	11	11%
	purpose explained	1	0.7%				explain purpose	1	1%
	tracking method explained	1	0.7%				explain tracking method	1	1%
	security information	1	0.7%						
	transparency provided	19	12.9%						
Control	download data	5	3.4%						
	clear history	10	6.8%	no coarse grained clear history	2	2.1%	add coarse-grained clear	1	1%
	manage future activity	17	11.6%	no coarse-grained manage future	1	1.1%			
	controls available	8	5.4%						
				(un-)consent mechanism	5	5.3%	easy (un-)consent	12	12%
				no real delete	1	1.1%			
Presentation	easy to use/understand	32	21.8%	difficult to use/understand	7	7.4 %	make it clear/easier	3	3%
	efficient	4	2.7%						
	interface	14	9.5%	interface	7	7.4%	improve interface visualizations	10	10%
Other	nothing	10	6.8%	nothing	45	47.9%	nothing	41	41%
	good for advertisers	1	0.7%						
	password protected	1	0.7%						
				awareness	17	18.1%	awareness	19	19%
							more regulation	1	1%

5.1 Minimal Transparency is not Enough

The principle of transparency as stated in the GDPR requires that:

> "*Any information addressed to the public or to the data subject be concise, easily accessible and easy to understand, and that clear and plain language and, additionally, where appropriate, visualisation be used*".

The OFA tool, as other TETs in the literature, provides limited, ambiguous, opaque, and difficult-to-access information. Getting more granular details requires users to download raw data exports, which are difficult to navigate and interpret, and need improvement to be useful [67]. Accuracy is a problem too, as the FB interface warns users that it receives "*more details and activity than what appears here*". Participants showed their discomfort regarding the fact that not all data are shown and pointed out that the warnings about it should be more visible. For effective transparency, more accurate and comprehensive information is required and users want these details. This finding echoes previous studies in the related area of ad transparency showing that users prefer detailed information [19, 69]. Furthermore, besides concrete information about data exchange, the OFA tool falls short at providing other pieces of information required by the GDPR, namely: information about risks and consequences, clear purpose of data collection and usage, and how the data is processed (e.g., inferences). Additionally, transparency should be usable. The current usability score for the OFA tool is marginally acceptable according to the standard SUS scale. In particular, the importance of visualizations in usable transparency communication is also highlighted in the GDPR [20] and previous work on tracking transparency confirmed that users obtain the most value from graphics [70], making it a relevant element to include and to explore in further research. It would be interesting to study usability patterns for effective transparency communication, considering other transparency domains and well-researched guidelines and privacy iconography, such as the "privacy nutrition labels" [18]. Additionally, participants in our user study suggested to improve OFA with notifications. It remains to be studied if users would benefit from this type of real time or "in-context" approach to transparency, i.e., informing when the sharing occurs and not only as an ex-post all-in-one-place list of trackers. Furthermore, one of the core findings in this study is the lack of awareness and the desire of people to get knowledgeable. It is vital to raise awareness of transparency tools, transparency rights, and media competence of the general public. More than a year after the OFA tool was realised, the vast majority of our participants have not heard about it. Indeed, as they suggested, the tool could be made easier to locate and frequently advertised.

5.2 Do Users Really Have Control over their Data?

The GDPR also reads that for transparency:

> "*Natural persons should be made aware of risks, rules, safeguards and rights in relation to the processing of personal data and how to exercise their rights in relation to such processing*".

The current state of transparency in FB, also illustrated in related work in different domains, sheds doubts over how easy is for users to actually exercise their rights regarding data control. While OFA provides means to take privacy actions, they are quite limited, specially regarding data erasure and consent: users can only disconnect the data from their account[15] but they can not delete the data or

---

[15]This is not even possible for all 3rd parties, e.g., Facebook does not allow to detach data coming from Oculus VR headsets and services.

stop data collection from the interface. Furthermore, when detaching their personal data, users are only warned about the loss of functionality that the decision might entail (e.g., same amount but less relevant ads) but privacy risks are not communicated. Our participants were surprised to see unexpected applications and webs sharing their data, which suggests that current consent notices are not working well, as otherwise users would be aware of their approval to share data. Additionally, opting out is overwhelmingly complex, as it has been pointed out also in previous research [30]. In our pilot study, the 6 participants that uploaded their anonymized data, had on average 425 companies sharing their information with the social network. Reading the privacy policies of each of them to understand how to restrict data sharing and sending formal requests to ask for it, would require an excessive and unmanageable amount of time [45, 52]. Even if a highly motivated user takes the time and effort to do this, it is only a tiny part of the overall surveillance picture, so their actions can feel irrelevant, which leaves users with no real path to choose privacy. Participants suggested on/off toggles to allow for efficient batch consent withdrawal. We need standardized protocols to enable and automate this functionality. Engagement from research, technologists, and policymakers is required to devise new ways to put users in control.

