When Does Context Harm Brands? Manipulating Feed Composition to Study Brand Safety

Case Study 1

November 14, 2024

Our first study demonstrates DICE's capability to study context effects with high experimental control and study realism. This study showcases how researchers can systematically manipulate the broader context in which users see a specific post (here, a sponsored post of a firm). Thus, we systematically varied the context (i.e., feeds) in which participants encountered a sponsored post.

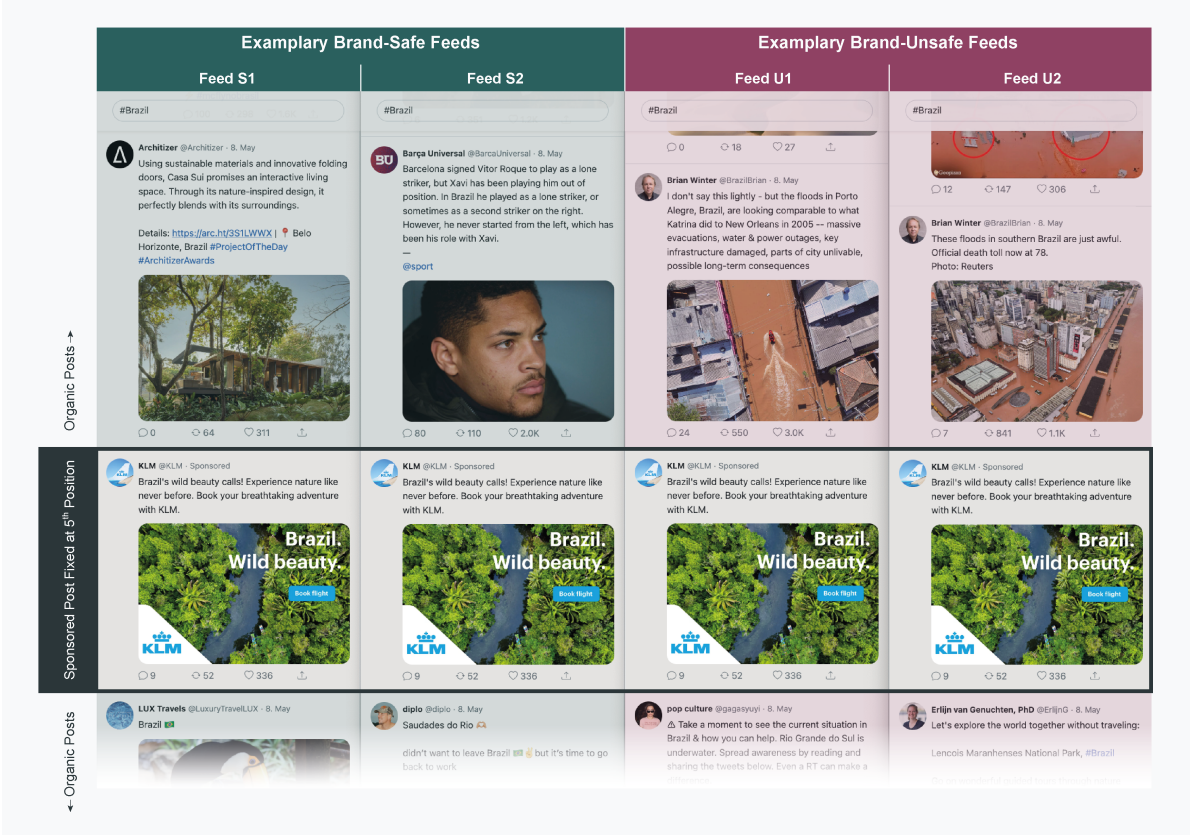
We demonstrate this capability by examining brand safety in social media advertising. Brand safety refers to the idea that advertising should not appear in contexts that could harm a brand's reputation (Fournier and Srinivasan 2023; Porter 2021). This concern is particularly relevant for social media advertising, where platforms use automated systems to place ads in dynamic, user-generated content environments. These systems often lack the nuanced understanding needed to identify potentially problematic contexts that could harm a brand. While industry reports suggest that up to 75% of brands have experienced such unsafe brand exposures (Ahmad et al. 2024; GumGum Inc. 2017), systematically examining these effects in the field risks apparent brand damage. DICE offers researchers full control to vary the feed composition participants experimentally are exposed to and test the presence of such context effects that are difficult (if not impossible) to assess with other methods, with tight experimental control and granular behavioral data to understand individual user experiences [Simonov, Valletti, and Veiga (2024);].

