Drivers and Persuasive Strategies to Influence User Intention to Learn About Manipulative Design

Pooria Babaei
Julita Vassileva*
pooria.babaei@usask.ca
jiv@cs.usask.ca
University of Saskatchewan
Saskatoon, Saskatchewan, Canada

ABSTRACT

The proliferation of e-commerce, game, and social networking sites, has brought to light the use of "dark patterns" or, more generally, manipulative designs (MDs), which exploit psychological effects and cognitive biases of users to channel their behavior toward outcomes that benefit the company or owner of the site, against the users' best interests. Previous research has categorized MDs, assessed their impact on users, gauged their prevalence, and attempted automated detection using computer vision and natural language processing techniques. However, limited attention has been given to understanding how to warn and educate users about MDs, guiding them to recognize and resist such manipulative tactics. To address this gap, we carried out a controlled study with n=134 participants, using a survey based on the Protection Motivation Theory (PMT) to better understand the motivations of people to learn about MDs. We also explored the effectiveness of two persuasive strategies, based on Cialdini's principles of influence (social influence and authority), to trigger attention towards MDs and intention to learn more about MDs and to avoid them. For this, we created a simulated application in a mobile app distribution platform modeled like Google Play Store containing a visual signal, a warning based on one of the two strategies, and simulated reviews from other users. The results indicate that two of the five PMT constructs - a higher Perceived Severity of MDs and a lower Perceived Response Cost of learning about MDs - have the most significant influence on the Intention to learn more about MDs. The participants in the experimental group, exposed to the two persuasive strategies exhibited a larger increase in their intention to seek information about MDs than the participants in the control group. Our study showcases the potential of a persuasive intervention, illustrating how mobile app distribution platforms can enhance user protection against MD exploitation. By implementing such interventions, these platforms can boost accountability and transparency of applications existing on their platform, and MD awareness among their users.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
FAccT '24, June 03–06, 2024, Rio de Janeiro, Brazil
© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 979-8-4007-0450-5/24/06
https://doi.org/10.1145/3630106.3659046

CCS CONCEPTS

• **Human-centered computing**; • **Security and privacy** → *Human and societal aspects of security and privacy*; • **Social and professional topics** → **Computing / technology policy**;

KEYWORDS

persuasive strategies, manipulative designs, dark patterns, protection motivation theory

ACM Reference Format:
Pooria Babaei and Julita Vassileva. 2024. Drivers and Persuasive Strategies to Influence User Intention to Learn About Manipulative Design. In *The 2024 ACM Conference on Fairness, Accountability, and Transparency (FAccT '24), June 03–06, 2024, Rio de Janeiro, Brazil.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3630106.3659046

1 INTRODUCTION

The main objective of designers when creating interfaces for mobile applications is to help users achieve their objectives by making them easy to understand and presenting information in a way that is easy to access. Lately, there has been a significant increase in the use of technology to affect how humans make decisions. By understanding psychological principles that govern human thinking, persuasive technologies and designs can create user experiences that intentionally and effectively influence people's behavior in ways that are beneficial for them, for example, encourage them to engage in physical activities, adhere to therapies, or be more engaged learners. While persuasive design strategies are not inherently alarming, the knowledge of how the human mind operates can also be exploited for unethical purposes by manipulating users' decision-making process in ways that go against their goals but benefit other stakeholders, for example, the owners of the company. Many websites, computer games, and mobile applications exploit the user by applying deceitful design elements, known as manipulative design or "dark patterns". Signaling, warning, and educating users about them is an important goal towards achieving accountability and transparency in application marketplaces. Our paper focuses on this problem. To avoid negative racial connotations associated with the word "dark", we use the term "manipulative design" (MD) as equivalent to the terms "dark pattern", "malicious design", or "deceptive pattern" that occur in the existing literature. So far, studies have categorized manipulative designs, studied their impact on users, and measured their prevalence. There are also a few recent works that aim to counteract MDs by auto-detecting them and creating friction around them. There is a lot of information regarding MD techniques available on the Internet for those

interested to learn about them. However, there has been little work on how to proactively inform/warn users about MDs and how to influence users' engagement with information about MDs. In other words, unlike other risk-related domains, such as cybersecurity, less is known about what motivates people to seek information and guidance related to avoiding risks from MD. Motivating people to avoid risky behaviors is an area of behavior change.

Persuasive technology [9, 10, 26] has been used effectively to promote behavior change in various domains, such as avoiding risky smartphone behaviors [18], irresponsible alcohol consumption [41], mental health [6], warning compliance [37, 51]. Remarkably, despite the extensive exploration of persuasive strategies in various contexts, there is a gap in existing literature concerning the application of these strategies to warn and educate users about the presence of manipulative designs. To the best of our knowledge, no prior research has systematically investigated the efficacy of employing persuasive strategies to alert users to the existence of MD and assess the impact of such strategies on enhancing users' engagement in seeking information about MDs.

In this paper, we explore the factors that influence whether users engage with information about MD. For this purpose, we use the self-reported intention of participants and utilize the Protection Motivation Theory (PMT) as the theoretical lens to discover the determinants of user behavioral intention (BI). Furthermore, we aim to examine the novel application of persuasive strategies, specifically *authority* and *social influence*, in the context of MD warnings inside an app distribution platform and evaluate the effectiveness of these strategies in promoting user awareness and information-seeking behavior regarding MDs. We chose Google Play Store as a highly visited platform. Therefore, we form our research questions as follows:

RQ1 - What are the PMT determinants of behavioral intention to seek information about manipulative designs based on the protection motivation theory?

RQ2 - Can behavioral intention for information seeking about MDs be amplified by the use of persuasive strategies?

The rest of the paper is organized as follows. In section 2, we provide background on persuasive strategies and MD, previous work on fighting MD, and the background on the Protection Motivation Theory (PMT). In section 3, we present the study design and in section 4 - the results. Finally, in section 5, we discuss the results of the study, the implications for intervention design, the limitations of our work, and possible future directions.

2 BACKGROUND AND RELATED WORK

Fogg introduced the term *Persuasive Technology* in 2002, describing it as "a computing system, device, or application intentionally designed to change a person's attitude or behavior in a predetermined way" [15]. Through the utilization of psychological principles in designing interfaces, designers can effectively share information with users, assist them in making decisions, encourage them to achieve their objectives, support the development of their skills, and potentially facilitate the formation or alteration of habits.

Different persuasion techniques have been offered, such as the 40 *persuasive strategies* introduced by Fogg [15], the six *principles of influence* by Cialdini (a seventh one has been added more recently) [9], or the 64 *compliance-gaining strategies* by Kellermann and Cole [27]. As far as we know, there is no prior work employing persuasive strategies in the interventions to increase user awareness about MDs. Therefore, in this study, we want to explore the use of two persuasive strategies, specifically, Cialdini's *authority* and *social influence* for this purpose. We chose to use Cialdini's strategies because they are well-known and applied widely in many domains and purposes. Authority and Social Influence seem to be the only appropriate strategies for the task of warning users about MD. Many studies have confirmed that people trust information more if it comes from a credible source (either from some authoritative source or from observing other people's opinions or behaviour) [2, 42, 57]. The *social influence* strategy suggests that people tend to copy the behaviors of others, especially when they are uncertain what to do [47]. In a recent study, Wang et al. [51] showed that authority and social influence strategies significantly improved the effectiveness of warnings against online fraud.

