More and Scammier Ads: The Perils of YouTube's Ad Privacy Settings

Cat Mai
New York University

Bruno Coelho*
New York University

Julia Kieserman*
New York University

Lexie Matsumoto*
New York University

Kyle Spinelli*
New York University

Eric Yang*
SUNY Buffalo

Athanasios Andreou
New York University

Rachel Greenstadt
New York University

Tobias Lauinger
New York University

Damon McCoy
New York University

Abstract

When users disable online ad personalization, they might be anticipate seeing fewer ads that are "relevant" to them as a trade-off for more privacy. In this paper, we show that the tradeoff can go much further than this intuition. We conducted controlled experiments on YouTube in Australia, Canada, Ireland, the United Kingdom, and the United States to investigate the impact of disabling ad personalization on the quantity and quality of ads that users receive. Through experiments where emulated users with different ad privacy settings watched sequences of 400 videos, we show that disabling ad personalization can lead to the user being shown as much as 1.30 times more pre-roll ads than the default (least private) setting. More concerning is that in our experiments, the proportion of predatory ads increased 2.69 times compared to the default setting, from 2.5 % to 8.7 % of ads. This result highlights that certain user demographics (in this case, privacy-conscious users) can be exposed to significantly higher rates of predatory ads, and suggests that the platform's efforts to curb such ads are still falling short.

Keywords

Targeted advertising, privacy dashboard, Google, ad settings

1 Introduction

Online platforms such as YouTube are offering privacy controls that allow users to manage how their personal data influence the ads they are shown. For example, users can disable ad personalization to avoid invasive targeting, and they can disable activity saving (e.g., the search or watch history) to prevent tracking. While changing these settings may improve privacy, platforms sometimes caution users that doing so might result in "less relevant ads" [29, 48]. However, there is reason to suspect that the implications could go beyond topic relevance, as high-quality advertisers may target their ads using mechanisms or attributes that are not available when tracking or personalization is disabled. Thus, ads that privacy-conscious users receive may not only be less relevant, but potentially also of lower quality. Prior work discussing ad targeting practices has focused on trade-offs between privacy and utility [1, 2, 16, 25, 67], but it is still an open question whether there is also a trade-off between privacy and ad safety.

In this paper, we investigate whether YouTube's ad privacy settings affect both the quantity and quality of ads in terms of ad *safety*, i.e., whether privacy-conscious users are exposed to more ads in general, and more scam and predatory ads in particular. Specifically, we investigate the following research questions:

- Does changing ad privacy settings affect ad load, i.e. lead to the user being shown more or fewer ads?
- Does changing ad privacy settings impact the rate of predatory ads, i.e. lead to the user being shown a higher or lower proportion of scam and predatory ads?

To investigate these questions, we run experiments emulating YouTube users watching sequences of 400 videos with three different ad privacy settings in parallel, extracting and manually annotating the pre-roll ads shown to them. Our experimental design controls for time, user location, and the type of video being watched as potential confounding factors. In total, our main data set consists of 450 watch sequences, amounting to 150 data points per privacy setting. Each watch sequence used one of five video types: Conspiracy, Popular, News, Kids, and Science. Our experiments ran in five vantage points: Sydney (Australia), Toronto (Canada), Dublin (Ireland), London (the United Kingdom), and Oregon (the United States). The user persona of our main experiment was fixed as a 34-year-old man. To address gender as an additional potential confounding factor, we conducted complementary smaller-scale experiments with a female persona to validate our findings on ad load and predatory ad rate. Overall, we manually annotated 10,628 unique pre-roll ads (corresponding to 38,851 impressions) and found 602 of them (2,610 impressions) to be predatory. Our code and data set of labeled ads is available at https://github.com/CybersecurityForDemocracy/youtube-ad-settings.

Our findings show that enabling stronger ad privacy settings on YouTube leads to users being shown an average of 1.30 times more pre-roll ads. More alarmingly, the share of ads labeled as predatory increases by an average of 2.69 times, from 2.5 % to 8.7 % of ads. A significant increase in predatory ads is observed across most countries and video types, with the exception of Canada and *Kids* videos, where the increase is positive but not statistically significant.

[*] These authors contributed equally to this study.

Corresponding author: cat.mai@nyu.edu.

Proceedings on Privacy Enhancing Technologies 2025(4), 1014–1038

https://doi.org/10.56553/popets-2025-0169

That is, privacy-conscious users are not only "penalized" by being shown more ads, but these ads are also more dangerous on average.

This result highlights that certain user demographics (in this case, privacy-conscious users) can be exposed to significantly higher rates of predatory ads, and suggests that the platform's efforts to curb such ads are still falling short. As such, our results contribute to quantifying an unexpected risk in enabling privacy controls, and raise the question of how they can be implemented more safely.

2 Background & Related Work

YouTube is a US-based video-sharing platform owned by Google. We use YouTube and Google interchangeably to denote the service provider. YouTube is the largest video-sharing platform by active users [22], and Google is the largest publisher by ad revenue [28]. YouTube primarily earns revenue from delivering online ads.[1] In the following, we provide a simplified overview of YouTube's ad system, along with the ad privacy options available to users.

2.1 Ad System

The primary function of the ad system is to deliver ads from advertisers to relevant audiences. Similar to other large platforms, YouTube's advertising model is programmatic—that is, ads are delivered dynamically. When a user initiates a video view, Google's ad load system determines whether to show an ad. In this paper, we focus on pre-roll ads since these are the most prominent ads on YouTube. Pre-roll ads play automatically before the actual video the user wishes to watch. They typically range from 5 seconds to several minutes long and are usually non-skippable for 5 seconds, after which viewers can skip the ad and proceed to the video.

When the ad load system determines that an ad should be shown to the user, an "auction" takes place to select the ad. The outcome of this ad selection process depends on several factors, which broadly include (1) how much an advertiser is willing to bid (pay) to reach a certain audience, (2) how the advertiser specified the desired audience (i.e., by which criteria the ad is targeted), and (3) ad delivery optimization based on criteria such as ad quality and how "relevant" the ad system estimates the ad to be for a given user (for example, the likelihood of the user interacting with the ad). These estimates likely become more accurate as more detailed information about the user becomes available to the ad system.

The bidding component of the ad system implies that advertisers assign different values to different audiences, often based on factors such as the audiences' estimated purchase power or likelihood of making a purchase. For example, toy advertisers may bid more for audiences who are inferred to be parents. Thus, a secondary implication is that users with little information available to the ad system are less likely to be linked to a high-value audience, and thus more likely to receive low-bid ads.

Advertisers can describe the desired target audience of an ad through two primary targeting mechanisms: contextual targeting and behavioral targeting. *Contextual targeting* involves the context of the content that the user is currently interacting with; in our study, the content is exclusively videos. Specifically, on YouTube, contextual targeting criteria include the time of day, the current video the user is watching, the user's general location, and the current search terms. In other words, contextual targeting does not utilize the user's past behaviors. *Behavioral targeting*, on the other hand, relies on the user's behavioral data, which include user demographics (age and gender) and historical data. Historical data include past search queries on Google, watch history on YouTube, location history in Maps, device type, and browsing IP addresses from Chrome, etc. This data also potentially includes past interactions with an advertiser, such as visits to their page, purchase history, ad clicks, etc. Based on user historical data, Google infers attributes about the user, such as interests, income range, parental status, relationship status, and location. Advertisers can also supply custom lists of user identifiers as the target audience of an ad.

Complementing (and potentially overriding some of) the advertiser-defined bids and targeting criteria, the ad system "optimizes" delivery based on the ad content and information available about users (such as contextual or behavioral user data). For example, prior work by Ali et al. showed that Facebook selected different audiences based on the ad image [4]. Delivery optimization attempts to maximize advertiser-specified goals such as ad clicks or purchases, while also minimizing the potential negative consequences of showing an ad that might upset a user and cause them to reduce their app use. This process can lead to a final audience that differs from what the advertisers originally intended [3–5, 8, 20, 35, 36, 38, 39].

2.2 Ad Privacy Settings

By default, YouTube collects all the aforementioned data about users while they are logged in and uses it to deliver personalized ads. Google allows users to configure stronger ad privacy settings,[2] which disable some forms of data collection and ad delivery [31]. Privacy settings and ad preferences configured for a Google account are applied across Google services, including YouTube. YouTube does not have separate privacy settings. The two settings that materially affect the type of information that can be used to target or deliver ads are ad personalization and activity history.

Web & App Activity stores activity from Google services, e.g., YouTube watch history or past Google searches. When disabled, advertisers can no longer target (and the ad system can no longer select ads) based on such historical data or its inferred attributes. Advertisers can still target (and ads can still be selected) based on user demographics (age and gender) and contextual data such as the current video being watched, or the inclusion of the user on an advertiser-provided user list.

