

CHI Paper Review

Josh Hochman

An Analysis of International Studies of Dark Patterns in Modern Apps

Hidaka, S., Watanabe, M., Kobuki, S., & Seaborn, K. (2023). Linguistic Dead-Ends and Alphabet Soup: Finding Dark Patterns in Japanese Apps. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*, 1–13. <https://doi.org/10.1145/3544548.3580942>

Chordia, I., Tran, L.-P., Tayebi, T., Parrish, E., Erete, S., Yip, J., & Hiniker, A. (2023). Deceptive Design Patterns in Safety Technologies: A Case Study of the Citizen App. *CHI '23: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–18. <https://doi.org/10.1145/3544548.3581258>

1) What are the research methods used in each paper? Describe each research method, and then compare and contrast the methods used in each paper. (Note: It is possible that the papers that you chose each use more than one research method)

In the realm of user experience design, researchers are increasingly identifying the presence of so-called *dark patterns (DPs)* in digital products. Although the exact definition of this term can vary from source to source, a dark pattern generally refers to a way of subtly but intentionally designing a user experience that manipulates the user's attention and decision-making process, resulting in a deviation from the user's original intent. The two studies examined here look at DPs through two different lenses: a macroscopic analysis of how dark patterns differ in Eastern vs. Western social context, and a microscopic analysis of dark patterns specific to an app marketed for user safety.

In "Linguistic Dead-Ends and Alphabet Soup: Finding Dark Patterns in Japanese Apps", Hidaka et al. explore dark patterns in the context of East Asian culture. By examining historical studies of this subject, the authors found that dark patterns had previously only been studied in the context of Western cultures. This study, they claim, serves as the first look at dark patterns specific to Japan. To learn more, the researchers adapted a procedure similar to that of a 2020 study by Di Geronimo et al., but applied it specifically to the context of Japanese users and adjusted the study's design accordingly (Hidaka et al., 2023). The primary objectives of the study included (a) identifying which dark patterns dominate Japanese apps, (b) understanding the distribution of DP types across different categories of apps, and (c) comparing the frequencies of DP types between the Japanese and US app markets (Hidaka et al., 2023). The researchers applied a hybrid of trace and observation methodology for their study by interacting with 200 apps and capturing the interactions with a screen recorder. They then analyzed each recording to identify which dark patterns were present, using the classification framework defined by Grey et al. and Di Geronimo et al. Interestingly, they worried about bias resulting from "*dark pattern blindness*", a term the authors coined to refer to a researcher's inability to identify when they were under the influence of a dark pattern (Hidaka et al., 2023). To mitigate this risk, two researchers analyzed every recording, instead of just one. As part of their analysis, the team identified two dark patterns unique to the Japanese context which they named *Untranslation* and *Alphabet Soup* (Hidaka et al., 2023).

Meanwhile, in "Deceptive Design Patterns in Safety Technologies: A Case Study of the Citizen App", Chordia et al. apply a very different methodology for examining dark patterns. This case study examines dark patterns specific to Citizen, a crime alert app, in the context of its usage in the US city of Atlanta. The researchers began by creating a user screener survey targeting Citizen users and posting it to social media sites; they then reached out to a subset of the respondents and requested user interviews. They found that the resulting user group lacked substantial diversity, so they performed a second round of screeners to target areas with predominantly Black populations. Although this improved the diversity of the final 15-person sample, the authors acknowledged that it still underrepresented Black populations (Chordia et al., 2023). The researchers recorded how long each user had been using the app and the amount of time they typically used the app each week. They then conducted user interviews with these individuals, inquiring about "which features they used in the app", "why and how often they used them", and "how these features influenced their behavior and beliefs" (Chordia et al., 2023). After the interviews, the team triangulated their study by performing an interface review of the Citizen app, analyzed their observations/findings, and then integrated the data from the user interviews and the interface review into a single dataset.

Both studies used a combination of qualitative and quantitative data. The main takeaway from the first study was a focus on the qualitative findings, namely that there are two distinct dark patterns that emerged in the Japan data that were not present in historical studies of Western data. The key finding of the second study also focused more on the qualitative effects of using the Citizen app than on their quantitative metrics (such as frequency of app use). However, the two studies differed significantly in the methodologies used to reach their conclusions. Hidaka et al. essentially served as the user group for their study; their team undertook the task of analyzing apps, categorizing the findings, and analyzing the results. They predominantly followed the procedure of a previous research team, but applied it to a different sector of the app market. On the other hand, Chordia et al. developed a new procedure to identify a representative sample of external users for their study. Interestingly, both studies used and referenced the same framework for defining types of dark patterns, published by Di Geronimo et al.