Experimental Design

To examine how brand-(un)safe contexts affect brand perceptions, we created two social media feeds that were identical in structure but varied in their content surrounding a sponsored post. Figure 1 displays exemplary screenshots of the implementation in DICE: both feeds

contained the same number of organic posts and one identical sponsored post (i.e., ad), with the only difference being the thematic focus of the organic content (brand-safe vs. unsafe content surrounding the ad). The sponsored post was an advertisement of an airline (i.e., KLM) promoting flights to Brazil. The post featured the tagline: *"Brazil's wild beauty calls! Experience nature like never before. Book your breathtaking adventure with KLM."* In the brand-safe condition, the sponsored post was surrounded by actual tweets covering Brazil that we scraped from the web. In the brand-unsafe condition, however, the sponsored post was surrounded by posts covering the severe flooding that happened during the time of the study (in the summer of 2024). The sponsored post was fixed in the fifth position in both conditions, while all organic posts varied randomly. This sponsored post becomes insensitive when juxtaposed against a feed focusing on the natural disaster in the region advertised. Such a situation is precisely the type of contextual mismatch that automated systems can create and managers fear because of adverse consequences for brands (Ahmad et al. 2024; GumGum Inc. 2017).

Figure 1: Exemplary DICE Feeds from Study 1



Procedure

As in the first case study, participants browsed the simulated feed on their own devices (75% desktop, 21% mobile, and 4% tablet). After scrolling through the feed, we redirected participants to a Qualtrics survey in which they first provided demographic information as a filler task. Next, we measured participants' brand perceptions by assessing their attitude toward KLM using three seven-point scales (1 = "Negative/Unfavorable/Dislike" and 7 = "Positive/Favorable/Like"; $\alpha = 0.96$). Finally, we assessed participants' awareness of the Brazil flooding before debriefing the study and redirected them to the recruitment platform (here Prolific). We provide our stimuli, materials, data, and analysis code on the Open Science Framework: https://osf.io/2xs5c/?view_only=4bf95d2a2c8449218b5fa7cd288f626a.

Configuration

Following the approach described in the app implementation section, we created a single CSV file containing all necessary information for the study (e.g., all organic and one sponsored post, randomization procedure, condition assignment, etc.). As shown in Figure 2, each post's content was specified in columns text (actual post) and username (account placing the post). The condition parameter assigns each row (i.e., each post) to its respective condition. We left the sequence column empty for organic posts to fully randomize them and fixed the sequence parameter to "5" for the sponsored KLM post. We marked this post as sponsored using the sponsored parameter and included a KLM landing page URL in the target column for participants who clicked on the ad. Finally, the resulting CSV file was uploaded to an online repository to generate a URL for the DICE app. The DICE app can handle links to any publicly accessible repository such as a Google Drive, GitHub, or other platforms.

Figure 2: Study Configuration

doc_id	username	handle	sponsored	target	sequence	condition	text	likes	reposts	replies
1	KLM	KLM	1	https://www.klm.co	5	safe	Brazil's wild beauty calls! Exper	336	52	9
2	KLM	KLM	1	https://www.klm.co	5	unsafe	Brazil's wild beauty calls! Exper	336	52	9
3	Architizer	Architizer	0			safe	Using sustainable materials an	311	64	0
4	LUX Travels	LuxuryTravell	0			safe	Brazil 🇧🇷	82	19	3
5	Barça Universal	BarcaUnivers	0			safe	Barcelona signed Vitor Roque t	2013	110	80
6	diplo	diplo	0			safe	Saudades do Rio 🫶didn't want	12123	571	123
7	Brian Winter	BrazilBrian	0			unsafe	These floods in southern Brazil	1123	841	7
8	Brian Winter	BrazilBrian	0			unsafe	I don't say this lightly - but the flo	3020	550	24
9	Erlijn van Genuchten,	ErlijnG	0			unsafe	Let's explore the world together	27	18	0
10	Evil Scientist	DarkLas_	0			unsafe	Brazil's ex president Bolsonaro	40	15	3

Participants and Randomization Checks

We recruited 982 US-American participants ($M_{age} = 39$ years; 56% female) from Prolific. Participants were randomly assigned to view either the brand-safe feed (featuring general Brazil-related content) or the brand-unsafe feed (featuring flood coverage). A critical advantage of DICE over observational and platform studies is its ability to implement true random assignment, allowing us to isolate the effect of context while canceling out other factors that might influence brand perception. To validate DICE's randomization functionality, we examined the balance between treatment groups, as illustrated in Table 1, the two treatment groups do not exhibit differences in observables. We found further support for effective randomization on observables via an omnibus test of joint orthogonality with randomization inference, $p = 0.39$ (Kerwin, Rostom, and Sterck 2024).

Table 1: Covariate Balance Across Conditions

Covariate	Appropriate	Inappropriate	Difference	p-value
Mean Age (Years)	38.848	38.522	-0.326	0.702
Female (Percent)	58.439	54.262	-4.177	0.194

Data

Our dataset comprises 982 participants and 19,640 observations at the participant × post level. We focus on a single sponsored post (i.e., the KLM ad) viewed by all participants. Hence, our final sample comprises 955 observations on the participant level, which is slightly less than the expected one observation per participant due to technical issues.