When a user disables the *Ad Personalization* setting, Google can no longer use signals related to the individual user to deliver ads (i.e., no behavioral data). The ad delivery system can still use contextual data to select ads, such as the time of day, the general location of the user, the current search terms, or the video being watched.[3] In the context of our experiments, we reiterate which user data can be used for targeting or ad delivery for each privacy setting in Section 3.2.2.

[1] YouTube also sells subscriptions, but they generate a negligible amount of revenue compared to online ads [6].

[2] While the use of the term "privacy" in the context of ad personalization is debatable, we follow Google's terminology of "ad privacy settings" for descriptive purposes.

[3] As an important nuance, we note that our results suggest that advertisers cannot directly target users who have disabled ad personalization based on the current video being watched, while the ad delivery system still appears to use this contextual signal to select ads for these users.

2.3 Ad Targeting Transparency

Ad targeting transparency, or ad explanations, are information snippets that accompany all ads on YouTube. They provide users with information about why the selected ad is being shown to them. On YouTube, these explanations contain the targeting category selected by the advertiser (e.g., "your age"), but not the actual values (not "users between 20 and 35"). Some platforms offer ad preference managers, which allow users to see what the platform has inferred about their demographics and interests [11]. On Google, this ad preference manager (called My Ad Center) is only available when ad personalization is enabled. Users can view past ad topics and demographic inferences, but their full interest profile remains hidden from them—all the while advertisers can use it for targeting [12, 24, 59]. This information asymmetry is also observed in other platforms [7, 69, 70].

2.4 Related Work

Our research aligns with recent endeavors in algorithm audits [10] to measure the personalization effect of content delivery, i.e. disparities in the distribution of video recommendations [41, 42, 56], search results [34], etc., across various demographics or contexts. In our study, the focus is on ads.

Our study adds to an expanding body of work that investigates the prevalence and distribution of problematic ads on social media. Audits of advertising on social media have covered a range of topics, such as measuring policy enforcement by the platform (e.g., for political advertising [40]), creating a taxonomy for problematic ads [72], as well as measuring how the prevalence of these ads varies across various demographics [3]. Researchers have studied "bad ads" of various kinds, ranging from blatant fraud schemes [3, 9, 37, 49] and malware distribution [44, 52, 58, 63, 64, 71] to ads that are more nuanced in their harms towards users, including privacy vulnerabilities [14, 53, 68], spreading misinformation and clickbait content [3, 27, 74], sensitive or triggering topics [26, 55, 60], or ads in unregulated industries, such as cryptocurrency [43, 57].

Literature discussing modern ad targeting practices focuses on utility and user privacy tradeoffs [1, 2, 16, 25, 67]. More precise targeting can mean more relevant and higher-quality ads for users, as well as more efficient campaigns for advertisers, at the expense of user privacy through large-scale data collection that enables these targeting tools. For scams, there is no utility–privacy tradeoff as they unequivocally harm users. To the best of our knowledge, there is no prior work measuring exposure to scam ads as a function of privacy settings.

Closest to ours but in a different domain is a study by Spinelli and Crovella measuring the impact of privacy settings on video recommendations on YouTube [61], showing that YouTube is recommending privacy-seeking users videos from less reliable sources. Video recommendations and ads are two distinct systems with different objectives and constraints. While video recommendations largely maximize user time spent on the platform and are exclusively determined by Google [19], ad delivery combines both the advertiser's constraints (e.g., budget and targeting criteria) and Google's ad delivery optimization algorithm.

Lastly, there is a line of work estimating or measuring the value of users (or information about them) to online advertisers [13, 14, 54, 73]. Privacy-conscious users are arguably less valuable to advertisers because their interests are unknown [1]; they are less likely to take action based on a randomly selected ad than a user being shown a personalized, highly "relevant" ad [45]. While not directly related to our work in their goals, these insights may help explain the phenomenon we observe in our experiments, i.e., when users with stronger ad privacy settings and less information known about them are of lesser value to advertisers, then they likely attract lower ad bids and thus lower-quality ads.

3 Methodology

We aim to test our hypotheses that disabling ad personalization on YouTube leads to the user being shown more ads, and those ads are also "scammier." To this end, we define the following metrics:

- The *ad load* is the average number of ads shown per video watched.
- The *predatory ad rate* is the proportion of predatory ads out of all ads served.

We use a binary assessment as to whether an ad is considered predatory or not (see Section 3.3). We measure differences in the ad metrics above to quantify the relative impact of three ad privacy settings. We furthermore test if this result holds across several dimensions, notably the viewer's location, the time of viewing, and the type of videos watched.

At a high level, our methodology consists of running controlled experiments with sock puppet accounts that emulate users watching YouTube videos in an instrumented browser. The basic building block of our experiments is a virtual machine with three browser instances, each with a different ad privacy setting, that watch the same experiment-specific list of 400 YouTube videos in parallel. We use this setup in order to reduce possible confounding factors when assessing differences in the above ad metrics. To measure the impact of location, we run five such virtual machines simultaneously with local IP addresses in Sydney (Australia), Toronto (Canada), Dublin (Ireland), London (the United Kingdom), and Oregon (the United States). We refer to this ensemble of experiments (i.e., consecutive repetitions with the different video types from different locations with three privacy settings each) as one run of the experiment. To mitigate the impact of time, we repeat each experiment six times with at least one week between consecutive runs.

3.1 Data Collection

We collected data for the main experiments over 8 weeks, from November 6, 2024 until January 17, 2025.

3.1.1 Input Videos. Our goal is to test whether our hypotheses (that stronger privacy settings lead to more ads and "scammier" ads) hold regardless of the type of video (or location). We aim to cover sufficiently different video categories without being exhaustive. We compiled five disjoint lists of YouTube videos: News, Science, Popular, Kids, and Conspiracy videos.

The *News* video list was compiled from Media Bias/Fact Check [46], an independent news rating agency, using data from September 9, 2022. Unlike the other video sets in our study, each country has its own *News* video list. We selected news sources classified as "Left-Center," "Right-Center," and "Least Biased" with high or very high factual reporting scores, along with high credibility. We excluded news sources without an active YouTube channel, determined by having fewer than 100 uploaded videos. For the US and Canada, where many sources meet our criteria, we randomly selected 50 active channels per country. In the other countries, we retained all active channels (12 in Australia, 4 in Ireland, and 28 in the UK). We then used the 100 most popular videos from each channel and shuffled the ordering to compile the final list.

The *Science* list is comprised of the most popular videos from a list of 138 manually curated science channels. We selected these channels using an informal depth-first search of relevant channels following public subscription relationships between channels, starting with Numberphile and Mark Rober. The number of videos per channel in our Science list roughly follows a normal distribution.

The *Popular* list contains 1,714 most viewed videos on YouTube as of March 14, 2024. We collected these videos through an auto-updating channel that indexes most-viewed videos on YouTube [51].

The *Kids* list contains 1,905 children-focused videos that were collected from the YouTube Kids app in November 2023. When compiling this list (and also the four other video type lists), we excluded content labeled as "made for kids" by YouTube on the main app as of November 2024. To detect videos "made for kids," we used the redirection banner to the YouTube Kids app that YouTube places under these videos in the main app. The rationale for excluding videos "made for kids" is that YouTube does not serve personalized ads on these videos [32] even when the account is configured to allow ad personalization, which would cause anomalies in our experiments. The Kids video list thus consists of thematically child-directed videos that are not explicitly labeled as "made for kids."

The *Conspiracy* video list is based on a prior study by Faddoul et al. [23]. From the original 6,752 videos, we removed 3,665 that were unavailable as of March 20, 2024.

In our experiment, we randomly sampled 400 videos from the corresponding video set to serve as the input to the crawler. Input videos were subsampled from the larger set to avoid using the same 400 videos across 6 runs, reducing potential biases caused by specific video selection, as well as mitigating video ordering effects.

3.1.2 Ad Collection Process. We extracted ads from YouTube using an instrumented browser inside virtual machines running Windows Server 2022. Specifically, we used Selenium with Google Chrome in headful mode (with the graphic user interface rendered). We collected pre-roll ads along with data from the "My Ad Center" iframe that accompanies all ads on YouTube. From this iframe, we collected the ad targeting criteria provided by Google, as well as the advertiser's name and location of registration. Per video, there can be 0–2 pre-roll ads; the crawler watched the ads up to a timeout of 10 minutes per ad. The crawler saved the ad creative—including the ad video URL and any accompanying text—as well as all associated data in the iframe, but did not interact with the ads directly to avoid potentially sending unintended signals. Whether or not the pre-roll ads were present, the crawler watched each input video for 30 seconds before moving on to the next one on the list.

