Measuring the Deceptive Potential of Design Patterns: A Decision-Making Game

Deborah Maria Löschner¹, Sebastian Pannasch²

¹ Technical University of Dresden, 01069 Dresden, Germany

Abstract
Recent years have seen exponential growth in research on deceptive design patterns (DDPs), revealing a clear impact on user behavior. The interdisciplinary research work has identified potential harms caused by DDPs, including financial, data, and attention losses, heightened frustration levels, and increased cognitive load. Since existing studies often employ realistic scenarios, they face various methodological challenges such as (i) limited statistical and explanatory power regarding the underlying mechanisms of deception, (ii) a lack of control over contextual factors, and (iii) difficulties in adapting to new developments in a rapidly transforming field as digital design. To address these challenges, we advocate for methodological innovations and introduce a decision-making game as a new experimental paradigm. The game incorporates various DDPs and contextual factors, allowing systematic exploration of their effects on decision-making processes. The paradigm aims to measure behavioral outcomes as well as underlying cognitive processes, providing a more nuanced understanding of DDP influence. By proposing a framework to build a reliable experimental approach, this work contributes to advancing the study of the influence of DDPs on user behavior and the understanding of potential countermeasures. The proposed paradigm offers flexibility, comparability between different user groups, and adaptability, providing a foundation for future investigations into the socio-digital vulnerability of users and the development of effective countermeasures.

Keywords
deceptive design, dark patterns, manipulation, experiment, decision-making

1. Introduction
As the internet becomes more and more a digital marketplace for various goods, users are increasingly challenged to autonomously deploy their money, data, or attention [1]. One of the contributing factors is the presence of deceptive design patterns (DDPs), which are design structures intended to influence user behavior in favor of companies’ interests. In recent years, literature on design practices has grown exponentially. Researchers from various disciplines have collaboratively contributed to a rich body of work that categorizes and empirically substantiates potential harms caused by various DDPs (e.g., losses with regard to finances [2], data [3], or attention [4], increased frustration levels [5], and heightened cognitive load [6]). Typically, the examination of DDPs takes place in realistic
scenarios where participants are assigned tasks such as making a purchase, booking a flight, or browsing a news portal (cf.,[2], [7], [8]). On a mockup website, participants encounter the DDP designed to influence behavior in line with the mockup company's goals.

A prototypical example is the study conducted by Luguri and Strahilevitz in 2021. In an elaborate scenario with a large sample (3,777 participants), the effects of various DDPs such as Confirmshaming, Interface Interference, or False Hierarchy were investigated in mild (fewer DDPs, easier to circumvent) and aggressive forms (many successive DDPs, harder to circumvent). Initially, participants completed a questionnaire regarding privacy-related attitudes and were informed that the purchase of a privacy protection software was recommended based on their individual preferences. Instructions implied that participants would have to pay for this software with their own money. The software offers contained either mild or aggressive DDPs. Recorded metrics included acceptance rates (i.e., how many subjects purchased the privacy software), comprising a single data point per DDP and participant, as each participant had just one interaction with a specific DDP. This example illustrates the typical methodology in mockup experiments: In order to create the illusion of a “real” decision-making situation, participants find themselves in lifelike website environments and are required to adhere specific instructions. Accordingly, their behavior is observed for a specific DDP within a particular website environment.

1.1. Current Research Challenges

This method of data collection has the advantage of capturing actual behavior in relatively realistic scenarios, but comes along with several methodological and conceptual challenges. Mockup-based measurements often examine the targeted DDPs in just one trial, i.e., once in a specific situation. The experimental context (e.g., website, instructions) is designed to be as plausible and realistic as possible, aiming to create the illusion of a lifelike decision. Repeating these decision-making situations would undermine the credibility of the entire experiment. For instance, participants in the setup by Luguri and Strahilevitz would quickly realize that the decision has no real consequences, if they were repeatedly prompted to purchase the privacy software. Additionally, presenting the exact same decision situation multiple times would lead to habituation effects (e.g., individuals already know where to click). Hence, this one-trial testing approach presents a methodological challenge: Experiments with only a few trials for investigating a specific condition (i.e., one specific DDP) have weak statistical test power. Among other things, this reduces the probability that a significant result actually reflects a real effect. Measurements are susceptible to interference, making it difficult to ascertain whether the observed effect is due to the experimental manipulation (the DDP) or situational circumstances. In this regard, it is hard to achieve large effect sizes and obtain robust, replicable results [10].