5.3 Limitations

Our study is to be interpreted considering its limitations. Like many survey-based studies, our results suffer from self-report biases. To limit desirability bias (respondents misreporting their answers to make themselves look better) [39], we did not mention that the study focus was on privacy until the debriefing information provided at the end of the survey. We asked demographic questions at the beginning of the survey to increase response rates, but this might have introduced stereotype biases [47]. Since we used a survey rather than an interview, we did not have the opportunity to follow up with participants regarding answers we found interesting. However, our initial results can be a starting point for designing interview or observation-based studies to further explore how users interact with the native FB transparency tool or alternative designs.

To guarantee the quality of the collected data, we followed best practices for conducting research with MTurk [55], using high-reputation workers, additional filters to target FB account holders, and requesting proof of usage of the transparency tool. Additionally, we implemented attention checks and filtered participant answers with implausible response times. While the duration of the study can be considered long, fairness in task time advertising and payment play an important role in increasing data quality even for long surveys [42], so we opted for an above average payment as an incentive. We also piloted with a large number of users to get an accurate estimate of the time to complete the survey, so the advertised time was strongly aligned with the actual time taken by our participants. According to existing research [58, 59], survey respondents get fatigued after 20 minutes, which might impact the data. In our case, however, there were two questionnaires with a break in between, in which participants had to use the tool under evaluation. Our participants spent an average of 11 minutes in the pre-usage survey, 9 minutes using the tool, and 12 minutes answering the post-usage survey, which yields a total average time of 23

minutes answering questions in a non-continuous manner, potentially having less impact on fatigue. In fact, checking the average time effort per question in the pre and post surveys, we did not observe a reduced time investment in the post-usage questions, but it was consistent across the whole study and aligned to expected average times per question [31].

Regarding the sample, we targeted participants located in UK, recruited through Mechanical Turk. It is known that MTurk workers are generally younger and more privacy-sensitive than the the overall population [35, 71], which can be a reason why the privacy concerns are very high from the beginning. Studies with other populations and a more representative sample are thus desirable.

Furthermore, to investigate behavioral reactions to transparency, we focus on self-reported qualitative feedback about intended future behavior. We explore this dimension as a proxy to actual privacy behavior because intentions are a prerequisite [17] to take action. Though correlated, the results are to be understood as an upper bound to actual behavior, which would be affected by multiple other factors, such as usability [1]. Follow-up studies on actual behavior modification would be required to understand if action follows intention and to better understand mediating factors hindering or facilitating change.

6 CONCLUSION

In contrast to common belief that citizens are increasingly careless about their privacy, users indeed have a vital interest in the ways their data are collected and processed. Our study on FB underlines the importance of institutional transparency for users to feel informed and capable of making privacy decisions and giving informed consent to the processing of their data. Indeed, it illustrates that current consent notices are inefficient: users were mostly surprised to discover to which extent third party companies are supplying their data to Facebook, despite the fact that they supposedly gave their fully informed consent to this practice. We conjecture that online transparency requires improvement, as the information asymmetry is going to put users at increased disadvantage and severely affects their right to self-determination. This is bound to exacerbate the current surveillance economy and to negatively impact free, and democratic societies. We posit that it is necessary to raise awareness about existing transparency tools towards this end. In addition, there is demand to develop new, usable tools that provide both detailed information and insight towards potential privacy impact, as well as effective means for users to exercise control over their data.

REFERENCES

[1] Alessandro Acquisti, Laura Brandimarte, and George Loewenstein. 2020. Secrets and likes: the drive for privacy and the difficulty of achieving it in the digital age. *Journal of Consumer Psychology* 30, 4 (2020), 736–758.

[2] Quentin André, Ziv Carmon, Klaus Wertenbroch, Alia Crum, Douglas Frank, William Goldstein, Joel Huber, Leaf Van Boven, Bernd Weber, and Haiyang Yang. 2018. Consumer choice and autonomy in the age of artificial intelligence and big data. *Customer Needs and Solutions* 5, 1-2 (2018), 28–37.

[3] Athanasios Andreou, Giridhari Venkatadri, Oana Goga, Krishna Gummadi, Patrick Loiseau, and Alan Mislove. 2018. Investigating ad transparency mechanisms in social media: A case study of Facebook's explanations.