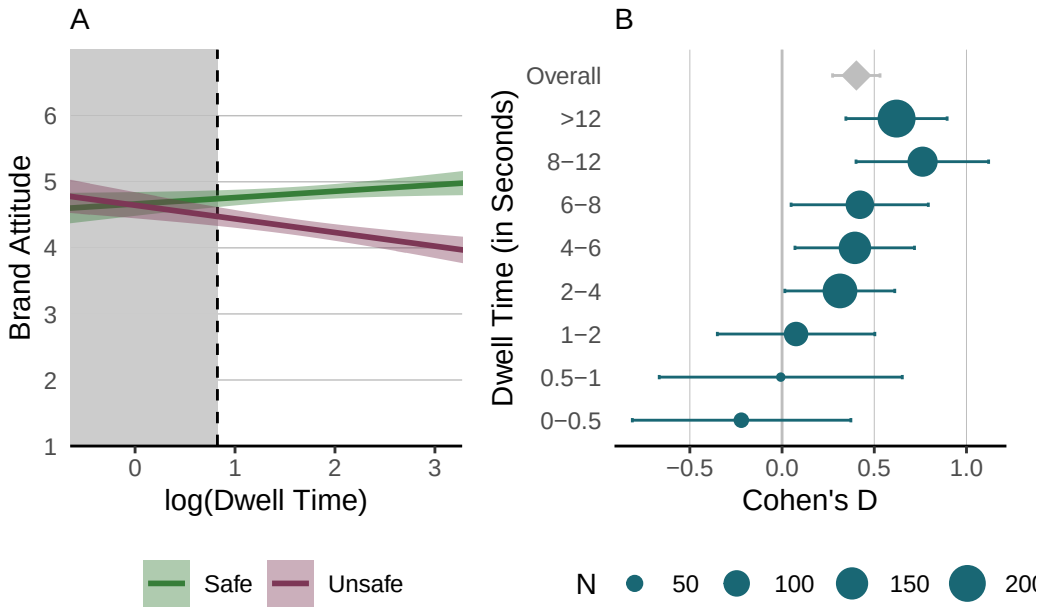
The resulting data structure has two convenient methodological implications. First, because we analyze one observation per participant rather than nested data, we can apply simpler methods, such as ordinary least squares (OLS) regressions. Second, because we only focus on one sponsored post, we do not need to divide our dwell time measure by the post's height as we did in Case Study 2. Both aspects increase the interpretability of our results.

Results

We found significantly more negative brand attitudes toward KLM in the unsafe feed condition ($M_u = 4.31$, $SD_u = 1.366$) were significantly less favorable than in the safe feed ($M_s = 4.821$, $SD_s = 1.161$, $\beta = -0.51$, $SE = 0.082$, $t = -6.217$, $p = 0.000$, $d = 0.403$).

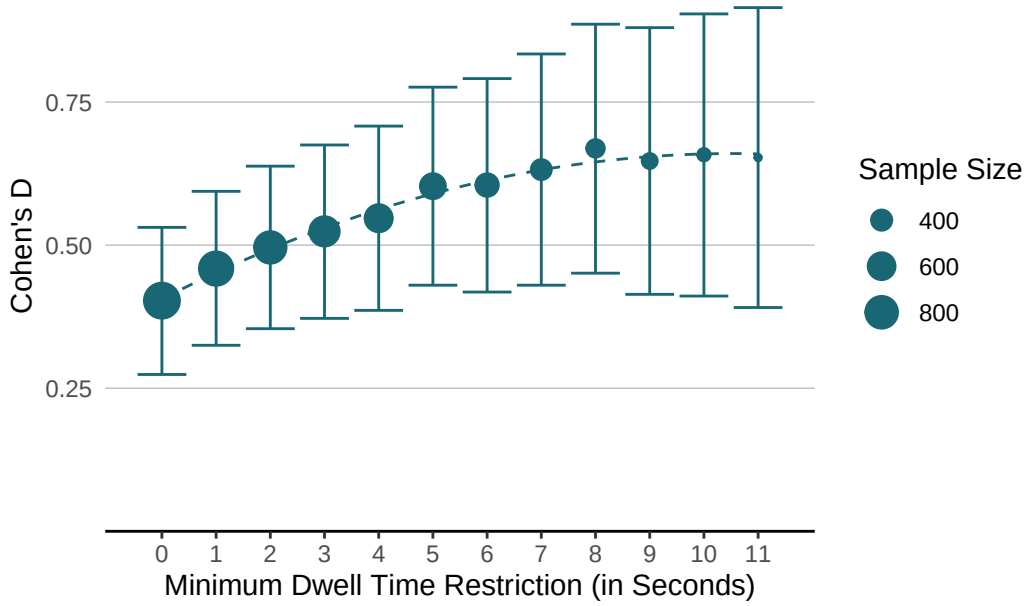
DICE's user interaction data further allows us to test whether this negative effect of unsafe content is conditional on the extent of attention that users devote to processing content in their feed. To further explore the interplay between context and brand attitudes, we examined whether the dwell time of the KLM ad (i.e., how long the sponsored post was inside the user's viewport) moderated the previously reported main effect. An OLS regression revealed a statistically significant interaction between the brand safety context manipulation and dwell time ($\beta = -0.302$, $SE = 0.068$, $t = -4.455$, $p = 0.000$), such that the negative effect of an unsafe context on brand attitude was more pronounced when participants spent more time viewing the sponsored post. In contrast, the effect was much smaller for participants who did not spend much time on the sponsored post. This moderation is robust to alternative model specifications where we repeated the same analysis while controlling for the dwell time dedicated to all organic posts ($\beta = -0.314$, $SE = 0.076$, $t = -4.112$, $p = 0.000$). We show this moderation in Panel A of Figure 3. In addition, we also show how the effect size varies as a function of the absolute dwell times in seconds (see Panel B, Figure 6) to further illustrate this interaction effect.