There are numerous degrees of freedom in the design of a mockup experiment, e.g., the layout of the website, the color of buttons, the wording of instructions, or the transaction costs such as data or money (cf., [11]). The effectiveness of DDPs can be increased by factors such as trust in a website or the appearance of the user interface (i.e., DDP-compliant behavior; [5]). Additionally, DDPs can have different real-world consequences: users may lose money, personal data, time, or attention. The exact type of these costs can influence the response to a DDP (cf. the proposed behavioral taxonomy in [6]). Hence, the specific design
of the test environment adds additional noise and complicates a clear inference of the experimental results to the use of DDPs. Systematic control for these factors would therefore be necessary to draw conclusions about the effectiveness of a design pattern and to eliminate the distortion of results potentially caused by contextual factors. With reference to the example study by Luguri and Strahilevitz, the results should be contextualized within the experimental design (e.g., by highlighting the privacy-related instructions). Examining certain specifically designed DDPs allows only limited generalizations regarding their effects on user behavior.

Furthermore, in a rapidly transforming digital world, new design structures that are classified as manipulative or deceptive emerge constantly. The classification of various DDPs has been a continuously debated issue since the term was coined in 2010 [12], with numerous proposals for categorizing different patterns (cf. [13], [14], [15]). Gray and colleagues recently proposed an ontology that classifies Dark Patterns at three different levels (from overarching categories to very specific examples) [16]. This illustrates the complexity of the DDP phenomenon and its continuous expansion. Accordingly, experiments investigating the effects of individual patterns must take into account the complexity and expansion of DDPs in order to save effort when conducting experiments and to ensure the comparability of results through consistent data collection methods. Experiments that explore a specific DDP in a specific mockup environment lack the flexibility and adaptability to implement new forms of manipulative design.

Conceptually, many DDPs are assumed to operate by exploiting specific cognitive biases ([17]; e.g., the default bias as the preference for the already selected option in the DDP preselection). Biases come into play particularly in complex decision-making situations when numerous decisions have to be made simultaneously, time is short and/or not all options can be thoroughly examined and weighed up [18]. This raises the question to which extent an experimental setting, with relatively simple decision situations, can truly capture the "nature" of DDP’s effectiveness. It seems that the above-described experimental scenarios are less suitable for testing the underlying mechanisms of DDPs. To assess these mechanisms, conditions that might intensify biases need to be varied systematically (e.g., the level of distraction and complexity or time pressure). This cannot be guaranteed in an experimental setup with only a few decision trials. The greater the diversity within the sample for instance in terms of digital literacy or privacy attitudes, the more pronounced these issues become. This challenge is particularly relevant when studying socio-digital vulnerability (i.e., user groups vulnerable to specific biases and forms of deception [19]). How can we generate reliable statements about who is particularly vulnerable to a specific type of DDP when we do not control for the influence of contextual factors such as the respective website at the same time?

1.2. Open Research Questions and Motivation

The scenario-based research on DDPs using website mock-ups and realistic pattern simulations provides insights into the effectiveness of specific DDPs under certain conditions. However, despite more than a decade of intensive research on DDPs, several questions remain unanswered:
To what extent do users' behaviors and responses depend on the DDP itself, the website, the emerging costs (money, data, attention, time, etc.) or the combination of these factors?

Which mechanisms regulate the impact of DDPs on user behavior?

How can the individual socio-digital vulnerability of users be measured effectively?

1.3. Requirements for an Experimental Paradigm

The research questions outlined above call the existing approaches into question and emphasize the need for a new experimental approach to study the effects of DDPs. Accordingly, the following requirements for an experimental paradigm can be formulated.

**Control of Experimental Factors.** We need controlled experimental environments in which different conditions can be systematically varied and numerous trials can be carried out for specific combinations of conditions. This would allow to examine the impact of various deceptive designs, the different transaction costs (time, money, data, etc.) and contextual factors (complexity of the decision situation or time pressure). It would furthermore enable a systematic investigation of the effect on variables, such as behavior (decision for or against the DDP-intended option), the time required for decision-making, or certainty of the decision.

**Comparability and Comprehensibility.** The experimental paradigm should be applicable to various user groups. Taking into account the expanding research interest in socio-digital vulnerability (cf. [19]), an experimental framework to gain insights into the needs of different user groups and the vulnerability associated with DDPs is required.