Although persuasive technology is often lauded as a tool for achieving behavior changes that benefit the users' health, the environment, or society, persuasive technology can be used for less noble purposes, so ethical concerns must be taken into account [21]. The designers' intentions need to be evaluated to distinguish between persuasive techniques and manipulative designs. When employing persuasive techniques, designers seek to motivate users to engage in actions that are intended to benefit them or their environment. Most commercial persuasive systems benefit both the user and the systems' owner. However, in manipulative design (MD), the persuasive techniques are no longer created in a user-centered manner (i.e. to benefit the user) but are rather business-centric [21] and deliberately designed to change the user behavior so that it can be exploited towards the goals of the system owners or shareholders without any benefit or to the detriment of the user. It is worth noting however that a negative user experience can occur unintentionally due to a lack of technical skills, inexperience, or little knowledge of the user needs by the designer [22]; this is known as an anti-pattern and is not in the scope of this research.

2.1 Manipulative designs and previous work on countering them

One of the earliest works on dark patterns or MDs, is the informative website by Harry Brignull in 2010 [24], where he introduced the term "dark patterns" and provided a framework for classifying them. So far, Brignull has identified 15 types of dark patterns in detail, including "trick wording", "hard to cancel", "forced action", "confirmshaming", "preselection", "fake urgency", "fake scarcity", "fake social proof", "obstruction", "hidden subscription", "comparison prevention", "nagging", "disguised ads", "visual interference", and "sneaking".

MDs are widely used despite the ethical concerns that have been raised. In [36], the authors conducted a study where they utilized a website crawler to examine 11,000 widely used e-commerce websites to determine the prevalence of MDs in e-commerce. The results of the study showed that 11% of the websites analyzed contained elements that qualified as MDs. Additionally, the study found that the use of deceitful patterns was more prevalent in more popular e-commerce websites. In another research [38], the authors ran an

experiment in which they investigated the frequency of impulse buying-inducing factors in the top 200 e-commerce websites in the US. The findings indicated that each website had a minimum of four features that encouraged impulse buying, while 75% of the websites had at least 16 features that prodded customers toward impulsive purchases. Similarly, [12] showed that out of the 240 apps that were analyzed, 95% contained at least one MD in their interfaces. Collectively, the researchers identified a total of 1,787 Dark Patterns across all the apps, averaging 7.4 harmful designs per app.

It is not unexpected that MDs are widely used because they have been proven in multiple studies to be highly effective in modifying user behavior to the benefit of the site owner. MDs have performed well in various types of testing, such as multivariate tests and A/B testing, as noted by Brignull in 2011 [5]. Using MDs in interface design can increase sales, generate higher revenues, and obtain more user data compared to a design without MD. Luguri and Strahilevitz [31] provided compelling evidence that MDs are effective in influencing consumers' decisions. They conducted a study on the acceptance rate of a security program using three levels of MD: no MD, mild MD, and aggressive MD. The results showed that when no MDs were used, only 11.3% of participants accepted the program, while more than double the participants (25.8%) accepted when mild MD tactics were used. With the use of aggressive MDs, the acceptance rate increased further to 41.9%. Utz et al. [50] investigated the impact of different consent pop-up designs on acceptance rates, including two MDs: Preselection and False Hierarchy. They found that users are more susceptible to sharing personal information when the "accept" button is visually prioritized over the "decline" button. Nouwens et al. [40] examined the effects of different consent banner designs on users' consent choices and found that removing the "reject all" button from the first page of a consent banner and hiding it on a second page, while keeping the "accept all" button present, increased the probability of a user accepting a privacy notice by 22%.

The principles of Accountability and Transparency would mandate that users and customers are informed about the presence of MD in the applications they are using or are about to start using. Unfortunately, this is not the case. In [12], the authors conducted an online study where they asked participants to identify MD elements in the user interfaces (UI) of various applications. The study found that more than half of the participants (55%) did not recognize MDs. This phenomenon was explained by the concept of "Dark Pattern Blindness," where these patterns are so common in today's applications that users have become accustomed to them and no longer notice them easily or at all. To make things worse, research shows that even after learning about manipulative practices, people are still susceptible to the persuasive influence of technology [52].

Because of the prevalence and impact of MD, more work on informing end users (consumers) of online services is necessary. In [8], authors showed that while user experience design students were thoughtful of the importance of user values, they frequently acted in opposition to these values by employing covert and manipulative techniques to influence user behavior and achieve the goals of shareholders. This fact adds to the importance of raising more user awareness around MD, warning users about MD, and even conceivably, considering some regulations to harness the influence of MDs. Fortunately, in the realm of addressing misleading practices in online businesses, governments are increasingly taking regulatory measures to safeguard consumers. Notably, both the European Union (EU) and the United States (US) have consumer protection laws targeting various deceptive tricks resulting in lawsuits against online businesses [14, 17, 24, 48].

Previous studies have proposed solutions to counteract MD. Graßl et al. [19] used nudges to flip the direction of MDs and lead user decisions towards the privacy-friendly choice. Based on a survey among impulse shoppers, Moser et al. [38] suggested friction techniques that neutralize manipulative mechanisms in purchase decisions (e.g., disabling urgency and scarcity messages). Bhoot et al. [32] and Mathur et al. [36] suggested a plug-in or browser extension that automatically detects MDs on websites and notifies the user. Laser [29] discussed the regulatory means that can be leveraged to restrict and fine manipulative tactics. Another conceivable way is the automated recognition of MDs. In a recent study, Mansur et al. [34] used a combination of computer vision and natural language processing to detect cues of ten unique visual and textual MDs in screenshots of applications, allowing for their detection, classification, and locating on the screen. They obtained an overall F1-score of 0.65. Although the results of this work have not had a considerable practical impact, it lays the foundation for further research toward the auto-detection of more types of MDs, with higher accuracy. There is also a book, *"Deceptive patterns – exposing the tricks that tech companies use to control you"* and website [24] by Brignull which aims to inform people of different types of MD and everything they need to know about it.

The currently existing research on counteracting MDs has limitations. First, it remains unclear and unexplored at present what factors influence people to consider such interventions (i.e. whether they seek out such information, whether they consider that information to be effective, and whether they are able to learn from it). Understanding this is needed to develop an evidence base regarding how people may be interacting with information about MD and how likely they are to engage with interventions that may be developed in the future. We address these questions as well in this paper using the PMT as the theoretical framework.

Second, the approaches discussed in the literature are valuable and worth considering, however, there are critical points to notice about them. Auto-recognition of the MDs is a problematic task as the instantiation of a single kind of MD can take many forms. It becomes even harder when we realize that some MDs are in the form of a process across multiple screens (e.g. *Hard to cancel* pattern), not necessarily a static user interface design. Hence, applying natural language processing or image processing tactics does not work at present.

Moreover, even if there were a precise enough tool to detect, locate, and classify MDs, how would the proposed solutions be applied in practice? Using nudges to lead users toward more privacy-friendly choices or using friction methods that neutralize MD effects requires installing an extra browser extension or perhaps an extra mobile application that serves as a filter by users. Although these approaches might be effective, they do not cover many users and would not have an impact on a large scale as they are limited to a browser extension or specific application. More practical approaches are to educate, inform, or warn users at a place that is

highly visited and requires no extra extension or application - e.g. in an app marketplace. This is the approach that we believe holds most promise for practice.