[4] Julia Angwin and Terry Parris Jr. 2016. Facebook lets advertisers exclude users by race. *ProPublica blog* 28 (2016).

[5] Aaron Bangor, Philip T Kortum, and James T Miller. 2008. An empirical evaluation of the system usability scale. *Intl. Journal of Human–Computer Interaction* 24, 6 (2008), 574–594.

[6] Susan B Barnes. 2006. A privacy paradox: Social networking in the United States. *First Monday* (2006).

[7] Susanne Barth and Menno DT De Jong. 2017. The privacy paradox–Investigating discrepancies between expressed privacy concerns and actual online behavior–A systematic literature review. *Telematics and informatics* 34, 7 (2017), 1038–1058.

[8] John Brooke et al. 1996. SUS-A quick and dirty usability scale. *Usability evaluation in industry* 189, 194 (1996), 4–7.

[9] José González Cabañas, Ángel Cuevas, Aritz Arrate, and Rubén Cuevas. 2020. Does Facebook use sensitive data for advertising purposes? *Commun. ACM* 64, 1 (2020), 62–69.

[10] José González Cabañas, Ángel Cuevas, and Rubén Cuevas. 2018. Unveiling and quantifying facebook exploitation of sensitive personal data for advertising purposes. In *27th {USENIX} Security Symposium ({USENIX} Security 18)*. 479–495.

[11] Carole Cadwalladr and Emma Graham-Harrison. 2018. *Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach (accessed: 15.09.2022) https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election.*

[12] Mary Jane Kwok Choon. 2018. Revisiting the privacy paradox on social media: An analysis of privacy practices associated with Facebook and Twitter. *Canadian Journal of Communication* 43, 2 (2018).

[13] Jacob Cohen. 1960. A coefficient of agreement for nominal scales. *Educational and psychological measurement* 20, 1 (1960), 37–46.

[14] Leucio Antonio Cutillo, Mark Manulis, and Thorsten Strufe. 2010. Security and privacy in online social networks. In *Handbook of Social Network Technologies and Applications*. Springer, 497–522.

[15] Djellel Difallah, Elena Filatova, and Panos Ipeirotis. 2018. Demographics and dynamics of mechanical turk workers. In *Proceedings of the eleventh ACM international conference on web search and data mining*. 135–143.

[16] Claire Dolin, Ben Weinshel, Shawn Shan, Chang Min Hahn, Euirim Choi, Michelle L Mazurek, and Blase Ur. 2018. Unpacking perceptions of data-driven inferences underlying online targeting and personalization. In *Proceedings of the 2018 CHI conference on human factors in computing systems*. 1–12.

[17] Serge Egelman, Marian Harbach, and Eyal Peer. 2016. Behavior ever follows intention? A validation of the Security Behavior Intentions Scale (SeBIS). In *Proceedings of the 2016 CHI conference on human factors in computing systems*. 5257–5261.

[18] Pardis Emami-Naeini, Yuvraj Agarwal, Lorrie Faith Cranor, and Hanan Hibshi. 2020. Ask the experts: What should be on an IoT privacy and security label?. In *2020 IEEE Symposium on Security and Privacy (SP)*. IEEE, 447–464.

[19] Motahhare Eslami, Sneha R Krishna Kumaran, Christian Sandvig, and Karrie Karahalios. 2018. Communicating algorithmic process in online behavioral advertising. In *Proceedings of the 2018 CHI conference on human factors in computing systems*. 1–13.

[20] EU. 2016. General Data Protection Regulation. (2016). Available at: https://gdpr-info.eu/ (accessed: 15.09.2022).

[21] Florian M Farke, David G Balash, Maximilian Golla, Markus Dürmuth, and Adam J Aviv. 2021. Are Privacy Dashboards Good for End Users? Evaluating User Perceptions and Reactions to Google's My Activity. In *30th {USENIX} Security Symposium ({USENIX} Security 21)*. 483–500.

[22] Franz Faul, Edgar Erdfelder, Axel Buchner, and Albert-Georg Lang. 2009. Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior research methods* 41, 4 (2009), 1149–1160.

[23] Geoffrey A Fowler. 2020. Facebook Will Now Show You Exactly How It Stalks You—Even When You're Not Using Facebook.". *Washington Post* (2020).

[24] Vaibhav Garg, Kevin Benton, and L Jean Camp. 2014. The privacy paradox: a Facebook case study. In *2014 TPRC conference paper*.