**Flexibility.** The experimental paradigm should allow for a flexible adaptation to dynamic digital changes and the investigation of newly emerging DDPs. This saves resources, as no completely new experiment has to be developed for a newly defined DDP test, and makes it easier to compare the effects of different DDPs and variations of a DDP.

1.4. Aim of a New Method to Measure the Deceptive Potential

A paradigm meeting these requirements should address the described issues. By systematically varying measures of behavioral influence (such as decision-making, time, cognitive load) and assessing them under different conditions, a more robust testing of theoretical assumptions about the effects of DDPs can be achieved. Additionally, on a psychological level, the deceptive potential of DDPs can be determined, classified, and compared across user groups. Understanding the decision-making process when interacting with DDPs can enrich discussions on their classification and regulation. Effective countermeasures against DDPs require a better understanding of their behavioral drivers. Technological interventions can only be effective when applied strategically within the decision-making process. Empirical insights into individual reactions to different types of influence under various contextual conditions could provide valuable guidance. Moreover, legal interpretation, particularly regarding laws like the Digital Services Act in the European Union, needs refinement, including clearer definitions of "manipulation" and "deception" in the digital realm. A comprehensive understanding of decision-making principles in digital
design, along with individual vulnerabilities, is crucial for developing effective technological and legal countermeasures.

2. The Experimental Paradigm: A Decision-Making Game

Based on the questions above, we have developed an experimental paradigm, consisting of a decision-making game in which individuals are tasked with making favorable decisions in order to maximize gains. This gamified experimental approach was chosen to simulate real-life decision consequences (e.g., financial loss associated with selecting a DDP), aiming to render these consequences as lifelike and perceptible as feasible. The game is implemented in pygame and combines a steady, playful narrative with different variations of DDPs and contextual factors (e.g., different costs like money or data). In the following, we discuss the structure of the paradigm.

2.1. Structure of the Game

The overarching story of the game is that the character (i.e., the participant) has to search for ingredients at various places in a village to prepare a meal (see Figure 1 for the scenario in the Baseline [BL] condition). Villagers keep putting food items in front of their doors, which the player can collect. The goal is to collect as many ingredients as possible to fill the ingredient bar (see Figure 1, right side the bar above the tomato). In each trial (iteration), the player starts in the center of the village and can choose between various food items located in front of the houses.

The food items possess diverse values, each contributing differently towards filling the player's food bar to its maximum capacity. Food items with high value (tomato, see Figure 1B) score more points in the ingredient bar, providing a big incentive. Spices (see Figure 1C) have a lower value, yielding fewer points in the ingredient bar and are thus a small incentive. But the game is not just about collecting food items and filling the ingredient bar. The player also carries a cookbook and must prevent losing its pages. Sometimes, villagers demand that, in exchange for a food item, a page of the cookbook is left behind. The player's task is to give up as few pages of their cookbook as possible. This is depicted on the right-hand sight in Figure 1A (i.e., the bar above the cookbook symbol). Hence, the aim of the game is to collect as many ingredients (i.e. tomatoes and spices) as possible while minimizing the loss of cookbook pages. Accordingly, there is a better option with a big incentive (the simple tomato, see Figure 1B) and a worse option with a small incentive (the tomato with a cookbook symbol, see Figure 1D).

The small and big incentives are located on opposite sides and are equidistant from the character. The character is controlled with arrow keys and can move in all directions, including diagonally by pressing two arrow keys simultaneously (e.g., left and up = diagonal up). Participants receive a comprehensive explanation of the game's functionality and significance at the beginning of the experiment. In order to motivate the participants, we plan to reward them with a performance-based payout (i.e., a monetary amount based on their achievements in the ingredient bar in relation to the number of lost cookbook pages). The current score is displayed on the right side of the screen. Thus,
participants can “track” their performance in real-time. Participants are informed in advance that they only have limited time to collect food item in order to prevent individuals from taking too much time for each decision.

Figure 1: Design of the Experimental Paradigm including the game field and player (A) and the different incentives (B – D).

2.2. Parameters Influencing Decision-Making in the Paradigm

The game should enable the measurement of factors that influence decision-making (i.e., contextual factors and specific DDPs) under stable conditions. Below we discuss possible use cases and limitations of applicability.