We propose to inform people of MDs inside mobile app distribution platforms (e.g. Google Play Store) by leveraging persuasive strategies. Statistics have shown there were over 6.5 billion smartphone users across the world [13] with downloading around 255 billion mobile apps to their connected devices in 2022 [11]. Moreover, according to statistics in 2020, 83% of time spent with tablets and 90% of time spent with smartphones is in apps [55]. However, browser extensions suggested by previous studies would not be an effective and practical approach, because of their limited user coverage and the fact that they work on the client side and need to be updated constantly to be able to warn their users of new apps (websites) with MDs and new types of MDs. In contrast, app distribution platforms are accessed by billions of users and have the power to inspect and vet apps before allowing them in and to apply rules or laws that demand MDs to be revealed in the app description. Submitting to such control by the distribution platform would ultimately benefit the platform in the long run, since, as [33] suggested, MDs diminish customers' trust in the brand's credibility in the long term.

2.2 PMT and information seeking

The Protection Motivation Theory (PMT) [45] is a theoretical approach that has shown promise in the cybersecurity field. Although originally developed in relation to fear appeals and principally applied to the health domain, PMT focuses on the factors that may influence people's intentions to engage in various behaviors. It revolves around the concept that the choice to react to a potentially harmful situation is shaped by two fundamental processes: threat appraisal and coping appraisal. Threat appraisal involves assessing the likelihood of the event happening (Perceived threat susceptibility) and its potential negative consequences (Perceived threat severity), which are based on an individual's direct experiences and indirect experiences (e.g., information from various sources) [35]. Coping appraisal, on the other hand, relates to the perceived effectiveness of taking protective action (Perceived response efficacy) and individuals' ability to perform an adaptive response (Perceived self-efficacy). Various sources of information, including environmental factors like persuasion and observational learning or personal experiences, can trigger these processes. Having both higher levels of threat appraisal and efficacy appraisal is considered necessary for people to be motivated to protect themselves, therefore, they must both perceive a threat and consider themselves able to effectively manage that threat. Rogers [46] further highlights the possible role of costs within this approach, whereby perceived response costs may influence overall coping appraisal (i.e. the perceived costs of engaging in a protective behavior). The lower the perceived cost of performing a protection task the more likely individuals are to engage in it. In other words, how 'severe' or 'likely' people perceive the MD threat to be, as well as the perceived effectiveness of protective information ('response efficacy') and the perceived ability to access such information (termed 'self-efficacy') may all influence their resultant engagement with protective information and other intervention materials. To our best knowledge, this theoretical framework has not been utilized to explore the factors that drive individuals to seek protective information about MDs. Figure 1 shows the PMT model we assess in this study.

3 RESEARCH DESIGN

A mixed study design was used, meaning the study was within-subject and between-subject in two different stages of the procedure. Before the actual study, we conducted a pilot study to refine the experiment design and survey questionnaire, evaluate the duration, and enhance the overall study design. All survey questions in the study were administered via a University-licensed SurveyMonkey site. After giving consent to participate in the study, the users were asked questions about their demographic background and shown a text briefly introducing manipulative designs as follows:

"Manipulative designs, also known as "dark patterns," are deceptive features in websites and apps that can trick or pressure you into making choices you might not want to. Here are some examples:

***Fake Scarcity**: You see a message like "Only 3 left in stock!" on a shopping site. This might not be true, but it makes you want to buy quickly.*

***Hard to Cancel**: Ever tried to cancel a subscription and found it really hard? That's by design, to keep you paying longer.*

***Hidden Privacy Choices**: Some apps or websites make it really easy to say "yes" to sharing your data but hide the "no" option. They might not be clear about what they'll do with your data.*

These are just 3 examples, but there are about 15 different types of these tricks, and they're common in online shopping, games, and social media apps. They can influence you to make choices that aren't in your best interest, like spending more money or giving away more personal information than you intended."

Then, participants were asked to fill out a survey containing items for intention and the PMT constructs' items. The survey items are shown in Table 1. They are adapted from a validated questionnaire developed by [53] regarding information seeking about phishing techniques. To keep the validity of the survey, we made only minimal changes to the items of the original survey by superseding the phrase *"falling victim to"* with "being affected by," replacing the word *"phishing"* with *"manipulative design"*, and *"Keeping up to date"* with *"learning about"*. PMT constructs' items were shown to the participants in a random order; however, the items related to participants' intention about information seeking about manipulative design were provided at the end of the questionnaire after the PMT items. We included an attention-check statement to filter out responses from participants who did not pay close attention while reading and answering the statements, in line with the findings of [28] that "attention-check" questions do not compromise scale validity. The attention-check was a statement similar to the other PMT items (*"I intend to learn more about manipulative designs in the long run, but this question is just to evaluate your attention. Please mark the lowermost option to let us know that you are paying attention"*).

After they responded to the questionnaire, the participants were provided with a link to the simulated app presentation (SAP) as it would appear in some app distribution platforms, e.g. Google Play Store. Note that we did not simulate the entire Google Play Store; we only simulated the representation page of a single shopping

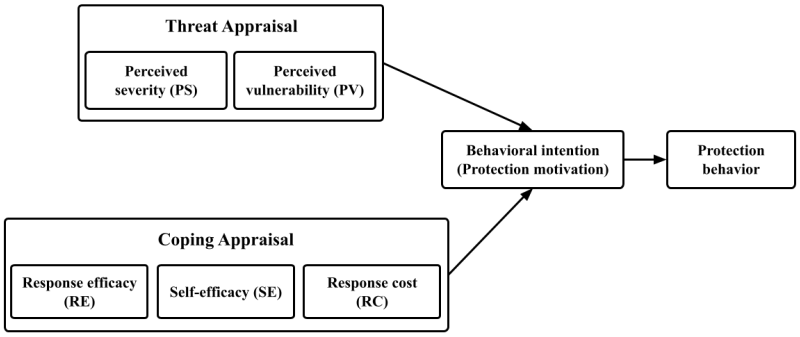

Figure 1: Protection Motivation Theory constructs

Table 1: PMT and intention items

Construct	Item
Perceived severity	If I were to be affected by a manipulative design, the consequences could be severe.
	Losing data privacy because of a manipulative design would be a serious problem for me.
	Being affected as a result of not detecting a manipulative design would be a serious problem for me.
Perceived vulnerability	It is possible that I will be affected by a manipulative design.
	I feel that I could be vulnerable to manipulative designs.
	It is likely that I will be affected by a manipulative design.
Self-efficacy	It would be easy for me to learn about manipulative design techniques.
	I feel confident in my ability to learn about manipulative design techniques.
	I am able to learn about manipulative design techniques.
Perceived response efficacy	If I learn about manipulative design techniques, I am less likely to be affected by them.
	If I learn about manipulative designs, I will lessen my chances of being affected by their trick.
	Learning about manipulative design techniques will prevent me from being affected by them.
Perceived response cost	Learning about manipulative design techniques takes a large amount of time.
	Trying to learn about manipulative design techniques would cause me many problems.
	Learning about manipulative design techniques requires significant effort.
Thinking of your future actions, indicate the degree to which you agree or disagree with the following statements regarding your likelihood of learning about manipulative design techniques in the future.	
Intention	I intend to learn about manipulative design techniques in the next 3 months.
	I am likely to learn about manipulative design techniques in the next 3 months.
	I expect to continue learning about manipulative design techniques in the future.

application. The SAP contains an MD warning - a red circle with a white "!" image and text "Manipulative Design" - along with the standard information about the app provided by the Google Play Store (number of reviews, downloads, and age group), as shown in Figure 2. This is because previous research showed that warnings containing a color or an icon are more effective in getting peoples' attention than warnings without such elements [51, 56].