	Experiment configurations
Location	Sydney (Australia), Toronto (Canada), Dublin (Ireland), London (the UK), and Oregon (the US)
Video list	News, Conspiracy, Science, Popular, Kids
Watch time per video	30 seconds
Videos per watch sequence	400
Randomized video sequence	Yes
Total watch sequences (Main experiment)	3 privacy settings × 5 video sets × 5 locations × 1 gender (Male) × 6 repeats = 450 watch sequences
Total watch sequences (Validation)	3 privacy settings × 3 video sets × 1 location × 1 gender (Female) × 4 repeats = 36 watch sequences

Table 1: Summary of the experiment configurations.

3.2 Experiments

Table 1 summarizes the configurations of our experiments.

3.2.1 Sock Puppet Personas. In order to create an account, Google requires users to provide their name, gender, date of birth, and a phone number for verification purposes. These account attributes are potential confounding factors in our experiments (e.g., results we observe in experiments with male accounts may not necessarily replicate with female accounts). Ideally, we would repeat our experiments with all combinations of account parameters (age, gender, location, etc.), but this is not a realistic option given the costs of running experiments. Instead, we varied a subset of parameters that we hypothesized could influence the delivery of predatory ads, namely, video type, location, and time. The remaining required account attributes were held constant in the experiments. Regarding gender, preliminary results indicated that female personas might exhibit slightly larger differences—both in ad load and predatory ad rate—between the ad privacy settings, thus we selected a male persona as a conservative experimental design choice: Any significant effects for the male persona would likely hold as well—or even be stronger—for a female persona. To validate that our findings are not entirely dependent on the chosen persona, we conducted a smaller-scale validation experiment with a female persona, detailed in Section 3.2.3. The remaining required parameters for account creation (i.e., name, date of birth, phone number) were selected arbitrarily. In summary, the main experiment uses Google accounts with identical name (John Green), gender (Male), date of birth (1/1/1990, making our account 34 years old during the study), and unique phone numbers (with +1 country code) for verification.

While the age and gender of the persona were fixed, we varied the location across five countries. As an indication that our accounts were properly localized, our non-US accounts observed the corresponding country code next to the YouTube logo. Additionally, our Irish and UK accounts had access to a special setting to link Google services,[4] a feature unavailable elsewhere. Although verifying the age of a Google account is optional, we used YouTube's system to proactively verify that all our accounts were over 18 years old before using them in experiments. This was necessary to enable

[4] https://myactivity.google.com/linked-services

personalized ads and view age-restricted content. For age verification, we used a mobile VPN and selfie verification in Australia, Ireland, and the UK, while in the US and Canada, it was enough to confirm the age by clicking on a button. Each experiment instance, represented by a browser instance, used a separate Google account. Within any active VM, three browser instances watched the same video list simultaneously, each with different ad privacy setting. After each iteration (which we call a *watch sequence*), we deleted all accounts' activity history. Srba et al. observed that there were no significant carry-over effects from similar account teardown processes [62]. We waited at least 12 hours to avoid bot detection and for the activity deletion to take effect before repeating the process for the next video list, although with randomized ad privacy setting assignments to minimize any potential account-specific effects.

In our preliminary experiments, we noticed that two out of the 60 accounts showed unusually high pre-roll ad loads, deviating significantly from the typical distribution. To address this, we implemented ad load tests for all accounts after creation. These tests involved each account watching 50 randomly selected videos from the *Popular* video list for 60 seconds. During this test, we enabled both ad personalization and using activity for ads. We removed accounts with pre-roll ad loads outside 2.5 standard deviations of their country-specific mean.[5] We also removed accounts that did not receive any ads.[6] Since most accounts produced approximately normal ad loads, these outlier accounts would introduce unwanted noise into our measurements.

3.2.2 Privacy Settings Experiments. For each watch sequence, the crawler sequentially viewed 400 videos in its predetermined ad privacy setting, watching 30 seconds of a video before moving on to the next available video. While Google does not have any detailed documentation on how they register a view, evidence suggests a watch time of 30 seconds is sufficient [15, 66].

With two activity settings (saving activity on and off) and two ad personalization settings (ad personalization on and off), there are four possible configurations. According to Google, turning off ad personalization means "your info won't be used to personalize ads, including: Your new and existing activity on Google sites and apps, including your general area while using them ..." [30]. We interpreted this explanation to mean that saving activity on and ad personalization off would have the same effect as configuring both off and only used the latter in the experiments. Datta et al. showed that Google complied with user preferences when users opted out of personalized ads [20]. Therefore, we tested for three configurations of ad privacy settings, which were activity on/personalization on, activity off/personalization on, and both off.

Within the constraints of our study design, the criteria that YouTube could always use to select ads were the general location based on IP address and the time of day. For ad personalization on and activity history off, the criteria additionally include the user demographics from the account (age and gender) and the current video being watched. For ad personalization on and activity history on, in addition to all the factors above, Google may also use YouTube watch history and interests or categories that Google may have inferred from the watch history, such as income, education level, or homeownership status.

3.2.3 Persona Validation Experiments. Since our main experiments were conducted using only male personas, we performed a validation test to examine whether the observed increases in ad load and ad predatoriness for personalization-disabled settings also hold for a female persona. Data for these validation experiments was collected from May 2–14, 2025. We replicated the account setup process described in Section 3.2.1, changing only the gender to female and the name to Jane Doe. Due to resource constraints, we performed the validation experiments with female personas in a more limited setup. Specifically, we focused only on the United States and the *Conspiracy*, *News*, and *Science* video sets, which showed the largest effect sizes in our main experiments. This setup was repeated 4 times, with each run spread at least 24 hours apart. In total, the validation covers 3 video sets × 1 location × 4 repeats.

3.3 Labeling of Ads

We define predatory (or scam) ads broadly as ads with malicious intent to manipulate the user into unfavorable market transactions, either through deception or undisclosed information. This is not necessarily dictated by the law, although many scam ads can be illegal. Scams in our study include, but are not limited to, ads that intentionally mislead users about the quality or features of the advertised products (e.g. get-rich-quick schemes, weight loss supplements) or deceive users into taking actions against their best interests (e.g. fake software updates, phishing attempts, predatory loans). Legitimate ads can exhibit a certain degree of deception; what characterizes scams in our definition is the disproportionate harm to the user, whereas in legitimate businesses, both the business and the user will benefit from the transaction [25]. While our interpretation is not narrowly defined in order to capture a wide range of scam ads, we exercised caution and adopted a conservative approach in our labeling to avoid false positives. In cases with uncertainty, we labeled the ads as non-scam. As an illustration, Figure 1 shows the landing pages of two predatory ads.

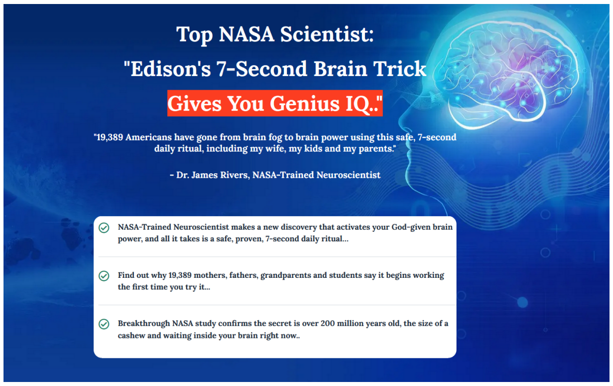
3.3.1 Inter-rater Agreement and Codebook. To assess the reliability of our predatory ad labels, we iteratively refined our codebook and measured inter-rater agreement. Appendix A and B detail our labeling practices, the codebook, and additional example ads.

We present a set of "red flags" observed in our experiments that may signal predatory intent. It is not within the scope of our study to be exhaustive about categorizing all possible predatory ads.

- **Ads in sectors that are controversial or prone to be predatory**, such as subprime credits and loans, speculative or unregulated industries e.g. cryptocurrency, foreign exchange market (forex), fortune-telling, alternative medicine etc.
- **Ads advertising businesses with strong evidence of malicious or deceptive practices** that we can verify through user reviews or Google searches of the business.
- **Ads advertising activities that are illegal or promote dishonest behaviors** confirmed through reviews or web searches.
- **Ads that use problematic or manipulative patterns** such as clickbait (e.g., promises of free products, attention-grabbing yet misleading thumbnails), misinformation, exaggerated language.

[5] We hypothesize that these accounts with abnormally high ad loads were part of an A/B test of elasticity.

[6] We hypothesize that these accounts might have been detected as inauthentic.