[25] Ghostery GmbH. 2020. *Ghostery website, (accessed: 15.09.2022).* https://www.ghostery.com/.

[26] José González Cabañas, Angel Cuevas, and Rubén Cuevas. 2017. FDVT: Data valuation tool for Facebook users. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 3799–3809.

[27] Greg Guest, Arwen Bunce, and Laura Johnson. 2006. How many interviews are enough? An experiment with data saturation and variability. *Field methods* 18, 1 (2006), 59–82.

[28] Barbara Guidi, Marco Conti, Andrea Passarella, and Laura Ricci. 2018. Managing social contents in decentralized online social networks: a survey. *Online Social Networks and Media* 7 (2018), 12–29.

[29] Seda Gürses and Claudia Diaz. 2013. Two tales of privacy in online social networks. *IEEE Security & Privacy* 11, 3 (2013), 29–37.

[30] Hana Habib, Sarah Pearman, Jiamin Wang, Yixin Zou, Alessandro Acquisti, Lorrie Faith Cranor, Norman Sadeh, and Florian Schaub. 2020. "It's a scavenger hunt": Usability of Websites' Opt-Out and Data Deletion Choices. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–12.

[31] Emmanuel M Ikart. 2019. Survey questionnaire survey pretesting method: An evaluation of survey questionnaire via expert reviews technique. *Asian Journal of Social Science Studies* 4, 2 (2019), 1.

[32] ISO. 1998. ISO 9241-11:1998 (en). Ergonomic requirements for office work with visual display terminals (VDTs) — Part 11: Guidance on usability. (1998). Available at: https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-1:v1:en, accessed: 15.09.2022.

[33] Sonia Jahid, Prateek Mittal, and Nikita Borisov. 2011. EASiER: Encryption-based access control in social networks with efficient revocation. In *Proceedings of the 6th ACM Symposium on Information, Computer and Communications Security*. 411–415.

[34] Milena Janic, Jan Pieter Wijbenga, and Thijs Veugen. 2013. Transparency enhancing tools (TETs): an overview. In *2013 Third Workshop on Socio-Technical Aspects in Security and Trust*. IEEE, 18–25.

[35] Ruogu Kang, Stephanie Brown, Laura Dabbish, and Sara Kiesler. 2014. Privacy Attitudes of Mechanical Turk Workers and the {US}. Public. In *10th Symposium On Usable Privacy and Security (SOUPS 2014)*. 37–49.

[36] Imrul Kayes and Adriana Iamnitchi. 2017. Privacy and security in online social networks: A survey. *Online Social Networks and Media* 3 (2017), 1–21.

[37] Saranga Komanduri, Richard Shay, Greg Norcie, and Blase Ur. 2011. Adchoices-compliance with online behavioral advertising notice and choice requirements. *ISJLP* 7 (2011), 603.

[38] Philip Kortum and S Camille Peres. 2014. The relationship between system effectiveness and subjective usability scores using the System Usability Scale. *International Journal of Human-Computer Interaction* 30, 7 (2014), 575–584.

[39] Frauke Kreuter, Stanley Presser, and Roger Tourangeau. 2008. Social desirability bias in cati, ivr, and web surveysthe effects of mode and question sensitivity. *Public opinion quarterly* 72, 5 (2008), 847–865.

[40] Jonathan Lazar, Jinjuan Heidi Feng, and Harry Hochheiser. 2017. *Research methods in human-computer interaction*. Morgan Kaufmann.

[41] Scott Lederer, Jason I Hong, Anind K Dey, and James A Landay. 2004. Personal privacy through understanding and action: five pitfalls for designers. *Personal and ubiquitous computing* 8, 6 (2004), 440–454.

[42] Matt Lovett, Saleh Bajaba, Myra Lovett, and Marcia J Simmering. 2018. Data quality from crowdsourced surveys: A mixed method inquiry into perceptions of Amazon's Mechanical Turk Masters. *Applied Psychology* 67, 2 (2018), 339–366.

[43] Naresh K Malhotra, Sung S Kim, and James Agarwal. 2004. Internet users' information privacy concerns (IUIPC): The construct, the scale, and a causal model. *Information systems research* 15, 4 (2004), 336–355.

[44] Paris Martineau. 2018. *Facebook is tracking you on over 8.4 million websites (accessed: 15.09.2022), https://theoutline.com/post/4578/facebook-is-tracking-you-on-over-8-million-websites.*

[45] Aleecia M McDonald and Lorrie Faith Cranor. 2008. The cost of reading privacy policies. *Isjlp* 4 (2008), 543.