We made three designs for SAP: one baseline design and two experimental design versions employing the authority and social influence persuasive strategies, respectively. Participants were randomly assigned to one of three design groups in the order they entered the study. In all three designs, to create a realistic list of user reviews, we added entirely positive reviews along with reviews reporting an MD.

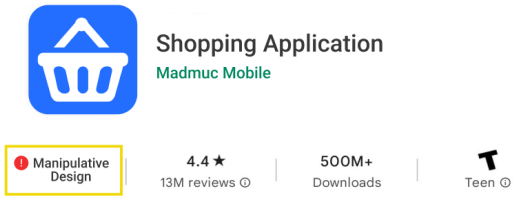

Figure 2: Warning notice at the top of the page

The *baseline* version simulates how apps are currently presented on app distribution platforms. There is no explicit section warning about any MDs used inside the SAP. Among the user reviews for the shopping app, there were two that described the experience

of other users with the MDs inside the shopping application ; the user may notice or overlook them. It is worth noting, however, that it is hard to design a totally neutral control condition inside an app distribution platform because users may presume that the fact that the app is available on the platform means that some authority has vetted the app as acceptable. User reviews provide additional information that can influence the users' decisions, acting as implicit social influence persuasion.

In the *authority* and *social influence* versions, there is a separate section (MD-Section) in the middle of the page, specifically to create MD awareness. A persuasive message is added to the beginning of this section, followed by a button labelled "Learn More," which leads to two examples of manipulative designs if the user clicks on it. The MD-Sections for the SAP's authority and social influence version are shown respectively in Figure 3 and 4.

In the brief introduction at the beginning of the experiment, participants were asked to decide whether to install the SAP or ignore it based on the information presented in the application presentation. After completing the experiment by clicking the install or ignore button on the SAP page and confirming that they had read the information, participants were instructed to return to the SurveyMonkey questionnaire. They were asked to respond to the same questionnaire containing the PMT and intention questions they had responded to before the experiment. The objective was to assess the impact of each warning (in the experimental versions) or lack thereof (in the baseline version) on the PMT constructs and intention for learning about MDs.

For the data analysis, we applied the Partial Least Square Structural Equation Modeling (PLS-SEM) [23] technique using Smart PLS-4 and the paired samples Wilcoxon test (also known as Wilcoxon signed-rank test) using SPSS respectively to answer the research questions of the study. The following section delves into the analysis of survey data.

4 DATA ANALYSIS AND RESULTS

The study was approved by the University's Behavioral Research Ethics Board (Beh-REB) under certificate Beh-ID 4304. All participants in the study were recruited through the University of Saskatchewan announcement board, newsletters, and social media posts in November and December 2023.

A total of 191 participants initially completed the study. However, individuals who did not respond correctly to the two attention-check questions in the pre- and post-experiment surveys were subsequently excluded, resulting in a reduced dataset with 134 participants. SPSS 27.0.1 was used for descriptive statistics analysis by obtaining the frequency of socio-demographic variables, such as age, gender, education, and familiarity with the MD concept. The study took, on average, 15 minutes to complete. 23.0% of the participants were between 18-24, 44.4% between 25-34, 19.3% between 35-44, 8.9% between 45-54, 2.2% older than 55, and 2.2% younger than 18 years old; 85.9% had an undergraduate and graduate degree, 8.1% had a high school degree, and 5.9 reported their level of education as "other." This suggests that the sample skews towards a particular demographic, with predominantly younger participants holding higher education degrees.

Moreover, 30.4% of the participants reported they were more than "somewhat familiar" with the MD concept, 47.4% of the participants were "somewhat familiar," and 22.2% did not know about the manipulative design concept, confirming the findings of [3, 20, 30] about the public's slight awareness of MD.

4.1 PLS-SEM for exploring determinants for Behavior Intention to learn about MD

Based on the first questionnaire, we utilized variance-based structural equation modeling (SEM), a widely employed approach to find the relationships between PMT constructs and intention to learn about MDs. SEM encompasses two primary techniques: covariance-based SEM (CB-SEM) and variance-based SEM. The latter uses the Partial Least Squares Method (PLS), which, according to [44], has some advantages over CB-SEM, including the ability to accept small samples and non-normally distributed data, which is more suitable for the validation of complex models. Consequently, because of having rather small data (n=134), we used PLS-SEM for our analysis.

We adhered to the two-step procedure proposed by Anderson and Gerbing [1]. The first step is outer model (measurement model) analysis, including reliability and validity testing, while the second step is inner model (structural model) analysis, including estimating and validating the structural model's path coefficients. The first step is intended to verify whether the constructs are reliable and valid, and the second step is to validate the relationships between constructs.

4.1.1 Reliability and Validity. As is shown in Table 2a, reliability was evaluated using the composite reliability (CR) and Cronbach's alpha. The CR and Cronbach's alpha values were all higher than 0.7, indicating a desirable reliability. The standardized factor loading and T-value of each item are shown in Figure 5 on each link between latent variables and their indicators. The validity test included a convergent validity test and a discriminant validity test. The convergent validity test was used by measuring the average variance extracted (AVE) and CR of each construct to determine the degree of similarity of different measures of the same construct. Along with the CRs higher than 0.8, the AVEs were all higher than 0.5, suggesting a desirable convergent validity [16]. The discriminant validity test was employed to check whether a very high correlation existed between the latent constructs. The heterotrait–monotrait (HTMT) ratio of correlations proposed by Henseler et al. was used to evaluate discriminant validity. As shown in Table 2b, values of HTMT were all lower than 0.9, demonstrating a favorable discriminant validity [25].