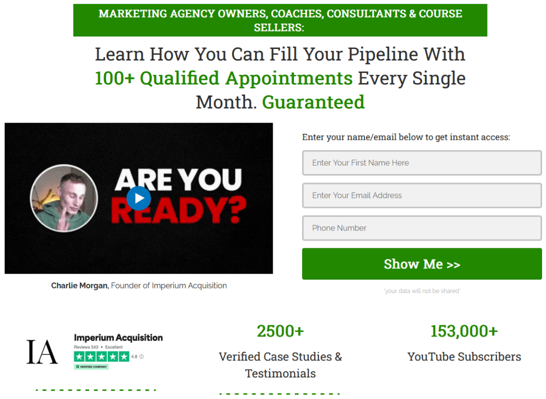
Top NASA Scientist:
"Edison's 7-Second Brain Trick
Gives You Genius IQ.."

"19,389 Americans have gone from brain fog to brain power using this safe, 7-second daily ritual, including my wife, my kids and my parents."

- Dr. James Rivers, NASA-Trained Neuroscientist

- ✓ NASA-Trained Neuroscientist makes a new discovery that activates your God-given brain power, and all it takes is a safe, proven, 7-second daily ritual...
- ✓ Find out why 19,389 mothers, fathers, grandparents and students say it begins working the first time you try it...
- ✓ Breakthrough NASA study confirms the secret is over 200 million years old, the size of a cashew and waiting inside your brain right now..

(a) Homepage for Purpose Pathway. The site advertises a 7-second ritual, supposedly confirmed by NASA, to activate your brain power. The site claims that roughly 20 thousand people have confirmed the efficacy upon their first attempt. We labeled this ad as predatory due to the unverifiable and scientifically baseless claims, such as “NASA-confirmed,” and likely fabricated testimony of 20 thousand users. The ad video on YouTube: https://www.youtube.com/watch?v=QMmidNe6R7M



MARKETING AGENCY OWNERS, COACHES, CONSULTANTS & COURSE SELLERS:

Learn How You Can Fill Your Pipeline With 100+ Qualified Appointments Every Single Month. Guaranteed

Charlie Morgan, Founder of Imperium Acquisition

Enter your name/email below to get instant access:

Enter Your First Name Here

Enter Your Email Address

Phone Number

'your data will not be shared'

2500+
Verified Case Studies & Testimonials

153,000+
YouTube Subscribers

(b) Homepage for Imperium Acquisition. The site boasts a program to transform users’ marketing business, specifically by finding them a “guaranteed” 100 qualified appointments per month. We labeled this as predatory due to the highly unlikely claim of “guaranteed” outcomes. The ad video on YouTube: https://www.youtube.com/watch?v=svNfcpurUmI

Figure 1: The landing pages of two example ads that were labeled as predatory.

An ad may exhibit one or more red flags without necessarily being labeled as predatory. This is either because after researching the advertiser, we found no compelling evidence for predatory practices, or because of the distinctions listed in Appendix B.

Three graduate students co-authors annotated the ads using the provided codebook. To quantify the annotation task’s difficulty, we measured inter-rater agreement on 50 randomly sampled pre-roll ads labeled by all three annotators. After two rounds of annotation, which included discussing disagreements and clarifying the codebook, the annotators reached a Fleiss kappa score of $\kappa = 0.74$ on a new 50 ad sample, indicating strong agreement.

All ads from our experiments received two independent annotations. In case of a tie, the ad was discussed with other annotators to reach a consensus. We deduplicated ad videos based on their video ID. Ads were combined across the three privacy settings before being distributed so that potential rater bias was less likely to affect the result. Raters were not informed of the actual video that the ads were shown on or the privacy settings that produced the ads to further limit any potential bias. Raters had access to the ad video and if available, the ad text (could be a link text or a call-to-action text, e.g., “Click here”), and the advertiser name and location. Raters were instructed to visit the advertiser’s website and perform a Google search for consumer reviews on sites such as Better Business Bureau, Reddit, or Trustpilot.

3.3.2 Label Propagation. After preliminary data labeling, we observed that legitimate advertisers almost exclusively published non-scam ads. To reduce the annotation workload, we propagated non-scam labels to the rest of an advertiser’s ads if an advertiser had at least 20 unique non-scam ads and no scam ads in our data set. To validate this approach, we manually labeled a batch of randomly sampled 500 ads and cross-checked our labels against the propagated labels with perfect agreement between the two.

	Pre-roll Ads	
	total (unique)	predatory (unique)
Run 1	5,578 (2,669)	307 (137)
Run 2	5,818 (2,259)	433 (152)
Run 3	5,748 (2,246)	372 (126)
Run 4	5,717 (2,339)	355 (113)
Run 5	5,879 (2,204)	497 (156)
Run 6	5,855 (2,323)	499 (147)
Persona Validation	4,256 (1,572)	147 (60)
Total	38,851 (10,628)	2,610 (602)

Table 2: Counts of pre-roll ad impressions in the six experiment runs. Unique ads (in parentheses) do not add up to the total because ads can re-appear across runs. Overall, 2,610 out of 38,851 ad impressions were classified as predatory.

3.4 Data Set Overview

Table 2 shows the total number of pre-roll ads, broken down by type, collected by the crawler in the six experiment runs. (Tables 10, 11, 12, 13, and 14 in the appendix contain the full data set broken down by privacy setting, time, and country for each of the five video types.) We identified 2,610 pre-roll impressions as predatory, accounting for 6.7 % of the total pre-roll impressions.

Among 38,851 ad impressions seen in our experiments, 21 did not have any listed reason in the targeting disclosures, either by data collection error or by Google’s design. We did not exclude these ad impressions from analysis since they would not materially impact the overall results. According to Google, some of the targeting disclosures are left blank because the ads are not personalized for viewers who might be under 18 years of age (an example is provided in Figure 3 in the appendix). Despite these instances of unintentional non-personalization, our accounts were still receiving personalized ad impressions otherwise. These unintentionally non-personalized ads appeared on videos that, upon manual inspection, seemed to be child-directed but were not explicitly labeled as content made for kids, similarly to what was observed in the study by Medjkoune et al. on YouTube ads for children [47]. For the ads that did have targeting disclosures, we grouped similar disclosures for ease of analysis and presentation. Specifically, we grouped the targeting reasons “The time of day” and “Your general location (like your country or city)” or non-English versions of these phrases together with the combined explanation “The time of day or your general location (like your country or city).”

3.5 Analysis Methodology

In our analysis, we aim to test several null hypotheses that are variations of *there is no difference in the ad load and predatory ad rate, respectively, between the three ad privacy settings.* Our data set consists of 450 watch sequences, amounting to 150 data points per privacy setting (see Table 1). Each data point in the test is either the ad load or predatory ad rate from one watch sequence (which represents a specific video type, location, experiment run, and privacy setting). We test our hypotheses on aggregated data (150 data points per setting), as well as on country-level and video type-level data (30 data points per setting). We use Holm-Bonferroni correction when performing multiple comparisons.

For our analysis, we consider privacy settings as “treatments” and aim to find whether different treatments have significantly different effects. Since we consider our accounts interchangeable within the same country, applying different treatments to the same subject is considered a repeated measurement experiment.

As the first step, we need to determine the appropriate test. For the hypothesis on ad load, we use one-way repeated measures ANOVA as the omnibus test since the ad load data follow a normal distribution. ANOVA tests for differences in three (or more) groups in a repeated measurement experiment like ours. Where we find significant results, we use Tukey’s HSD test with Holm-Bonferroni correction. This allows us to identify which pair of privacy settings exhibits significant differences.

For the hypothesis on predatory ad rates, since the rates do not approximately follow a normal distribution, we use Friedman test as the omnibus test, which is a non-parametric alternative to ANOVA. Where we find significant results, we use the Conover post-hoc test with Holm-Bonferroni correction to find which pair of privacy settings exhibits significant differences.

We also perform regression analysis to investigate the effect sizes of the two stronger ad privacy settings compared to the default setting in terms of ad load and predatory ad rate. For each of these two metrics, we fit the models on aggregated data (150 data points per setting), as well as on country-level and video type-level data (30 data points per setting). We use logistic regression models, where the dependent variable is the metric, and the predictors are country, video type, and privacy setting. Specifically, for analyses on ad load, we use Poisson regression, which is used to model count data—the input to the model is the count of pre-roll ads. For analyses on predatory ads, we use negative binomial regression, which is the alternative to Poisson regression for data with high variances—the input to the model is the count of predatory pre-roll ads. To normalize the count of predatory ads, we add an offset term that is the log of the total ad count. When we report the results, we report the expected rate ratios instead of the coefficients from the model. The rate ratio is calculated as e^{β}, where β is the corresponding coefficient from the model estimations, along with a 95 % confidence interval for these ratios. Along with the rate ratios, we also report Cox-Snell’s pseudo-R^2—a measure of relative fit, which ranges from 0 to 1. It is important to note that pseudo-R^2 should not be interpreted the same way as R^2 in linear regression; while it offers an approximation of goodness of fit, it does not directly quantify the proportion of variance explained by the model [33].