[46] Matthew B Miles and A Michael Huberman. 1994. *Qualitative data analysis: An expanded sourcebook*. Sage.

[47] Barbara Mirel. 1990. Usability and hardcopy manuals: evaluating research designs and methods. In *Proceedings of the 8th annual international conference on Systems documentation*. 69–77.

[48] Saif M. Mohammad and Peter D. Turney. 2013. Crowdsourcing a Word-Emotion Association Lexicon. 29, 3 (2013), 436–465.

[49] Patrick Murmann and Simone Fischer-Hübner. 2017. Tools for achieving usable ex post transparency: a survey. *IEEE Access* 5 (2017), 22965–22991.

[50] Jakob Nielsen and Rolf Molich. 1990. Heuristic evaluation of user interfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 249–256.

[51] Jason RC Nurse and Oliver Buckley. 2017. Behind the scenes: a cross-country study into third-party website referencing and the online advertising ecosystem. *Human-centric Computing and Information Sciences* 7, 1 (2017), 1–21.

[52] Jonathan A Obar and Anne Oeldorf-Hirsch. 2020. The biggest lie on the internet: Ignoring the privacy policies and terms of service policies of social networking services. *Information, Communication & Society* 23, 1 (2020), 128–147.

[53] Financial Express Online. 2020. *Signal downloads see record surge amid growing concerns over WhatsApp's updated privacy policy,(accessed: 15.09.2022).* https://www.financialexpress.com/industry/technology/signal-downloads-see-record-surge-amid-growing-concerns-over-whatsapps-updated-privacy-policy/2169500/

[54] Thomas Paul, Martin Stopczynski, Daniel Puscher, Melanie Volkamer, and Thorsten Strufe. 2012. C4PS-helping Facebookers manage their privacy settings.

In *International Conference on Social Informatics.* Springer, 188–201.

[55] Eyal Peer, Joachim Vosgerau, and Alessandro Acquisti. 2014. Reputation as a sufficient condition for data quality on Amazon Mechanical Turk. *Behavior research methods* 46, 4 (2014), 1023–1031.

[56] S Camille Peres, Tri Pham, and Ronald Phillips. 2013. Validation of the system usability scale (SUS) SUS in the wild. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 57. SAGE Publications Sage CA: Los Angeles, CA, 192–196.

[57] Philip Raschke, Axel Küpper, Olha Drozd, and Sabrina Kirrane. 2017. Designing a GDPR-compliant and usable privacy dashboard. In *IFIP international summer school on privacy and identity management.* Springer, 221–236.

[58] Melanie Revilla and Jan Karem Höhne. 2020. How long do respondents think online surveys should be? New evidence from two online panels in Germany. *International Journal of Market Research* 62, 5 (2020), 538–545.

[59] Melanie Revilla and Carlos Ochoa. 2017. Ideal and maximum length for a web survey. *International Journal of Market Research* 59, 5 (2017), 557–565.

[60] Martina Angela Sasse, Sacha Brostoff, and Dirk Weirich. 2001. Transforming the 'weakest link'—a human/computer interaction approach to usable and effective security. *BT technology journal* 19, 3 (2001), 122–131.

[61] Florian Schaub, Aditya Marella, Pranshu Kalvani, Blase Ur, Chao Pan, Emily Forney, and Lorrie Faith Cranor. 2016. Watching them watching me: Browser extensions impact on user privacy awareness and concern. In *NDSS workshop on usable security.* 1–10.

[62] M. Schufrin, S. Reynolds, A. Kuijper, and J. Kohlhammer. 2021. A Visualization Interface to Improve the Transparency of Collected Personal Data on the Internet. *IEEE Transactions on Visualization and Computer Graphics* 27, 02 (feb 2021), 1840–1849. https://doi.org/10.1109/TVCG.2020.3028946

[63] Till Speicher, Muhammad Ali, Giridhari Venkatadri, Filipe Ribeiro, George Arvanitakis, Fabrício Benevenuto, Krishna Gummadi, Patrick Loiseau, and Alan Mislove. 2018. Potential for discrimination in online targeted advertising. In *FAT 2018-Conference on Fairness, Accountability, and Transparency*, Vol. 81. 1–15.

[64] UK.GOV. 2022. *National Minimum Wage and National Living Wage rates, (accessed: 15.09.2022).* https://www.gov.uk/national-minimum-wage-rates

[65] Tobias Urban, Martin Degeling, Thorsten Holz, and Norbert Pohlmann. 2019. "Your hashed IP address: Ubuntu." perspectives on transparency tools for online advertising. In *Proceedings of the 35th Annual Computer Security Applications Conference.* 702–717.