4.1.2 Structural model. Bootstrapping was operated to estimate the significance (α = 0.05) of each path coefficient. Path coefficients (β) and significance (p-value) were used to define the intensity and direction of the variable relationships to reveal the correlation between PMT constructs and BI. Additionally, explanatory power was assessed by the R^2 value. The results are shown in Figure 5. The analysis shows that **perceived severity (PS) positively and significantly affects behavioral intention** (β = 0.386, p-value < 0.001), and the **perceived response cost (RC) negatively and significantly impacts behavioral intention** (β = -0.243, p-value

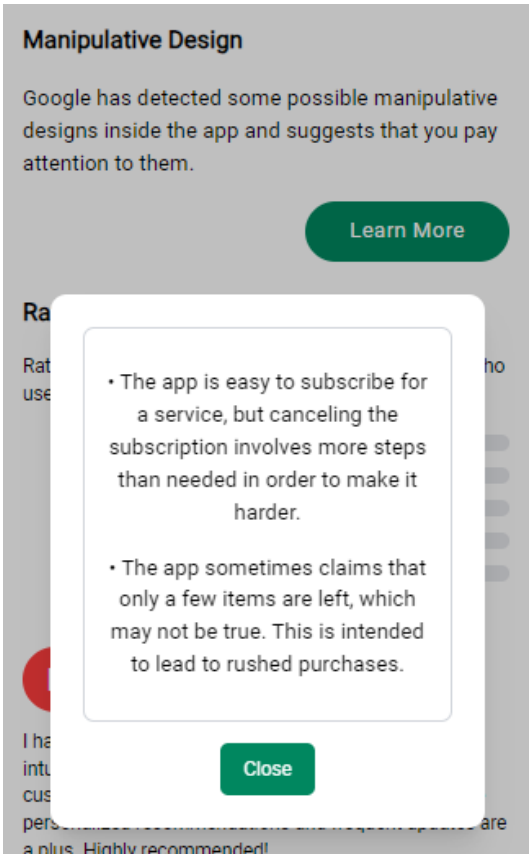

Figure 3: MD awareness in authority version

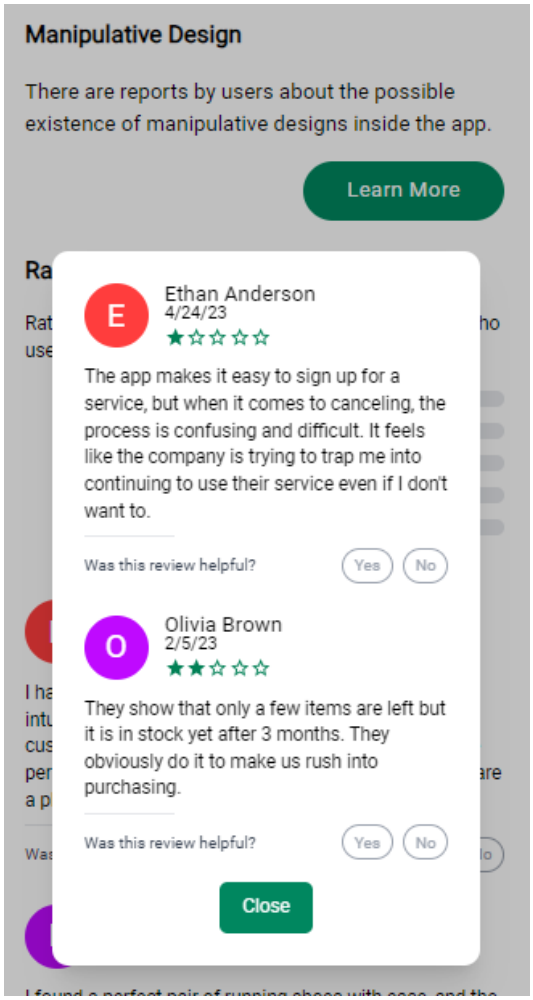

Figure 4: MD awareness in social influence version

Table 2: Reliability and validity tests

(a) Reliability and convergent validity

Construct	Cronbach's α	Composite reliability	AVE
PS	0.704	0.834	0.628
PV	0.868	0.916	0.789
RE	0.913	0.945	0.853
SE	0.849	0.911	0.775
RC	0.821	0.890	0.731
BI	0.790	0.878	0.706

(b) Discriminant validity - Heterotrait-monotrait (HTMT)

	BI	PS	PV	RC	RE	SE
BI						
PS	0.569					
PV	0.205	0.328				
RC	0.298	0.079	0.097			
RE	0.192	0.430	0.143	0.141		
SE	0.189	0.147	0.130	0.333	0.055	

< 0.05). This result provides an answer to our first research question **RQ1**.

The perceived vulnerability, perceived response efficacy, and self-efficacy had positive, but no significant influence on the intention (p-value > 0.05). It is also necessary to determine the variance indicated by the R^2 value in the structural model analysis. The R^2 value of behavioral intention is 0.260, indicating that 26.0% of the variance of behavioral intention is explained by the PMT constructs, which is a totally acceptable explanatory power [43]. This is because it is not uncommon for research centered on human behavior to yield lower R^2 values than other domains (given the inherent

complexities in predicting behavior). Nevertheless, additional research investigating this relationship would be helpful for drawing more robust conclusions.

4.2 Paired samples Wilcoxon test for analyzing whether persuasive strategies can amplify BI

The results of the questionnaires before (pre) and after (post) participants worked with the SAP were used to calculate the change each persuasive strategy made in the PMT constructs and BI. We ran a Paired samples Wilcoxon test [54] for each construct pair (e.g. pre-PS and post-PS) in each design group (baseline, authority, and social influence) and compared the results. In selecting the Wilcoxon signed-rank test as a non-parametric alternative to the paired-sample t-test for analyzing Likert scale data in this study, several considerations were taken into account. Likert scale responses, being ordinal, often deviate from normal distribution. Likewise, in our case, the assumption of normality was violated in our dataset for most of the variables in all groups (Shapiro-Wilk p-value(s) < 0.05); therefore, we adopted the Wilcoxon test. The results of the test are shown in Table 3. The negative mean difference value indicates an increase in the construct mean value from pre-intervention to post-intervention. Hence, a negative value is favorable for all constructs except for the response cost, in which a positive mean difference indicates a decrease, which is preferable in response cost. As the results suggest, the baseline and two experimental interventions all caused favorable changes to the constructs, although the changes were not significant for all cases. **Authority and social influence persuasive strategies outperformed the baseline version for all constructs except self-efficacy**. This result provides an answer to the second research question, **RQ2**.

In the baseline version, there was no significant change in any constructs. In the authority group, only the change in behavioral intention was significant (p-value = 0.004), and there was no significant change in the PMT constructs. However, in the social influence group, perceived severity (p-value = 0.008) and behavioral intention (p-value < .001) experienced significant changes before and after the intervention. The response cost (p-value = 0.053) was also close to experiencing a significant change. Accordingly, **the persuasive strategy social influence outperforms the authority strategy** in the task of influencing people's intention to seek information about manipulative designs.

5 DISCUSSION

The first part of this study focused on examining the relation of PMT-based measures to self-reported intentions as determinants of protective behavior to access information about manipulative designs. Although previous work has explored PMT in relation to a range of security behaviors [4], this study is the first to examine this approach in the context of manipulative design information seeking. Understanding the motivations that drive individuals to seek information about manipulative designs (dark patterns) is neglected in this domain, although this often provides a fundamental means of informing users about how they are being manipulated to serve someone else's interests. The overall goal of our approach is to create transparency and accountability regarding the presence of MDs in apps on distribution platforms. Our approach will not necessarily make users avoid systems with manipulative design because often there are no better choices, or on balance, the positive features of an app dominate over the possible harms of MDs. The intention behind this study is to discover how to persuade users to pay attention and seek information about MDs by finding underlying determinants of intention to learn about manipulative designs and employing possible persuasive strategies to notify users of MD. In this way, users could be aware and make informed decisions.