3.5.1 Persona Validation Experiments. We complement our main experiments (which all used a male persona) with a set of validation experiments with a female persona. Our null hypotheses are similar in nature to the main experiments: for a female persona, there is no difference in the ad load and predatory ad rate, respectively, between the three ad privacy settings. Identical to the main experiment analysis methodology, for the hypothesis on ad load, we use one-way repeated measures ANOVA with Tukey’s HSD test with Holm-Bonferroni correction. For the hypothesis on predatory ad rates, we use the Friedman test with the Conover post-hoc test with Holm-Bonferroni correction. Since this experiment is designed to validate the main findings of the male personas and due to the smaller validation data set, we do not perform statistical tests on data broken down by video sets or countries.

4 Analysis

Since our overall hypotheses are that stricter ad privacy settings lead to an increase in the quantity of pre-roll ads shown while decreasing their quality, we will first examine the impact of ad privacy settings on ad load, followed by an analysis of the predatory ad rate. For both quantity and quality analysis, we investigate our research questions using both the aggregated data and the data broken down by country and by video type. To analyze pre-roll ad loads, we use statistical methods that assume normal data distribution, whereas for pre-roll ad predatoriness, we use methods suited for non-normal data. We include the output for all the statistical tests we used in Appendix C.

4.1 Pre-roll Ad Load

First, we ask whether changing the ad privacy settings has a significant impact on *pre-roll ad load*, i.e., the number of pre-roll ads shown per video watched. Our null hypothesis $H_{0\text{–ad load}}$ is: *there is no difference in the pre-roll ad load between the three privacy settings.*

To assess differences in the pre-roll ad load between the three ad privacy settings, we conducted a one-way repeated measures ANOVA with the watch sequences grouped into the three ad privacy settings, i.e. 150 watch sequences per privacy group. The omnibus test revealed significant differences between pre-roll ad load for the three ad privacy settings ($F(2, 298) = 172.47, p < 10^{-50}$, $\eta^2 = 0.212$), i.e. we reject $H_{0\text{–ad load}}$. According to Cohen’s guidelines [18], $\eta^2 = 0.212$ represents a large effect size, meaning that most of the total variance in the pre-roll ad load is explained by the differences across the three ad privacy settings. Furthermore, a tiny residual error ($MSE = 0.0006$) in our test indicates minimal

	Pers. On/Act. Off		Pers. Off/Act. Off			
	Ratio	95% CI	Ratio	95% CI	R^2_{CS}	n
News	1.29	[1.20, 1.38]	1.25	[1.17, 1.34]	0.98	30
Popular	1.23	[1.16, 1.30]	1.23	[1.16, 1.31]	0.65	30
Science	1.29	[1.22, 1.37]	1.34	[1.27, 1.42]	0.77	30
Kids	1.27	[1.20, 1.34]	1.28	[1.22, 1.35]	0.79	30
Consp.	1.32	[1.25, 1.40]	1.36	[1.28, 1.44]	0.82	30
AU	1.29	[1.22, 1.37]	1.29	[1.22, 1.37]	0.85	30
CA	1.30	[1.22, 1.38]	1.33	[1.25, 1.41]	0.88	30
IE	1.14	[1.07, 1.21]	1.17	[1.10, 1.25]	0.99	30
UK	1.32	[1.25, 1.40]	1.33	[1.25, 1.41]	0.93	30
US	1.34	[1.26, 1.42]	1.35	[1.28, 1.43]	0.97	30
All	1.28	[1.24, 1.31]	1.30	[1.26, 1.33]	0.96	150

Table 3: Pre-roll Ad Load: Ratios of pre-roll ad load for stronger ad privacy settings relative to the default setting. Stronger privacy settings are personalization on & activity saving off, and personalization off & activity saving off, respectively. These are results from fitting a Poisson model where the dependent variable is the ad load, and the predictors are country, video type, and privacy setting, with n watch sequences per privacy setting. The table includes the ad load ratios using the default setting as the baseline, along with a 95 % confidence interval for the ratios and pseudo-R^2 (a measure of relative fit, ranging from 0 to 1). Ratios are significant at $p = 0.001$. Disabling ad personalization led to an average of 1.30 times more pre-roll ads than the default privacy setting.

variability within the ad privacy setting, reflecting high consistency across our measurements.

Post-hoc pairwise comparisons using Tukey's HSD test with Bonferroni correction showed that the default (least private) ad privacy setting had the *lowest* pre-roll ad load, significantly lower than the other two more private settings ($p < 10^{-30}$). However, the two stronger privacy settings did not differ significantly from each other ($p > 0.01$). Pairwise comparisons with the default (least private) privacy setting also revealed large effect sizes (Cohen's $d > 1.12$), meaning that the average pre-roll ad load of the default setting is at least 1.1 standard deviations lower than the other two settings. In conclusion, when users disable either using activity for ads or ad personalization (i.e., choose a stronger ad privacy setting than the default), they will be shown significantly more pre-roll ads. We did not find a significant difference between these two settings, i.e. in terms of ad load, just disabling the use of activity for ads is no "middle ground" between ad personalization on and off.

To make these test results more concrete, Table 3 quantifies the increase in ad load between the two stronger settings versus the default setting. Overall, both disabling the use of activity for ads and disabling ad personalization altogether led to an average of 1.30 times more pre-roll ads than the default privacy setting.

4.1.1 Persona Validation Experiments. To validate whether the association observed in the previous subsection—stronger privacy settings lead to higher ad loads—also holds for a female persona, we analyzed the validation experiment using a one-way repeated measures ANOVA (12 watch sequences per privacy setting). The omnibus test revealed significant differences between pre-roll ad load for the three privacy settings ($F(2, 22) = 47.67$, $p < 10^{-7}$, $\eta^2 = 0.40$), i.e. we reject $H_{0\text{–ad load}}$ for the female data. Pairwise comparisons ($p = 0.0005$) showed similar patterns to our main experiments: the two more private settings are significantly different from the default (least private) setting, while not being materially different from each other. The test calculated Cohen's $d > 1.63$, meaning that the average pre-roll ad load of the default setting is at least 1.63 standard deviations lower than the other two settings. This means that the ad load effect size in the female persona experiments is slightly larger than that of the main experiments using a male persona (where Cohen's $d > 1.12$).

4.1.2 Analysis by Country and Video Type. While the aggregated data from multiple countries and video types shows that enabling ad privacy settings leads to ad load increasing significantly, we also investigate whether this effect depends on the country or video type. To more formally determine whether the relationship between privacy setting and ad load remains significant across different data subsets, we tested $H_{0\text{–ad load}}$ within each country and video set. For these tests, we subset the data to include only watch sequences that match the respective criteria. For example, testing the hypothesis for Australia used only sequences from that country, covering all five video sets. Each analysis had 30 data points per privacy setting, as opposed to 150 data points in the aggregated analysis above. Similar to the aggregated analysis, we used one-way repeated measures ANOVA to test the overall differences, followed by Tukey's HSD method to find pairwise differences; all p-values corrected for multiple comparisons with Holm-Bonferroni method. We found that the association between stronger privacy settings and higher ad loads holds consistently across all countries and video sets ($p < 0.00001$). Pairwise comparisons revealed similar differences to the aggregated data: both stronger privacy settings differ significantly from the default (least private) setting but not from each other.

All the pairwise comparisons were significant at $p < 10^{-25}$ after adjusting for multiple comparisons, except for Ireland and the *News* video set. For these two subsets of data, there was a lack of significance in pairwise comparisons at $p = 0.01$, meaning that while the ad load distributions of three privacy settings differ significantly overall, it is less clear which specific pair(s) drive this difference. This aligns with our understanding of the data set, as watch sequences with either of these two attributes consistently produced the lowest ad load in all three privacy settings compared to other attributes. The low ad load in Ireland and *News* watch sequences may have limited our ability to detect significant pairwise differences between the privacy settings, despite a statistically significant overall effect. Another possible explanation for the *News* watch sequences is that some news channels may have opted out of monetization, thereby limiting the frequency of ads shown on their videos. On average, news channels had pre-roll ads on 5.4 % of their videos, though there were some outliers (since videos can repeat between the experiment runs, here we calculate the total video count without deduplicating). For instance, two major Irish news outlets—RTÉ (national broadcaster) and The Journal—had ads in fewer than 0.8 % of their videos. RTÉ, with 2,523 video views in our data set, had ads on only 21 of them; similarly, The Journal had just 8 ad-serving videos out of 2,106 video views, resulting in an ad-to-video ratio of less than 0.5 %. These outliers with much lower ad-to-video ratios than other channels are likely indications of (intentional) demonetization, where the channel opts out of earning from ads on their content while YouTube still displays (limited) ads. Regardless of the actual cause, the considerably lower ad load in Ireland and *News* data reduces the interpretability of pairwise comparisons in this subset, and should be interpreted with caution.