[66] Frank J van Rijnsoever. 2017. (I Can't Get No) Saturation: A simulation and guidelines for sample sizes in qualitative research. *PloS one* 12, 7 (2017), e0181689.

[67] Sophie Veys, Daniel Serrano, Madison Stamos, Margot Herman, Nathan Reitinger, Michelle L Mazurek, and Blase Ur. 2021. Pursuing Usable and Useful Data Downloads Under {GDPR/CCPA} Access Rights via {Co-Design}. In *Seventeenth Symposium on Usable Privacy and Security (SOUPS 2021).* 217–242.

[68] Ken Ward. 2018. Social networks, the 2016 US presidential election, and Kantian ethics: applying the categorical imperative to Cambridge Analytica's behavioral microtargeting. *Journal of media ethics* 33, 3 (2018), 133–148.

[69] Miranda Wei, Madison Stamos, Sophie Veys, Nathan Reitinger, Justin Goodman, Margot Herman, Dorota Filipczuk, Ben Weinshel, Michelle L Mazurek, and Blase Ur. 2020. What Twitter Knows: Characterizing Ad Targeting Practices, User Perceptions, and Ad Explanations Through Users' Own Twitter Data. In *29th {USENIX} Security Symposium ({USENIX} Security 20).* 145–162.

[70] Ben Weinshel, Miranda Wei, Mainack Mondal, Euirim Choi, Shawn Shan, Claire Dolin, Michelle L Mazurek, and Blase Ur. 2019. Oh, the Places You've Been! User Reactions to Longitudinal Transparency About Third-Party Web Tracking and Inferencing. In *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security.* 149–166.

[71] Huichuan Xia, Yang Wang, Yun Huang, and Anuj Shah. 2017. " Our Privacy Needs to be Protected at All Costs" Crowd Workers' Privacy Experiences on Amazon Mechanical Turk. *Proceedings of the ACM on Human-Computer Interaction* 1, CSCW (2017), 1–22.

[72] Alyson Leigh Young and Anabel Quan-Haase. 2013. Privacy protection strategies on Facebook: The Internet privacy paradox revisited. *Information, Communication & Society* 16, 4 (2013), 479–500.

[73] Elena Zheleva and Lise Getoor. 2011. Privacy in social networks: A survey. In *Social network data analytics.* Springer, 277–306.

A   SURVEY INSTRUMENT

A.1   Pre-Usage Questionnaire

A.1.1   Demographics and Background.

Q1   With what gender do you identify? (□Male □Female □Non-binary □Other □Prefer not to say)

Q2   How old are you? (□18-24 □25-34 □35-44 □45-54 □55-64 □65 or older □Prefer not to say)

Q3   What is the highest degree or level of education you have completed? (□Some high school □High school □Some college □Trade, technical, or vocational training □Bachelor's degree □Master's degree □Professional degree □Some graduate School □Doctorate □Prefer not to say)

Q4   What nationality do you most identify with?

Q5   Which of the following best describes your educational background or job field? (□I have an education in, or work in, the field of computer science, engineering, or IT. □I do not have an education in, or work in, the field of computer science, engineering, or IT. □Prefer not to say.)

Q6   Approximately, how often do you access your Facebook account? (□Every day □A few times a week □About once a week □A few times a month □Once a month □Less than once a month)

A.1.2   Privacy Attitudes.   **IUIPC Awareness, Collection and Control Subscales**

(Answer choices: □Strongly agree □Agree □Somewhat Agree □Neither agree nor disagree □Somewhat Disagree □Disagree □Strongly disagree)

Q7   Companies seeking information online should disclose the way the data are collected, processed, and used.

Q8   A good consumer online privacy policy should have a clear and conspicuous disclosure.

Q9   It is very important to me that I am aware and knowledgeable about how my personal information will be used.

Q10   It usually bothers me when online companies ask me for personal information.

Q11   When online companies ask me for personal information, I sometimes think twice before providing it.

Q12   It bothers me to give personal information to so many online companies.

Q13   I'm concerned that online companies are collecting too much personal information about me.

Q14   Consumer online privacy is really a matter of users' right to exercise control and autonomy over decisions about how their information is collected, used, and shared.

Q15   Consumer control of personal information lies at the heart of consumer privacy.

Q16   You believe that online privacy is invaded when control is lost or unwillingly reduced as a result of a marketing transaction.

General Attitudes

Q17   In your own words, how do you explain privacy?