2) What are the key results from each paper? Describe each paper, and then compare the results.

By examining dark patterns in the context of the Japanese app market, Hidaka et al. identified three key findings. Their first objective question asked if dark patterns exist in the Japanese context; the study concluded that they most certainly do. Second, the team defined a new class of dark patterns not previously observed in studies of dark patterns in Western contexts, which they named *Linguistic Dead Ends* (Hidaka et al., 2023). They further broke this category down into the categories of *Untranslation* and *Alphabet Soup*. *Untranslation* refers to the practice of advertising an app (often on an app store) as running in one language, but leaving select features within the app in another (often unfamiliar) language, typically without an easy way to translate them (Hidaka et al., 2023). This practice typically benefits the app developer and harms the user. As an example, the researchers identify the terms of service page as a common culprit source of this pattern in apps on the Japanese market. *Alphabet Soup* describes a pattern in which an app displays a foreign word using the written characters of the local language, resulting in an obfuscation of the word's meaning (Hidaka et al., 2023). Finally, the researchers compared the frequency and distribution of the different types of dark patterns identified in the 200 Japanese apps versus the corpus of the original Di Geronimo study. While both studies showed similar percentages of apps that had at least one dark pattern (93.5% for Hidaka et al., 95% for Di Geronimo et al.), the average number of DPs per app was on average about half in the Japan study (3.9 DPs per app) as compared to the original study (7.4 DPs per app) (Hidaka et al., 2023).

Similarly, in their analysis of the Citizen app, Chordia et al. definitively identify the presence of dark patterns in an app marketed as a way to keep users safer. Interestingly, since user race plays a key role in their findings, they choose to use the verbiage “deceptive design patterns” instead of “dark patterns” to avoid associating the word “dark” with race, or implying that the word “dark” describes a form of negative behavior. After interviewing their 15 participants, the researchers created one dataset by reviewing the interview content and organizing it by codes and overarching themes. Next, they created a second dataset by analyzing and identifying dark patterns within Citizen, using affinity diagramming to identify the deceptive patterns and align the common themes. Finally, they merged the two datasets together based on these common themes and passed the data to an independent group of researchers for the final analysis and feedback. From a high-level perspective, the study concluded that Citizen creates “an inflated sense of danger” among its users, while also marketing itself as a way to assuage that fear (Chordia et al., 2023). Additionally, they identify several notification features in the app that

directly contribute to this fear. First, the app warns users about alleged dangers that aren't even in their geographic area, as well as isolated events that pose no risk to public safety. It also warns users about events that happened in the past, and notifies them about some events without providing any detail that would allow the user to determine whether those events warrant any concern. Finally, the app uses "forced action" by requiring the user "to turn on alerts in order to read the Notification Feed" (Chordia et al., 2023).

3) Compare and contrast the writing style of each of the two papers. (Formal? Informal? Technical? Data-based? Persuasive? Do they use lots of figures and graphics? Or numbers and formulae? In general, how are the authors making their arguments?)

Both papers present their data with the type of formal language that I would expect in a complex research text. However, in "Deceptive Design Patterns in Safety Technologies: A Case Study of the Citizen App", Chordia et al. appear more willing to intersperse some sections with an informal vibe, opting for less technical jargon in favor of a more relatable flavor of storytelling. In their analysis of their user interviews, the authors include a number of italicized direct user quotes; the use of informal language in this context (for example, when Participant P2 says "*there's always a little action right around me because I'm by Edgewood*") provides a welcome relief from the more data-heavy portions and instills a sense of human familiarity in the reader (Chordia et al., 2023). The paper ultimately sounds as if the authors are telling a story as much as they are sharing their study and its findings. They introduce us to their work not by defining the study objectives or highlighting its conclusions, but rather by orienting us to the subject material through summary of the relevant history and ensuring that all readers begin their reading of the document with an equal understanding of the definition of deceptive design. While the paper does include information about the quantitative side of the study, the bulk of the text focuses more on asking the question of why the team might have observed the things they observed, as well as proposing possible answers. In a way, the tone almost seems to speak directly to the reader, encouraging and challenging them to look for answers as well.

In "Linguistic Dead-Ends and Alphabet Soup: Finding Dark Patterns in Japanese Apps", on the other hand, Hidaka et al. undoubtedly place a greater emphasis on their numerical findings, presenting information with an air of formality. Both studies use colorful images to emphasize key points. However, the imagery presented by Hidaka et al. has a more noticeable flair. By using bright colors with a conservative amount of text in their numbered figures, they seemingly invite readers to ponder the meaning of the content for a few extra moments. In contrast, the majority of the color imagery employed by Chordia et al. appears as arrays of darker colored app screenshots. Given their relatively small size, and without knowing the full context of what was happening when these images were captured, they are easier for the reader to skip over quickly. If one were to summarize the tone of Hidaka et al.'s presentation using everyday language, it might read as follows: *We wanted to find out if designers use the same kinds of sneaky tricks in Japanese apps as Western apps. Guess what? Not only were these patterns present in the Asian context, but we also found some new forms of deceit not previously seen in data from the Western context.*

4) What are your perceptions of this research topic, and the specific contributions made by the authors in their papers?