Table 3 further breaks down the ratios of mean pre-roll ad load by video type and country and shows that the increase in ad load for stronger privacy settings persists across these factors, even though the magnitude of this increase may slightly differ.

In all the subsets into which we spliced the data, users who opt for stronger privacy settings would consistently be shown more ads—30 % more ads on average compared to the default setting. This suggests that YouTube indirectly makes users “pay” for more privacy with more ads (a more disruptive experience).

4.2 Predatory Pre-roll Ad Rate

The previous subsection showed that privacy-conscious users are exposed to more advertising on YouTube, but it has remained unclear whether the ads differ in terms of quality. In this subsection, we utilize our annotations of pre-roll ads to quantify differences in the *predatory* ad rate, i.e., the proportion of pre-roll ads observed during a watch sequence that are labeled as predatory. Our null hypothesis $H_{0\text{–predatory ad rate}}$ is: *there is no difference in the predatory ad rate between the three ad privacy settings.*

To test for statistically significant differences in the predatory ad rate between the three ad privacy settings, we use the Friedman omnibus test with the watch sequences divided into three groups according to their privacy settings, i.e. $n = 150$ watch sequences per group. The test reveals significant differences between predatory ad rates for the three privacy settings ($\chi^2(2) = 68.47$, $p < 10^{-15}$), thus, we reject $H_{0\text{–predatory ad rate}}$. Conover’s post-hoc test with Holm correction shows that the predatory ad rate is the *lowest* with the default privacy setting ($p < 10^{-7}$), while the two stronger privacy settings result in higher predatory ad rates that are not statistically different from each other at the $p = 0.05$ level. In other words, stronger ad privacy settings lead not only to more ads but also to “scammier” ads, and similarly to the ad load, the two stronger privacy settings do not differ significantly from each other.

In addition to our statistical tests, Table 4 highlights that overall, the most private setting led to an average of 2.69 times more predatory ads than the default (least private) setting; while for the “middle” private setting (ad personalization enabled, activity disabled), this increase is 2.87 times. In conclusion, our experiments show that reducing or disabling the ad system’s ability to personalize ads led to a different composition of ads in terms of ad safety, i.e. an amplification of predatory ads.

4.2.1 Persona Validation Experiments. To assess whether the amplification of predatory ads also applies to a female persona, we applied the Friedman test to the validation experiment data (12 watch sequences per privacy setting). The test reveals statistically significant differences between predatory ad rates for the three privacy settings ($\chi^2(2) = 10.8$, $p < 0.005$), thus, we reject $H_{0\text{–predatory ad rate}}$ for the female persona data. Pairwise comparisons at $p = 0.05$ showed that the most private setting produced significantly more predatory ads than the other two settings. A subtle difference from the main experiments is that, for the male persona, the “middle” setting (ad personalization enabled, using activity for ads disabled) yields similar results to the most private setting. In contrast, for the female persona in the validation experiment, the “middle” setting is more similar to the least private (default) setting. Notwithstanding these differences, the validation experiment confirms that increased predatory ad rates due to the strongest ad privacy setting also occur for the female persona.

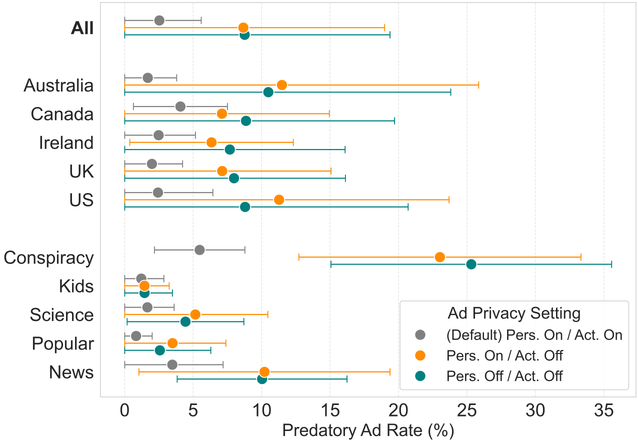


Figure 2: Predatory ad rates by ad privacy setting across countries and video categories. Each point shows the mean predatory ads for a given setting, with error bars indicating ±1 standard deviation.

4.2.2 Analysis by Country and Video Type. Aggregating data from all countries and all video sets, we have shown that stronger privacy settings are significantly associated with higher rates of “scammy” ads. To further investigate whether this association holds across different parameters, we tested $H_{0\text{–predatory ad rate}}$ within each country and video set.

For these tests, we subset the data to include only watch sequences matching the respective criteria. For example, testing the hypothesis for Australia used only sequences from that country, covering all five video sets. Each analysis had 30 data points per privacy setting, compared to 150 data points in the aggregated analysis above. Similar to the aggregate analysis, we used Friedman test to detect overall differences, and Conover post-hoc test for pairwise differences; all p-values corrected for multiple comparisons with Holm-Bonferroni method. At $p = 0.005$, the hypothesis holds true for all countries except Canada ($p = 0.8$). Among video sets, the hypothesis holds for *Conspiracy* and *News* ($p < 0.00005$), and *Popular* and *Science* ($p < 0.05$), but not for *Kids* ($p = 0.8$). This is in line with what we expected as the effect sizes—the differences in scam rates between the privacy settings—are most prominent in the *Conspiracy* and *News* video sets.

As the next step, we conducted post-hoc tests for pairwise comparisons between the privacy settings. However, unlike the aggregated data or the ad load data, these results did not reach the same

	Pers. On / Act. Off		Pers. Off / Act. Off			
	Ratio	95% CI	Ratio	95% CI	R^2_{CS}	n
News	2.35	[1.30, 4.25]	2.54	[1.40, 4.60]	0.20	30
Popular	3.89	[1.86, 8.13]	2.66	[1.26, 5.64]	0.20	30
Science	3.17	[1.67, 6.00]	2.73	[1.44, 5.19]	0.15	30
Kids	1.28^*	[0.60, 2.70]	1.24^*	[0.59, 2.64]	0.27	30
Conspiracy	4.11	[2.38, 7.09]	4.55	[2.64, 7.85]	0.32	30
Australia	6.10	[3.01, 12.37]	5.26	[2.59, 10.69]	0.77	30
Canada	1.46^*	[0.80, 2.70]	1.67^*	[0.91, 3.07]	0.50	30
Ireland	2.71	[1.43, 5.12]	3.05	[1.62, 5.74]	0.34	30
United Kingdom	3.33	[1.72, 6.47]	3.69	[1.90, 7.14]	0.55	30
United States	3.12	[1.66, 5.88]	1.93	[1.02, 3.67]	0.70	30
All	2.87	[2.17, 3.81]	2.69	[2.03, 3.57]	0.55	150

Table 4: Predatory Pre-roll Ad Rate: Ratios of predatory ad rate of stronger ad privacy settings relative to the default setting. Stronger privacy settings are personalization on & activity saving off, and personalization off & activity saving off, respectively. These are results from fitting a negative binomial model, where the dependent variable is the predatory ad rate, and the predictors are country, video type, and privacy setting, with n watch sequences per privacy setting. The table includes the predatory rate ratios using the default setting as the baseline, along with a 95 % confidence interval for the ratios and the model's pseudo-R^2 (a measure of relative fit, ranging from 0 to 1). Ratios with (*) are not significant at $p = 0.05$. An illustrative reading of the table: in the data subset that contains only *News* watch sequences, compared to the default setting, the ad personalization enabled/saving activity disabled setting observed 2.35 times more predatory ads with a 95 % confidence interval of [1.30, 4.25]. When all personalization is disabled, the increase is 2.54 times with a 95 % confidence interval of [1.40, 4.60]. The model for 90 *News* data points (30 per privacy setting) has a pseudo-R^2 of 0.20. Overall, disabling ad personalization increased the share of predatory pre-roll ads by an average of 2.69 times compared to the default privacy setting.

level of significance (a detailed table of post-hoc test output is included in Table 9 in Appendix C). We suspect this is due to several factors: the limited sample size of 30 data points per privacy setting (compared to 150 data points for the aggregated data), the need to correct for 33 comparisons, and the non-normal distribution and high variances of the predatory rates compared to ad load data. Under these constraints, our tests on data segmented by country and video type lacked the statistical power to achieve significance. Despite this, even in these data segments, the overall trend of more private settings leading to more predatory ads remains evident, as shown in Tables 4 and 5. Table 4 presents the ratios of predatory ad rates of the two more private settings relative to the default setting, as estimated by the regression model, whereas Table 5 lists the raw mean predatory ad rates (without the ratios).