Q18   In general, how do you feel about your information being collected? (Answer choices: □Extremely uncomfortable, □Somewhat uncomfortable, □Neither comfortable nor uncomfortable, □Somewhat comfortable □Extremely comfortable)

Q19   You would only provide accurate and personal information at SNSs if their control policy is verified / monitored by a reputable third party.(Answer choices: □Strongly agree □Agree □Neither agree nor disagree □Disagree □Strongly disagree)

Q20   If it were available, you would like to use a system that shows you what information has been collected about you online.(Answer choices: □Strongly agree □Agree □Neither agree nor disagree □Disagree □Strongly disagree)

Trust and Risk Beliefs

(Answer choices: □Strongly agree □Agree □Neither agree nor disagree □Disagree □Strongly disagree)

Q21   SNSs are in general predictable and consistent regarding the usage of the information.

Q22 SNSs are always honest with users when it comes to using the information that users would provide.

Q23 In general, it would be risky for a user when he/she gives his/her personal information in SNSs.

Q24 We use this question to discard the answers of people who are not reading the questions. Please select "Strongly Agree" to preserve your answers. [*Check question*]

Q25 There would be high potential for privacy loss associated with giving personal information to SNSs.

A.1.3   Awareness. **Experience with Off-Facebook Activity**

Q26 Have you ever heard about the "Off-Facebook Activity" privacy tool?(□Yes □No)

Q27 If Q26=Yes: Please list the information that you already knew about the "Off-Facebook Activity" tool.

Q28 Have you ever used the "Off-Facebook Activity" tool?(□Yes □No)

Q29 If Q28=Yes: How did you use the "Off-Facebook Activity" tool? (□You have cleared history of your activities □You have disconnected your activities from your Facebook account by turning of the "Off-Facebook Activity" privacy tool □You have downloaded your activities from your Facebook account □You have accessed the information that Facebook has about you)

Q30 If Q28=Yes: Do you like to clear your "Off-Facebook Activity" for privacy reasons?(□Yes □No)

Knowledge about Facebook Data Collection

Q31 How many companies do you think that share their information about you with Facebook? (□0 companies □Less than 20 companies □20-60 □60-100 □More than 100 companies)

Q32 Which types of information do you think Facebook collects as you use it? (□Websites that you visit □Every advertisement topic you have clicked □A list of all companies that has your contact information from the advertisements you have clicked □All of your contact information from your phone book □Your online purchases □Your exact location □Every social event you are invited to through Facebook □Your personal information (name, email, gender, ...) □When you use an app in your smartphone □Other)

Q33 In your own words, please describe the purposes for which you think the information you selected above is collected.

Privacy Behavior in Facebook

Q34 Who are your friends on Facebook? (□Friends □Friends of friends □Professors □Colleagues □Boss □Parents/uncles/aunts □People you have never met □Patients □Other)

Q35 Do you use restricted lists to limit access to your profile?(□Yes □No)

Q36 Do you use your Privacy Settings on Facebook?(□Yes □No)

Q37 If Q36=Yes: How do you configure Facebook settings to protect your privacy?

Q38 If Q36=Yes: How often do you change your Privacy Settings on Facebook?(□Never □I did it once Frequently □When something changes (e.g., when you see a policy change notification, when you add a new friend, when you enter a relationship, etc). Please specify)

Q39 If you follow any other strategy to protect your privacy in Facebook, please briefly explain it. (For example: not using your real name to avoid being searchable, configuring Tinder so that you are not discoverable for Facebook friends, not tagging your pictures)

Q40 Have you ever looked at your Facebook ad preferences?(□Yes □No)

Q41 Have you ever configured Privacy Settings related to Facebook advertisements?(□Yes □No)

Q42 If Q41=Yes: How did you change your ads Settings?(□You do not allow ads based on the data from Facebook partners □You do not allow ads based on your activity on Facebook company product that you see elsewhere □You do not allow ads that include your social actions □Other)

Q43 Do you follow any strategies to prevent Facebook from collecting information about you?(□Yes □No)

Q44 If Q43=Yes: Please explain these strategies.

A.2   Intervention

Now, please access your Facebook account on a computer and access the Off-Facebook Activity tool following these steps:

- Click on the top right icon of your Facebook profile. Select "Settings & Privacy" > "Settings".
- Click on "Your Facebook Information" on the left column.
- Click on "Off-Facebook Activity". From this screen, you can also click "Manage Your Off-Facebook Activity" for more information. You'll be asked to re-enter your password.