As a future UX designer, it comes as no surprise to learn of the pervasive nature of dark patterns in modern app design. I found it particularly surprising when the authors claimed that they were the first to look at this issue in apps on the Japanese market. The discovery of two new and unique DPs by Hidaka et al. raises many questions about their global presence. As explained by Henrich, Heine, and Norenzayan and referenced by Hidaka et al., most studies into dark patterns have examined nations described as “WEIRD”: Western, Educated, Industrial, Rich, and Democratic. Given that “Linguistic Dead-Ends and Alphabet Soup: Finding Dark Patterns in Japanese Apps” identifies *linguistic dead ends* as a category of dark patterns not previously observed, it begs the question of why we haven’t yet identified them in the Western context. For example, let us briefly return to Hidaka et al.’s concept of Untranslation: the practice of advertising an app on an app store as running in one language, but leaving select features within the app in another (often unfamiliar) language. What would I find if I looked for apps on Google Play where the primary language is Spanish? Would these apps include terms and conditions in English? It seems hard to believe that this phenomenon is unique to Japanese apps.

Reflecting on Chordia et al.’s findings, I find it concerning that Citizen, an app whose goal is to make users feel safer by keeping them informed of nearby crime-related incidents, appears to be making them feel less safe by increasing their anxiety. Although the case study centered on Atlanta, it seems likely that anyone living in proximity to a city known for crime could expect similar results. As someone who lives on the outskirts of Baltimore, I occasionally venture into the city. During these outings, I experience a heightened sense of alertness and caution, knowing the city’s reputation as a dangerous place. My general knowledge of which areas of the city to avoid has historically boiled down to Google searches for “most dangerous Baltimore neighborhoods” and adopting a habit of avoiding the city after dark, if possible. After examining Chordia et al.’s findings, it makes me question to what extent this unease has been fueled by learning about crimes that would never have had a chance of affecting me. By comparison, I rarely read crime reports for the Baltimore suburb I call my home, as I perceive it as safer than the inner city. Crime is everywhere; it’s unlikely that my town is perfectly safe. As someone already naturally prone to anxiety, if I began using Citizen to increase awareness of incidents occurring here in my suburb, would I begin to feel less safe?

5) What do you think the next steps in future research on this topic should be? (Note: We are asking your opinion, not the opinions of the paper authors. Please do not repeat what the authors listed as “next steps for research.”)

In “Linguistic Dead-Ends and Alphabet Soup: Finding Dark Patterns in Japanese Apps”, Hidaka et al. examined how cultural differences played into patterns observed in the Asian context. What might the findings look like if we were to do a similar study in other Asian contexts that only partially fit the WEIRD definition? For example, by American standards, the People’s Republic of China and the Russian Federation could both be described as educated, industrial, and rich, but probably not as Western or Democratic.

In “Deceptive Design Patterns in Safety Technologies: A Case Study of the Citizen App”, Chordia et al. explicitly state that they chose Atlanta as their focus city because people living there were already

concerned about crime before participating in the study. This inherently leads me to a glaring question: what might the researchers find if they executed a similar study in a city less associated with crime? If you do a Google search asking for the world's safest cities, many bloggers will list Reykjavik, Iceland or Tokyo, Japan. If Chordia et al. executed the same procedure with residents of Tokyo as they did in Atlanta, how might those residents describe the experience? Would they similarly demonstrate increased levels of anxiety? Or might they instead be more apt to not believe the information being put out by the app?

Additional sources:

Di Geronimo, L., Braz, L., Fregnan, E., Palomba, F., & Bacchelli, A. (2020). UI Dark Patterns and Where to Find Them: A Study on Mobile Applications and User Perception. *CHI '20: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–14. <https://doi.org/10.1145/3313831.3376600>

Gray, C. M., Kou, Y., Battles, B., Hoggatt, J., & Toombs, A. L. (2018). The Dark (Patterns) Side of UX Design. *CHI '18: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–14. <https://doi.org/10.1145/3173574.3174108>

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Most people are not WEIRD. *Nature*, (466), 29–29. <https://doi.org/10.1038/466029a>

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, 33(2–3), 61–83. <https://doi.org/10.1017/S0140525X0999152X>