Altogether, the results of this investigation demonstrate some consistency with findings from other risk-related areas with the support shown for the role of perceived severity as a threat appraisal and perceived response cost as coping appraisal in influencing future intentions to take protective behavior [7, 39]. However, in line with the results of [53] focusing on phishing, we did not find firm support for the role of perceived vulnerability within the MD context. It may be that those who consider themselves vulnerable to manipulative designs will not be motivated to seek more information about them unless they perceive the consequences as severe (perceived severity). Unlike the findings of [53] in the phishing context, however, we did not find rigorous evidence that self-efficacy and perceived response efficacy have an effect on information seeking in the MD context. An interpretation of this fact might be that since many people have a vague perception of manipulative designs and feel the consequences are not as extreme as phishing threats to themselves in many MD cases, they have no desire to learn more about them even though they believe they can easily access such information. In addition, it is possible that many people have encountered some types of MDs [3, 20, 30] such as those highlighted in our study, and these MDs have not necessarily had a negative impact on them individually (lack of perceived severity), and as a result, they may regard manipulative designs not sufficiently crucial to need learning about them even though they think that learning about them is helpful (response efficacy).

The study answered the second research question (RQ2), showing that using persuasive strategies increases the intention to seek information about manipulative designs. Our findings align with previous studies in Human-Computer Interaction (HCI) that demonstrate the effectiveness of persuasive strategies, particularly those leveraging social influence tactics. Notably, research by Fogg [15] and Cialdini [9] has highlighted the power of social influence in shaping user behavior within interactive systems. The experimental groups (either authority or social influence) showed more favorable changes in self-reported intention and PMT constructs (except for self-efficacy) compared to the baseline, although not all of the changes in PMT constructs were significant. The social influence strategy made not only more positive changes to self-reported intention but also induced more favorable changes to perceived severity and perceived response cost, which are the most significant drivers of seeking information about MD yielded from the previous section. One can argue that in addition to the two strategies, there is a third one - "fear" - that exists implicitly in both the persuasive and the baseline warning messages, but we do not plan to evaluate its effect separately in our study because any warning or explanation of MD implicitly suggests the possible risks and harms for users, and would evoke fear.

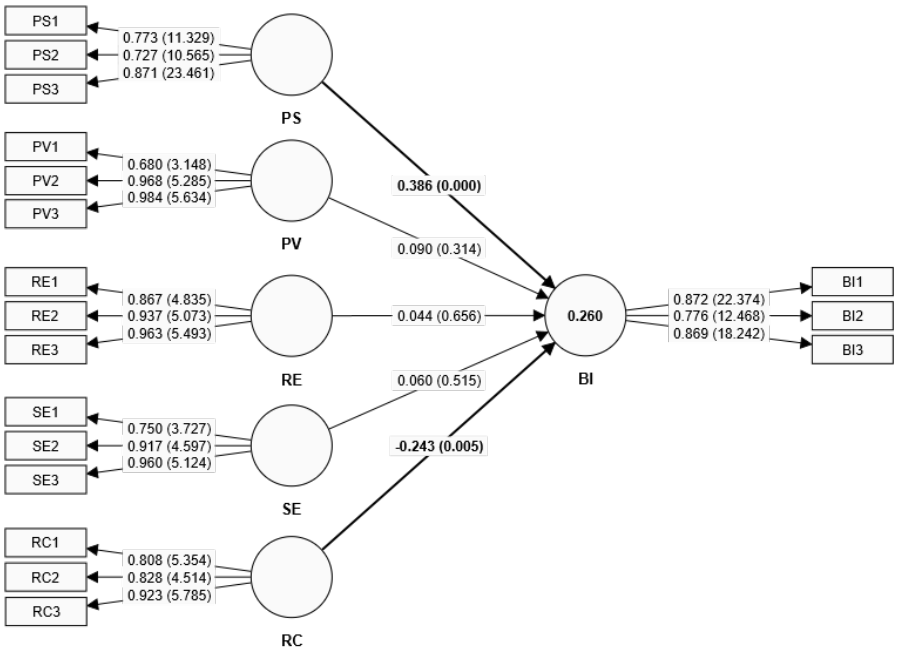

Figure 5: The PLS-SEM of the PMT Scale

Table 3: Mean differences and (p-values) in PMT and BI constructs between pre and post-intervention (ADP) in the three groups

		Baseline (n=41)	Authority (n=49)	Social influence (n=45)
PS-pre	PS-post	-0.333 (0.240)	-0.167 (0.404)	-0.333 **(0.008)**
PV-pre	PV-post	-3.00e-5 (0.739)	-0.167 (0.120)	-0.167 (0.318)
RE-pre	RE-post	-0.167 (0.217)	8.98e-6 (0.863)	-0.167 (0.160)
SE-pre	SE-post	-0167 (0.084)	-3.33e-5 (0.992)	-0.167 (0.542)
RC-pre	RC-post	0.500 (0.122)	0.333 (0.078)	0.333 (0.053)
BI-pre	BI-post	-0.333 (0.122)	-0.333 **(0.004)**	-0.333 **(<.001)**

5.1 Implications for interventions

The findings of the first part of the study suggest that in designing MD awareness approaches, it is helpful to emphasize the severe potential risks that MDs cause for individuals by bringing the most harmful MD types to the users' attention in order to signpost them to further information. However, it is also necessary to provide MD information to the user in a straightforward and concise way and to minimize the cost of learning about MD, e.g. *"It only takes 5 minutes to learn about manipulative design."*

The study showed the effectiveness of persuasive strategies in increasing self-reported intention to learn about MD. Therefore, it is recommended to employ the social influence strategy to positively influence people's intention to seek information about manipulative designs as much as possible. A practical use of this strategy (as utilized in this study) could be to present other authentic users' reports about their encounters with MD in the form of reviews and noting this by phrases such as *"There are reviews suggesting the presence of manipulative designs in this app"*.

An important implication of our work is suggesting the app presentations on app distribution platforms as a good medium to build awareness about manipulative design. We demonstrated a possible design to inform users about MD in a simulated app presentation for the Google Play Store. App distribution platforms (ADPs) are accessed by billions of users and have the power to apply rules to reveal MDs in the app description pages in the form of user reviews. ADPs, visited by many users, have the potential to significantly raise awareness about MDs. As suggested by [49], increased awareness is crucial to help unsuspecting users make decisions that are in their best interest. Another benefit of including MD information within ADPs is that users encounter this information when installing or learning more about an app. This contextual presentation enhances the effectiveness of this approach. ADPs can deploy text analytic tools over these reviews to flag apps that use MDs and incorporate warnings in their presentations. This will ensure transparency and accountability for ADPs and increase customer trust. Therefore we believe that ADPs provide a feasible opportunity to inform users about MD on a large scale. Submitting to such control by the distribution platform would ultimately benefit also the applications offered on the platform in the long run, since, as [33] suggested,

MDs diminish customers' trust in the brand's credibility in the long term.

5.2 Limitations and Future work

Our study did not educate the participants about the different types of MD. It only contained a short description of what an MD is in the introduction to inform the participants about the meaning of the terms used in the questionnaire. Adding an educational section and elaborating on MD to familiarize users with the different types of MD would have biased the results of the study. How to educate users about MD is an important challenge left for future studies because information seeking in itself may not always be beneficial but is instead entirely dependent on accessing appropriate, valuable, and up-to-date information and guidance.