If you could not find the option, here is a direct link: https://www.facebook.com/off_facebook_activity/

Once you access the Off Facebook Activity interface, please take a screenshot of the summary information. It should look similar to the image in Figure 7.
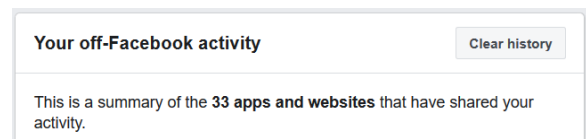

Figure 7: Sample Screenshot of the Off-Facebook Activity tool

Don't include any personal information. Please **spend some time using and exploring the Off-Facebook Activity tool** and come back to answer the remaining questions about your experience.

A.3   Post-Usage Questionnaire

A.3.1   Open-Ended Reactions.

Q45 Please list the new information that you gained by using this tool.

Q46 Please list three feelings you have after using the "Off-Facebook Activity" tool.

Q47 Why do you feel this way? (Please explain for each of the listed feelings.)

Q48 Did you find any surprising information by using this tool? (□Yes □No)

Q49 If Q48=Yes: Please explain these surprising information in your own words.

Q50 Do you have any question about what you saw in the tool?

A.3.2   Post-Usage Intended Behavior. Compared to before you used the tool,

Q51 How likely are you to seek out more information about possible privacy-related settings on Facebook?

Q52 How likely are you to configure more private settings on your Facebook?

Q53 How likely are you to restrict Facebook's partners sharing your information? Answer choices for all questions: □Much more likely □More likely □Somewhat more likely □About the same as before □Somewhat less likely □Less likely □Much less likely □Don't know

Q54 Please, mention any other action that you have taken or will take to protect your privacy as a result of using this tool.

A.3.3 Usefulness.

Q55 Compared to before you used the tool, you feel that you are able to make better decisions about your privacy-related actions.(□Strongly agree □Agree □Neither agree nor disagree □Disagree □Strongly disagree)

Q56 Why do you use Facebook?

Q57 After using the tool, you think the benefits of using Facebook outweigh the privacy risks. (□Strongly agree □Agree □Neither agree nor disagree □Disagree □Strongly disagree)

A.3.4 Usability.

System Usability Scale (SUS). During the rest of this survey, we use the term "system" to refer the Off-Facebook activity tool. Next, we show statements about your experiences with the system. Please select the answer choice that best describes your agreement. (Answer choices: □Strongly agree □Agree □Neither agree nor disagree □Disagree □Strongly disagree)

Q58 I think that I would like to use this system frequently

Q59 I found the system unnecessarily complex

Q60 I thought the system was easy to use

Q61 I think that I would need the support of a technical person to be able to use this system

Q62 I found the various functions in this system were well integrated

Q63 I thought there was too much inconsistency in this system

Q64 I would imagine that most people would learn to use this system very quickly

Q65 I found the system very cumbersome to use

Q66 I felt very confident using the system

Q67 I needed to learn a lot of things before I could get going with this system

Likes, Dislikes and Suggestions

Q68 List the aspects, if any, that you liked about the Off-Facebook Activity/Facebook Data Visualization tool.

Q69 List the aspects, if any, that you disliked about the Off-Facebook Activity/Facebook Data Visualization tool.

Q70 Do you have any suggestion(s) to improve the Off-Facebook Activity/Facebook Data Visualization tool? If so, please elaborate below.

Q71 We use this question to discard the answers of people who are not reading the questions. Please select "Strongly Agree" to preserve your answers. [*Check question*]

A.3.5 Post-Usage Privacy Attitudes and Awareness.

Q72-Q86 We repeat questions Q7-Q16, Q18, Q21-Q23, and Q25.

A.3.6 General Feedback Question.

Q87 Is there anything else you would like to add about Facebook privacy practices or this survey in general?

B DETAILED DEMOGRAPHICS

Table 6: Detailed participant demographics

		n = 100
Gender	Female	38 %
	Male	61%
	No Answer	1%
Age	18-24	29%
	25-34	43%
	35-44	16%
	45-54	8%
	55-64	2%
	>= 65	1%
	Prefer not to say	1%
Education	Some high school	1
	High school	10
	Some college	24
	Bachelor's degree	38
	Trade, technical, or vocational training	4
	Professional degree	2
	Master's degree	18
	Doctorate	2
	Prefer not to say	1
IT	No	52
	Yes	45
	Prefer not to say	3
FB Usage	Every day	60
	A few times a week	27
	About once a week	1
	A few times a month	5
	Once a month	5
	Less than once a month	2