Implementation of Contextual Factors. Contextual factors refer to the real consequences or costs incurred by individuals through the use of DDPs. Given the widespread adoption of DDPs in e-commerce and privacy-related contexts, costs can include financial harms and loss of personal data. These costs can engage two different psychological mechanisms: experiencing loss (e.g., data or actual financial losses) or receiving less (e.g., selecting a "worse" offer) [9]. Since there is a difference between losing something you already have and receiving less than expected, both mechanisms are implemented (high vs. low gain, loss vs. no loss). The food items represent the gain condition (i.e. money), and the cookbook represents the loss condition (i.e. data). This implementation serves as a heuristic for real-world mechanisms (reduced gain vs. loss). Therefore, an absolute interpretation, such as quantifying financial losses, is not feasible. However, these abstract mechanisms are present in many DDPs and decision situations, ensuring the broad applicability of the paradigm. Costs at the psychological level (such as capturing attention or negative emotions) are not covered by the current implementation. However, measurement could be facilitated, for instance, through eye-tracking analysis (tracking visual attention) or additional ratings on emotion questionnaires during the experiment.

Implementation of DDPs. Our paradigm focuses on decisions involving two options that can be "better" or "worse" for the player’s goal. Thus, the utilization of deceptive design elements aims to enhance the small incentive or devalue the big incentive. Various
deceptive strategies can be employed on different levels: visual (e.g., color highlighting or concealing information), cognitive (e.g., misleading language or symbolism), and motivational (e.g., countdowns to increase pressure for a decision). So far, the paradigm focused on situations in which a choice has to be made between two options. However, it is conceivable to implement attention-related DDPs (directing attention to a specific option), or DDPs that leave no choice (e.g., forced action, forced continuation). Since the paradigm also allows for the measurement of time or the character's movement trajectory (see section 2.4), data could be collected on whether and when participants accept an offer and how the character moves accordingly.

2.3. Experimental Design

First, we focus on the investigation of two different costs (money vs. data) and six different DDP-conditions (one baseline and five DDP conditions), resulting in a 2x6-within subject-design. These five patterns were selected for practical reasons to enable the validation of the paradigm with data from realistic scenarios (see section 2.4). Each pattern exhibits a clearly identifiable visual signal and is frequently employed in real-world applications. Therefore, due to their feasibility and practical relevance, we have chosen these patterns. The five employed DDPs are further elucidated below. Each represents an abstraction of real-world DDPs, accompanied by initial proposals for their implementation. Additionally, we measure a baseline condition (BL), already depicted in Figure 1.

**Aesthetic Inference (AI).** On the visual level, the small incentive is highlighted (bright circle around it), while the large incentive is less visible (see Figure 2A). This DDP is often used in cookie banners on websites, where the option to accept all cookies is visually accentuated, while the "reject" option is hard to perceive.

**Countdown (CD).** Alongside the small incentive, a countdown of 10 seconds is initiated (see Figure 2C). Once expired, a new trial begins—thus, the countdown carries a real consequence. This time expiry mirrors the real-world implementation of the frequently occurring DDP on shopping websites, where a countdown aims to prompt users into quicker decision-making.

**Obstruction (OS).** In this condition, the large incentive is surrounded by obstacles (tree trunks, see Figure 2D), making it more challenging to reach. OS is observed wherever the execution of an action is impeded, such as when extra clicks are necessary to deselect an option (e.g., cancel newsletter subscription). It requires an extra click, hence an additional effort, to reach the non-DDP-conforming option. This extra effort to access the "higher-value" option is represented through the tree trunks.

**Social Proof (SP).** Represented by five stars, the smaller incentive is upgraded (see Figure 2E). This attempts to simulate SP from website contexts, where portraying a (perceived) majority opinion suggests the popularity of a product. We propose that stars, being a widely recognized symbol for the social rating of a product, can simulate a similar mechanism.

**Wrong Signal (WS).** This DDP misleads through the use of symbols that mean something different, such as using a lock as a symbol for data protection or privacy next to an option
that is less data-friendly [20]. This DDP is represented as a “thumbs up” next to the smaller incentive (see Figure 2F), aiming to make it appear more positive.

![DDP Images](image1.png)

**Figure 2**: Implementation of the DDPs (A) Aesthetic Inference, (C) Countdown: “Just 5 seconds left!”, (D) Obstruction, (E) Social Proof, and (F) Wrong Signal.