Furthermore, it is important to acknowledge that the study results were derived from a relatively limited sample size of 135 participants. The participants were predominantly young and educated individuals. This demographic composition raises concerns regarding the generalizability of the findings, as the sample may not fully represent the diversity of perspectives and experiences present in the broader population. Consequently, there is a risk of bias in the study results, as they may not accurately reflect the attitudes, behaviors, and responses of individuals from different age groups, educational backgrounds, and socio-economic statuses.

Finally, only self-reported intentions to learn about MD techniques were assessed rather than actual information-seeking behavior, so future work is needed to explore the extent to which our survey constructs also relate to future actions. Since past behavior may provide a useful measure in this context, including an assessment of past behavior related to this aspect would also be beneficial in future work.

6 CONCLUSION

Overall, this study provides a first step in uncovering factors that affect whether people intend to learn about manipulative designs (MDs). We applied the Protection Motivation Theory, which is a theoretical construct used in other risk-related domains. The findings reveal that the perceived severity of manipulative design and the perceived response cost of learning information about that can directly influence future intentions to engage with that information. We also evaluated the efficacy of two persuasive strategies, authority, and social influence strategy, showing that the latter is more effective in warning users about the presence of MDs in apps on app distribution platforms. The designs presented can be used to provide accountability and transparency about manipulative designs.

ACKNOWLEDGMENTS

This work was partially supported by the Natural Sciences and Engineering Research Council (NSERC) Discovery Grant Program grant of the second author (RGPIN-2021-03521).

REFERENCES

[1] James C Anderson and David W Gerbing. 1988. Structural equation modeling in practice: A review and recommended two-step approach. *Psychological bulletin* 103, 3 (1988), 411.

[2] Jonathan Ben-Naim, Jean-François Bonnefon, Andreas Herzig, Sylvie Leblois, and Emiliano Lorini. 2017. Computer-mediated trust in self-interested expert recommendations. *Cognition beyond the brain: Computation, interactivity and human artifice* (2017), 233–250.

[3] Kerstin Bongard-Blanchy, Arianna Rossi, Salvador Rivas, Sophie Doublet, Vincent Koenig, and Gabriele Lenzini. 2021. " I am Definitely Manipulated, Even When I am Aware of it. It's Ridiculous!"-Dark Patterns from the End-User Perspective. In *Designing Interactive Systems Conference 2021*. 763–776.

[4] Scott R Boss, Dennis F Galletta, Paul Benjamin Lowry, Gregory D Moody, and Peter Polak. 2015. What do systems users have to fear? Using fear appeals to engender threats and fear that motivate protective security behaviors. *MIS quarterly* 39, 4 (2015), 837–864.

[5] Harry Brignull. 2011. Dark patterns: Deception vs. honesty in UI design. *Interaction Design, Usability* 338 (2011), 2–4.

[6] Ting-Ray Chang, Eija Kaasinen, and Kirsikka Kaipainen. 2013. Persuasive design in mobile applications for mental well-being: multidisciplinary expert review. In *Wireless Mobile Communication and Healthcare: Third International Conference, MobiHealth 2012, Paris, France, November 21-23, 2012, Revised Selected Papers 3*. Springer, 154–162.

[7] Tim Chenoweth, Robert Minch, and Tom Gattiker. 2009. Application of protection motivation theory to adoption of protective technologies. In *2009 42nd Hawaii International Conference on System Sciences*. IEEE, 1–10.

[8] Shruthi Sai Chivukula, Jason Brier, and Colin M Gray. 2018. Dark Intentions or Persuasion? UX Designers' Activation of Stakeholder and User Values. In *Proceedings of the 2018 ACM Conference Companion Publication on Designing Interactive Systems*. 87–91.

[9] Robert B Cialdini. 2001. The science of persuasion. *Scientific American* 284, 2 (2001), 76–81.

[10] Robert B Cialdini and Noah J Goldstein. 2002. The science and practice of persuasion. *Cornell Hotel and Restaurant Administration Quarterly* 43, 2 (2002), 40–50.

[11] Data.ai. 2023. Number of mobile app downloads worldwide from 2016 to 2022. https://www.statista.com/statistics/271644/worldwide-free-and-paid-mobile-app-store-downloads Accessed: 2023-04-25.

[12] Linda Di Geronimo, Larissa Braz, Enrico Fregnan, Fabio Palomba, and Alberto Bacchelli. 2020. UI dark patterns and where to find them: a study on mobile applications and user perception. In *Proceedings of the 2020 CHI conference on human factors in computing systems*. 1–14.

[13] Ericsson. 2022. Number of smartphone mobile network subscriptions worldwide from 2016 to 2022, with forecasts from 2023 to 2028. https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide Accessed: 2023-04-25.

[14] Federal Trade Commission. 2023. *Complaint of the Federal Trade Commission v. Epic Games, Inc.* https://www.ftc.gov/system/files/ftc_gov/pdf/1923203epicgamesfinalconsent.pdf Accessed: November 03, 2023.

[15] Brian J Fogg. 2002. Persuasive technology: using computers to change what we think and do. *Ubiquity* 2002, December (2002), 2.

[16] Claes Fornell and David F Larcker. 1981. Evaluating structural equation models with unobservable variables and measurement error. *Journal of marketing research* 18, 1 (1981), 39–50.

[17] French Data Protection. 2022. *Complaint of the French Data Protection Authority (DPA) v. TikTok.* https://www.legifrance.gouv.fr/cnil/id/CNILTEXT000046977994?page=1&pageSize=10&query=2016%252F679&searchField=ALL&searchType=ALL&sortValue=DATE_DECISION_DESC&tab_selection=cnil&typePagination=DEFAULT Accessed: November 03, 2023.

[18] Anirudh Ganesh, Chinenye Ndulue, and Rita Orji. 2021. PERMARUN-a persuasive game to improve user awareness and self-efficacy towards secure smartphone behaviour. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–7.

[19] PAJ Graßl, HK Schraffenberger, FJ Zuiderveen Borgesius, and MA Buijzen. 2021. Dark and bright patterns in cookie consent requests. (2021).

[20] Colin M Gray, Jingle Chen, Shruthi Sai Chivukula, and Liyang Qu. 2021. End user accounts of dark patterns as felt manipulation. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021), 1–25.

[21] Colin M Gray, Yubo Kou, Bryan Battles, Joseph Hoggatt, and Austin L Toombs. 2018. The dark (patterns) side of UX design. In *Proceedings of the 2018 CHI conference on human factors in computing systems*. 1–14.

[22] Saul Greenberg, Sebastian Boring, Jo Vermeulen, and Jakub Dostal. 2014. Dark patterns in proxemic interactions: a critical perspective. In *Proceedings of the 2014 conference on Designing interactive systems*. 523–532.

[23] Joe F Hair, Christian M Ringle, and Marko Sarstedt. 2011. PLS-SEM: Indeed a silver bullet. *Journal of Marketing theory and Practice* 19, 2 (2011), 139–152.

[24] Brignull Harry. 2021. Deceptive Design. https://deceptive.design Accessed: 2023-04-03.

[25] Jörg Henseler, Christian M Ringle, and Marko Sarstedt. 2015. A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the academy of marketing science* 43 (2015), 115–135.

[26] Maurits Kaptein, Panos Markopoulos, Boris de Ruyter, and Emile Aarts. 2009. Can you be persuaded? individual differences in susceptibility to persuasion. In *Human-Computer Interaction–INTERACT 2009: 12th IFIP TC 13 International Conference, Uppsala, Sweden, August 24-28, 2009, Proceedings, Part I 12*. Springer, 115–118.