Table 4 shows that while the overall increase in predatory ads is observed across all countries and video types, there are some exceptions—specifically *Kids* videos and Canada—that did not reach statistical significance after correcting for multiple comparisons. This lack of significance in certain cases aligns with the pairwise comparisons above; given that these two analyses are conducted on the same data set, they are subject to the same statistical constraints, including limited sample sizes and high variances.

This subsection has illustrated that while the overall trend—stronger privacy settings leading to more predatory ads—is observed across all countries and video types, there are some exceptions when it comes to significance levels of this effect. The driving factor of these exceptions is the varying difference in effect size between the video types and countries. (In contrast, the increases in ad load are more similar across these dimensions.) Our next analysis investigates whether ad targeting explanations could provide clues as to why there are differences in the predatory rates both across privacy settings and between video types.

4.3 Ad Targeting Transparency

The previous subsections have shown that YouTube tends to deliver a significantly higher proportion of predatory ads to our accounts with stronger ad privacy settings. In order to begin to understand why this may be the case, we now investigate the targeting explanations of the ads seen in our experiments, since they are supposed to explain to the user why they are seeing a certain ad. Specifically, we are interested in understanding whether predatory advertisers target their ads in a way that is associated with the differences in privacy settings and video types that we have observed.

Table 6 shows the 10 most common targeting reasons across three privacy settings. While a range of targeting criteria appeared for ads received under the default privacy setting and the personalization on/activity off setting, ads received under the most private

	Default: Pers. On/Activity On	Pers. On/Activity Off	Pers. Off/Activity Off	*n*
News	3.5, [0.0, 17.1], $\sigma = 3.7$	10.2, [0.0, 39.5], $\sigma = 9.2$	10.0, [0.0, 25.7], $\sigma = 6.2$	30
Popular	0.8, [0.0, 4.1], $\sigma = 1.2$	3.5, [0.0, 14.8], $\sigma = 3.9$	2.6, [0.0, 12.5], $\sigma = 3.7$	30
Science	1.7, [0.0, 9.5], $\sigma = 2.0$	5.2, [0.0, 24.2], $\sigma = 5.3$	4.4, [0.0, 14.3], $\sigma = 4.3$	30
Kids	1.2, [0.0, 5.8], $\sigma = 1.7$	1.4, [0.0, 6.1], $\sigma = 1.8$	1.4, [0.0, 6.7], $\sigma = 2.0$	30
Conspiracy	5.5, [0.0, 13.3], $\sigma = 3.3$	23.0, [6.4, 53.1], $\sigma = 10.3$	25.3, [11.5, 41.9], $\sigma = 10.2$	30
Australia	1.7, [0.0, 6.7], $\sigma = 2.1$	11.5, [0.0, 53.1], $\sigma = 14.4$	10.5, [0.0, 41.9], $\sigma = 13.3$	30
Canada	4.1, [0.0, 12.5], $\sigma = 3.4$	7.1, [0.0, 29.3], $\sigma = 7.8$	8.9, [0.0, 40.4], $\sigma = 10.9$	30
Ireland	2.5, [0.0, 9.3], $\sigma = 2.7$	6.3, [0.0, 20.5], $\sigma = 6.0$	7.7, [0.0, 32.9], $\sigma = 8.4$	30
United Kingdom	2.0, [0.0, 8.7], $\sigma = 2.2$	7.1, [0.0, 25.8], $\sigma = 8.0$	8.0, [0.0, 33.3], $\sigma = 8.1$	30
United States	2.4, [0.0, 17.1], $\sigma = 4.0$	11.3, [0.0, 43.8], $\sigma = 12.4$	8.8, [0.0, 37.6], $\sigma = 11.9$	30
All	2.5, [0.0, 17.1], $\sigma = 3.1$	8.7, [0.0, 53.1], $\sigma = 10.3$	8.8, [0.0, 41.9], $\sigma = 10.6$	150

Table 5: Predatory Pre-roll Ad Rate: Percentage of pre-roll ads that are predatory per watch sequence. Entries are: *mean, [min, max], standard deviation*; they are aggregated over *n* watch sequences. With the default settings, an average of 2.5 % of ads shown to users are labeled as predatory; with ad personalization disabled, the rate increases to 8.7 %.

Reason	Default: Pers. On/Act. On	Pers. On/Act. Off	Pers. Off/Act. Off
The time of day or your general location (like your country or city)	88.4	90.2	100.0
Your age	53.6	40.6	0.0
Google's estimation of your interests, based on your activity while you were signed in to Google	43.3	0.0	0.0
The video you're watching	37.2	48.8	0.0
Google's estimation of your areas of interest, based on your activity	13.7	10.6	0.0
Google's estimation of your approximate current location	13.7	10.6	0.0
Your activity, while you were signed in to Google	12.0	6.4	0.0
Google's estimation of your interests	10.8	29.4	0.0
Your gender	10.1	6.4	0.0
Household Income range	7.0	1.7	0.0

Table 6: Top 10 most common targeting disclosure reasons for each privacy setting, shown as percentage. Each percentage represents the share of ads under the corresponding setting that included a given targeting reason. Since one ad disclosure can include multiple reasons, percentages do not sum to 100 %. When ad personalization was disabled, advertisers could not target ads based on the current video being watched (or any past activity) according to these targeting disclosures.

setting contained ad explanations only for targeting by time of day and general location. That is, assuming ad targeting explanations provided by YouTube included all advertiser-specified criteria, ads received while ad personalization was disabled were not targeted by video placement, for instance. Yet, as we have shown in Table 5, the predatory ad rate varied depending on the type of videos being watched from an average of 1.4 % for *Kids* videos to 25.3 % for *Conspiracy* videos. Although these predatory ads were not explicitly targeted by their advertisers, they appeared on different video types at a different rate. Reconciling these two observations leads us to hypothesize that YouTube's ad system, or ad delivery optimization, may be the driving force of this phenomenon by selecting different ads depending on the video where they will appear. When the ad system is not allowed to utilize information about the user (because ad personalization is disabled), it might optimize ad delivery by using behaviors from other information-rich users who previously engaged with the same content (i.e., the prior audience of the video). This is similar to the argument made by Ali et al. [3], who argue that a lack of specific targeting parameters indicates that the advertiser has chosen default targeting, and any skew in ad distribution is likely driven by the ad system's delivery optimization and other aspects of the ad auction, such as advertisers' spending power.

The "middle" privacy setting, where ad personalization is enabled but web & app activity is disabled, presents a more complex picture. Table 6 shows that the ad explanations observed were broadly similar to the ad explanations with the default privacy setting, except that no ads were disclosed as being shown due to "Google's estimation of your interests, based on your activity while you were signed in to Google." (However, two variants of this disclosure, "Google's estimation of your interests" and "Your activity, while you were signed in to Google," were still present.) Despite the similarity in targeting explanations to the default setting, the predatory ad rate for the "middle" setting was closer to that of the strongest privacy setting (which disables ad personalization entirely), i.e., significantly higher than in the default setting. Based on these data alone, it is unclear what might have caused this outcome.

While the different predatory ad rates between video types (when ad personalization was disabled) might arise from Google's ad delivery optimization, differences within a video type—between the default privacy setting and the two stronger settings—require a different explanation. Within the same video type, Google has the same attribute inferences for our sock puppet users and the same general audience estimates for the video type, therefore differences in predatoriness cannot be attributed to ad delivery optimization alone (nor direct advertiser targeting, since they cannot target privacy settings). For example, Table 4 shows that predatoriness could increase as much as 4 times that of the default setting when users watched *Conspiracy* videos with stronger privacy settings enabled. We speculate that this may be driven by economic incentives: with personalization enabled, users would be eligible for targeted ads, which typically command higher prices due to increased effectiveness [65]. Given Google's ad auction mechanism, advertisers willing to pay higher prices would be preferentially shown to users whose behavioral data is available for ad selection (i.e., users who have ad personalization and activity saving enabled). Consequently, if we assume that predatory advertisers are cost-sensitive and bid on the lower spectrum, information-rich users are more likely to receive higher-quality (often legitimate) ads, which naturally displace lower-quality (e.g., predatory) ads. In contrast, when personalization and activity are disabled, the lack of information renders the user less valuable [14, 45], and more within reach of low-bidding predatory advertisers.