### 2.4. Measures of Deception

Considering the measurement of the deceptive potential of DDPs, it appears crucial to identify measures for the decision itself but also for processes underlying the decision, enabling a more nuanced characterization.

**Behavior.** Principal behavior is measured in the form of the decision, indicating which reward was collected (i.e., DDP-conforming vs. non-DDP-conforming, small vs. big incentive).

**Cost of a Decision.** It can be assessed based on the time it takes to collect a reward; shorter times would reflect lower decision costs, based on the assumption that less thought was given to individual options or the decision as a whole.

**Decision Uncertainty.** The unique design of the game allows us, prospectively, to measure decision uncertainty (although it would require switching the game control from keyboard to mouse). In each trial, the character starts from exact the same position, with rewards consistently positioned at the same distance. By tracking the character’s coordinates on the path to the reward (*mouse tracking*), the precise movement trajectory can be reconstructed. Does the character move directly toward the collected reward or after approaching another direction (see Figure 3)? This allows for the representation of cognitive processes underlying the decision, especially in terms of how conflicted or certain the decision is (for a review on process-tracing methods in decision making see [21]).

![Movement Trajectories](image2.png)

**Figure 3**: Hypothesized movement trajectories of the character for (a) certain decisions (dashed black line) and (b) less certain decisions (solid white line).

### 2.5. Validation

The next step is to validate the game. Therefore, participants will first engage in an online experiment featuring website mockups and realistic decision scenarios. Each of the
described DDPs will be presented in two trials. Subsequently, individuals will be invited to
a laboratory experiment where they will navigate the described paradigm in the form of a
decision-making game. The results of both study parts will be compared using equivalence
tests [22] with regard to decision behavior. In order to confirm the validity of the game
measures, we would predict the following findings: (i) DDPs that show a higher influence
on participants' behavior in the online experiment should similarly impact their behavior
in the game, and (ii) participants who were more strongly influenced by DDPs in the
online experiment should also demonstrate this in the game. Initial results regarding the
validity of the game should be available by May 2024 and could be discussed throughout
the workshop.

2.6. Open Questions

The experimental paradigm presented here offers an initial solution to the methodological
and conceptual challenges in assessing DDPs. However, several open questions arise that
need further discussion or investigation in experimental pilot studies. The abstract
implementation of DDPs is one way to depict them, but alternative design forms should be
explored to ensure that the "deceptive core" of each respective DDP is captured. The
validation study will reveal how well the DDPs implemented in the game resemble those
found on websites. Additionally, we would like to explore further contextual factors or
extend the paradigm to include DDPs primarily targeting user attention. The precise
implementation in these cases needs careful consideration and conceptualization.

3. Conclusion

Experiments have revealed numerous harmful effects of DDPs on users. Nevertheless,
collecting data in real-world scenarios on mockup-websites inherently encompasses
certain methodological weaknesses, hindering the generalization of results, testing the
underlying mechanisms of deceptive designs, and accurately comparing user groups to
better understand vulnerability to DDPs. For this reason, we propose a new
methodological direction and advocate for an experimental paradigm that can control
influencing factors effectively, adapt flexibly to new DDPs, and systematically investigate
socio-digital vulnerability. To deal with these challenges, we have developed a decision-
making game in which participants choose between a small and a large incentive and are
tasked with collecting as many incentives as possible. The decision-making process is
impeded by the deployment of DDPs, which either devalue the large incentive (e.g., by
making it more challenging to attain) or enhance the small incentive (e.g., by making it
more visibly prominent). So far, five different DDPs are implemented, but the list of
patterns can be extended and is interchangeable. The decision-making game experiment
serves as an initial approach to examine DDPs under these requirements. With this
paradigm, we aim to make robust statements about the nature of influence and the
underlying mechanisms behind DDPs, thus contributing to a more precise understanding
of how DDPs operate. To apply technological and legal countermeasures effectively, we
must measure the deceptive potential of design structures more precisely.
Acknowledgements

The project as part of the Disruption and Societal Change Center at the Technical University of Dresden can be realized through funds from the State of Saxony, Germany.

This Word template was created by Tiago Prince Sales (University of Twente, NL) in collaboration with Manfred Jeusfeld (University of Skövde, SE). It is derived from the template designed by Aleksandr Ometov (Temper University of Applied Sciences, FI). The template is made available under a Creative Commons License Attribution-ShareAlike 4.0 International (CC BY-SA 4.0).

References