[27] Kathy Kellermann and Tim Cole. 2006. Classifying Compliance Gaining Messages: Taxonomic Disorder and Strategic Confusion. *Communication Theory* 4, 1 (03 2006), 3–60. https://doi.org/10.1111/j.1468-2885.1994.tb00081.x arXiv:https://academic.oup.com/ct/article-pdf/4/1/3/22294641/jcomthe0003.pdf

[28] Franki YH Kung, Navio Kwok, and Douglas J Brown. 2018. Are attention check questions a threat to scale validity? *Applied Psychology* 67, 2 (2018), 264–283.

[29] MR Leiser. 2020. 'Dark Patterns': the case for regulatory pluralism. *Available at SSRN 3625637* (2020).

[30] Yuwen Lu, Chao Zhang, Yuewen Yang, Yaxing Yao, and Toby Jia-Jun Li. 2023. From Awareness to Action: Exploring End-User Empowerment Interventions for Dark Patterns in UX. *arXiv preprint arXiv:2310.17846* (2023).

[31] Jamie Luguri and Lior Jacob Strahilevitz. 2021. Shining a light on dark patterns. *Journal of Legal Analysis* 13, 1 (2021), 43–109.

[32] Aditi M. Bhoot, Mayuri A. Shinde, and Wricha P. Mishra. 2020. Towards the identification of dark patterns: An analysis based on end-user reactions. In *IndiaHCI'20: Proceedings of the 11th Indian Conference on Human-Computer Interaction*. 24–33.

[33] Maximilian Maier and Rikard Harr. 2020. Dark design patterns: An end-user perspective. *Human Technology* 16, 2 (2020), 170.

[34] SM Hasan Mansur, Sabiha Salma, Damilola Awofisayo, and Kevin Moran. 2023. Aidui: Toward automated recognition of dark patterns in user interfaces. In *2023 IEEE/ACM 45th International Conference on Software Engineering (ICSE)*. IEEE, 1958–1970.

[35] Ingrid M Martin, Holly Bender, and Carol Raish. 2007. What motivates individuals to protect themselves from risks: the case of wildland fires. *Risk Analysis: An International Journal* 27, 4 (2007), 887–900.

[36] Arunesh Mathur, Gunes Acar, Michael J Friedman, Eli Lucherini, Jonathan Mayer, Marshini Chetty, and Arvind Narayanan. 2019. Dark patterns at scale: Findings from a crawl of 11K shopping websites. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–32.

[37] David Modic and Ross Anderson. 2014. Reading this may harm your computer: The psychology of malware warnings. *Computers in Human Behavior* 41 (2014), 71–79.

[38] Carol Moser, Sarita Y Schoenebeck, and Paul Resnick. 2019. Impulse buying: Design practices and consumer needs. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–15.

[39] Kurt Neuwirth, Sharon Dunwoody, and Robert J Griffin. 2000. Protection motivation and risk communication. *Risk Analysis* 20, 5 (2000), 721–734.

[40] Midas Nouwens, Ilaria Liccardi, Michael Veale, David Karger, and Lalana Kagal. 2020. Dark patterns after the GDPR: Scraping consent pop-ups and demonstrating their influence. In *Proceedings of the 2020 CHI conference on human factors in computing systems*. 1–13.

[41] Abdul-Hammid Olagunju, Marcella Ogenchuk, and Julita Vassileva. 2021. Mobile Persuasive Application for Responsible Alcohol Use: Drivers for Use and Impact of Social Influence Strategies. In *International Conference on Persuasive Technology*. Springer, 102–114.

[42] José Osvaldo De Sordi, Manuel Meireles, and Marcia Carvalho de Azevedo. 2014. Information selection by managers: priorities and values attributed to the dimensions of information. *Online Information Review* 38, 5 (2014), 661–679.

[43] Peterson K Ozili. 2023. The acceptable R-square in empirical modelling for social science research. In *Social research methodology and publishing results: A guide to non-native english speakers*. IGI Global, 134–143.

[44] Christian M Ringle, Marko Sarstedt, and Detmar W Straub. 2012. Editor's comments: a critical look at the use of PLS-SEM in" MIS Quarterly". *MIS quarterly* (2012), iii–xiv.

[45] Ronald W Rogers. 1975. A protection motivation theory of fear appeals and attitude change1. *The journal of psychology* 91, 1 (1975), 93–114.

[46] Ronald W Rogers. 1983. Cognitive and physiological processes in fear appeals and attitude change: A revised theory of protection motivation. *Social psychology: A source book* (1983), 153–176.

[47] Jingzhi Shang, Debra Z Basil, and Walter Wymer. 2010. Using social marketing to enhance hotel reuse programs. *Journal of Business Research* 63, 2 (2010), 166–172.

[48] year = 2022 note = Accessed: November 04, 2023 SUPERIOR COURT OF THE DISTRICT OF COLUMBIACIVIL DIVISION, url = https://oag.dc.gov/sites/default/files/2022-12/2022.12.30%20Consent%20Judgment%20and%20Order.pdf. [n. d.]. *Complaint of DISTRICT OF COLUMBIA v. GrubHub, Inc.*

[49] Kiemute Oyibo Tasneem Naheyan. 2024. The Effect of Dark Patterns and User Knowledge on User Experience, Decision Making and Website Reputation. In *International Conference on Persuasive Technology*. Springer, 36–50.

[50] Christine Utz, Martin Degeling, Sascha Fahl, Florian Schaub, and Thorsten Holz. 2019. (Un) informed consent: Studying GDPR consent notices in the field. In *Proceedings of the 2019 acm sigsac conference on computer and communications security*. 973–990.

[51] Jie Wang, Jiaming Shi, Xin Wen, Liang Xu, Ke Zhao, Fuyang Tao, Wenbiao Zhao, and Xiuying Qian. 2022. The effect of signal icon and persuasion strategy on warning design in online fraud. *Computers & Security* 121 (2022), 102839.

[52] Susan Weinschenk. 2013. *How to get people to do stuff: Master the art and science of persuasion and motivation*. New Riders.

[53] Emma J Williams and Adam N Joinson. 2020. Developing a measure of information seeking about phishing. *Journal of Cybersecurity* 6, 1 (2020), tyaa001.

[54] Robert F Woolson. 2007. Wilcoxon signed-rank test. *Wiley encyclopedia of clinical trials* (2007), 1–3.

[55] Yoram Wurmser. 2020. The Majority of Americans' Mobile Time Spent Takes Place in Apps. https://www.insiderintelligence.com/content/the-majority-of-americans-mobile-time-spent-takes-place-in-apps Accessed: 2023-04-25.

[56] Stephen L Young. 1991. Increasing the noticeability of warnings: Effects of pictorial, color, signal icon and border. In *Proceedings of the Human Factors Society Annual Meeting*, Vol. 35. Sage Publications Sage CA: Los Angeles, CA, 580–584.

[57] Tao Zhou and Yanjun Sun. 2009. An empirical analysis of online consumer initial trust building based on elm. In *2009 International Conference on Networks Security, Wireless Communications and Trusted Computing*, Vol. 2. IEEE, 59–62.