Figure 3: Moderation of the Effect of Context on Brand Attitudes by Dwell Time



Additionally, researchers could also consider using the dwell time measures as a screening device. Figure 4 illustrates how the effect size varies as a function of using different thresholds (i.e., minimum dwell times) for excluding participants. These analyses are consistent with the moderation results reported previously as the effect sizes increases with more restrictive minimum dwell times.

Figure 4: Using Dwell Time To Screen Study Participants



Implications

Our first study demonstrates DICE's ability to test relevant and difficult-to-study context effects with high experimental control. We also show how the behavioral data generated by the DICE app allows researchers to better understand how attention (as proxied through dwell times) affects brand perceptions, showcasing DICE's capacity to reveal behavioral mechanisms that are difficult to study with existing paradigms. From a substantive perspective, the findings provide experimental support for brand safety concerns and demonstrate that contextual misplacements can actively harm brand perceptions.

On a methodological level, our first study demonstrates how researchers can manipulate not just single social media posts but entire feed compositions and how DICE's behavioral data can serve multiple methodological purposes: using dwell time data as a moderator to assess the strength of an effect (i.e., showing that the negative effect of unsafe feeds increases with enhanced dwell time) or as a data quality or attention check (i.e., demonstrating variation in effect sizes due to different thresholds for filtering out participants that did not properly engage with the focal post). Using the behavioral tracking data for moderation analyses and data quality checks serves as a practical example of how dwell time data can be used as an unobtrusive proxy to measure *manipulation intensity* (Krefeld-Schwab, Sugerman, and Johnson 2024).

References

Ahmad, Wajeeha, Ananya Sen, Charles Eesley, and Erik Brynjolfsson. 2024. "Companies Inadvertently Fund Online Misinformation Despite Consumer Backlash." *Nature* 630: 123–31. https://doi.org/10.1038/s41586-024-07404-1.

Fournier, Susan, and Shuba Srinivasan. 2023. "Monitoring Marketing Sources of Brand Reputation Risk." *NIM Marketing Intelligence Review* 15 (1): 46–53. https://doi.org/doi:10.2478/nimmir-2023-0007.

GumGum Inc. 2017. "The New Brand Safety Crisis: A Fractured Environment." GumGum. https://web.archive.org/web/20220317063148/https://insights.gumgum.com/hubfs/Brand_Safety_GumGum.pdf.

Kerwin, Jason, Nada Rostom, and Olivier Sterck. 2024. "Striking the Right Balance: Why Standard Balance Tests over-Reject the Null, and How to Fix It." IZA Discussion Paper 17217. Institute of Labor Economics (IZA). https://www.iza.org/publications/dp/17217/striking-the-right-balance-why-standard-balance-tests-over-reject-the-null-and-how-to-fix-it.

Krefeld-Schwab, Antonia, Eli Rosen Sugerman, and Eric J. Johnson. 2024. "Exposing Omitted Moderators: Explaining Why Effect Sizes Differ in the Social Sciences." *Proceedings of the National Academy of Sciences* 121 (12): e2306281121. https://doi.org/10.1073/pnas.2306281121.

Porter, Jonathan. 2021. "Commentary: Inefficiencies in Digital Advertising Markets: Evidence from the Field." *Journal of Marketing* 85 (1): 30–34. https://doi.org/10.1177/0022242920970133.

Simonov, Andrey, Tommaso Valletti, and Andre Veiga. 2024. "Attention Spillovers from News to Ads: Evidence from an Eye-Tracking Experiment." *Journal of Marketing Research* 0 (0): 00222437241256900. https://doi.org/10.1177/00222437241256900.