In conclusion, while our previous analyses have shown that disabling all ad personalization led to a higher rate of predatory pre-roll ads compared to the default setting, this increase is not adequately explained by Google's ad targeting explanations. While our findings highlight the effects of the privacy settings on ad quality and quantity, a thorough explanation of these effects would require a deeper understanding of the interaction between ad delivery optimization and economic incentives, and in turn access to more fine-grained data on ad pricing, delivery and targeting, which is currently not available.

5 Discussion

YouTube has provided users with ad privacy settings to configure how personalized they want their ads to be, but cautions that stronger ad privacy settings will result in less relevant ads. What YouTube does not state, but our experiments have shown, is that configuring stronger ad privacy settings also impacts users in two additional major ways: More ads and often disproportionately more dangerous ads.

Disabling ad personalization led to an average of 1.30 times more pre-roll ads, and this effect occurred consistently across nearly all scenarios with low variances. This suggests an ad load system purposefully engineered to extract a target ad revenue per user watch time, which means that users receiving "cheaper" ads (such as non-personalized ads) need to endure more of them. Receiving more ads is a negative consequence that, on its own, is not necessarily concerning.

More concerningly, disabling ad personalization also led to our sock puppets receiving 2.69 times more ads that we considered predatory. We observed higher variances in the predatory ad rate across video types, which suggests that this phenomenon may be a side effect of other factors in the ad system rather than a deliberate design decision. Due to limited visibility into the internals of the ad delivery system, we cannot determine with certainty which parts of it are to blame for this outcome. However, we can rule out explicit targeting by predatory advertisers based on no such targeting appearing in the ad explanations when all ad personalization was disabled. Instead, it is likely a combination of multiple factors, including the ad delivery optimization system and economics. Ad delivery optimization may cause users who disable ad personalization to receive ads based on the aggregate preferences of the general audience of the content they are consuming; thematically speaking, predatory ads are likely more engaging to the audience of *Conspiracy* videos than the audience of *Kids* videos, for instance. (We observed the corresponding trend in our experiments that there are more predatory ads for *Conspiracy* videos, even when ad personalization is disabled.) On the other hand, economic forces in online advertising can also put users who disable ad personalization at a disadvantage. We hypothesize that the kind of predatory advertiser observed in our experiments is cost-conscious and tends to place ad bids on the lower end of the spectrum. Less valuable audiences, such as users with ad personalization disabled, are more likely to receive these more indiscriminate, "bottom-of-the-barrel" ads because they are not eligible to receive higher-bid personalized ads that could displace them.

We stress that the higher predatory ad rates were observed when all ad personalization was disabled, and therefore the ad delivery system likely had a larger role in routing these predatory ads to specific users than the predatory advertisers. The question, then, is how privacy controls in an online ad platform can be implemented *safely*, that is, in a way that strengthens privacy without inadvertently increasing exposure to dangerous content.

Our results bear conceptual similarity to those of Ali et al. [3], who found that the distribution of "problematic" ads is biased towards certain demographics such as older people and minority groups, with the difference that privacy-conscious users are an entirely self-selected population. That is, our results likely reflect a broader, more fundamental side effect of personalized ad systems. Unfortunately, there is no strong incentive for platforms such as Google to invest more resources into safer privacy controls or to investigate and detect scams beyond a minimum level that keeps legitimate advertisers around, as platforms face little liability, and most of the harms of scams are borne by the users.

Lastly, our experiments strongly suggest that YouTube's ad explanations are incomplete. When ad personalization was disabled, videos of different types received identical ad explanations, yet different types of ads (as evidenced by the varying scam ad rates, for example). Consequently, these ad explanations did not actually explain to users why a certain ad was shown to them. We hypothesize that YouTube's ad explanations currently include only advertiser-defined targeting criteria, and omit any additional signals used by the ad delivery system to select the respective ad. Specifically for Ireland, this could potentially be in violation of Article 26 of the European Union's Digital Services Act.

5.1 Limitations

Our experiments were designed to isolate the effects of privacy settings on the observed ad load or predatory ad rate from other confounding factors, such as the type of videos being watched, user location, or the time of measurement. While certain age groups are more likely to see problematic ads [3, 50], and women are often more targeted [70], we chose to prioritize diversifying our experiments across countries and video types instead, thus, we kept age and gender as a fixed variable. Furthermore, while we did not provide any account information beyond what was outlined in Section 3.2.1, Google made demographic inferences about our accounts based on our watch history when ad personalization was enabled. We deleted these inferences before reusing an account. Google's inferences sometimes differed between nearly identical accounts watching the same videos simultaneously from the same IP address (Degeling and Nierhoff observed similar unexpected variations in inferred interests [21]). Due to this inconsistency, we chose not to further analyze the demographic inferences made by Google during the experiments. Separately, Google also makes ad interest inferences, which however are not revealed to users. Similarly, Google's ad explanations mention only the broad category of targeting but not any concrete value, which limited our analysis of targeting reasons.

Our experiments do not explain *why* privacy settings influence ad load or predatoriness. Furthermore, we did not measure how these effects change over time—for example, whether predatory ad rates decrease or converge as targeting attributes accumulate. It is conceivable that the predatory ad rate with the default privacy settings is even lower when more targeting attributes are available, i.e. with a longer activity history than in our experiments. As a result, the consequences of strengthening privacy settings might be more severe in practice for real users with a long watch history and rich interest profile. On the other hand, it is conceivable that users with a real interest in what we classify as predatory ads (or an interest presumed by Google) might see a reduction in predatory ads when they switch off ad personalization.

We acknowledge that the definition (and determination) of which ads are to be considered predatory (i.e., where exactly to draw the line) is subject to debate. We attempted to address this issue with a conservative definition in order to underestimate rather than overestimating the rate of predatory ads, and by testing for agreement among annotators.

Lastly, we note that our results apply to pre-roll ads on YouTube. It is an open question to what extent they might also apply to other ad systems on different platforms.

5.2 Ethics

This research is based on emulated "sock puppet" users and did not collect human subject data. The study was classified as exempt by our institution's IRB. To conduct this research, we created accounts that were machine-operated and therefore siphoned ad impressions that were intended for humans. We acknowledge that this activity runs counter to the immediate economic interests of YouTube and their advertisers. We limited negative impacts of our experiments by not clicking on ads, and argue that the benefits of better understanding undesired and potentially dangerous side effects of ad privacy settings outweigh the costs that our experiments have caused. We will share our findings with Google to make them aware of the issue and allow them to implement mitigations. We will also publish our code and data for the benefit of the community.

6 Conclusion

In this paper, we measured the impact of ad privacy settings on the number and predatoriness of pre-roll ads on YouTube. Disabling ad personalization led to 1.30 times more pre-roll ads on average than the default setting. The percentage of ads labeled as predatory increased 2.69 times, from 2.5 % in the default setting to an average of 8.7 %. That is, privacy-conscious users "pay" for their choice by seeing more ads in general, and disproportionately more predatory ads in particular. The increased number of ads is likely by design given the consistency of this effect, whereas the selection of more predatory ads may be a side effect of the ad system. We suspect that it arises from a combination of ad delivery optimization (especially when ad personalization is disabled) and economic forces within the ad ecosystem, as more private users are less valuable to advertisers. Our results raise questions about how ad privacy controls can be implemented without increasing risks to users.

Acknowledgments

This work is supported by the NSF under grant no. 2151837. We thank Michael Satz, Ian Goldstein, and Laura Edelson for their consultation on statistical analysis, as well as the anonymous reviewers for their feedback and suggestions that helped improve the paper. The authors used Writefull (in Overleaf) and Grammarly to correct typos, grammatical errors, and awkward phrasing.

Our code and data set of labeled ads is available at https://github.com/CybersecurityForDemocracy/youtube-ad-settings.

References

[1] A. Acquisti, C. R. Taylor, and L. Wagman. The Economics of Privacy. *Journal of Economic Literature*, 54(2), 2016. DOI: 10.1257/jel.54.2.442.

[2] E. Aguirre, A. L. Roggeveen, D. Grewal, and M. Wetzels. The personalization-privacy paradox: implications for new media. *Journal of Consumer Marketing*, 33(2), Jan. 2016. ISSN: 0736-3761. DOI: 10.1108/JCM-06-2015-1458.

[3] M. Ali, A. Goetzen, A. Mislove, E. M. Redmiles, and P. Sapiezynski. Problematic Advertising and its Disparate Exposure on Facebook. In *32nd USENIX Security Symposium*, 2023. ISBN: 978-1-939133-37-3. URL: https://www.usenix.org/system/files/usenixsecurity23-ali.pdf.

[4] M. Ali, P. Sapiezynski, M. Bogen, A. Korolova, A. Mislove, and A. Rieke. Discrimination through Optimization: How Facebook's Ad Delivery Can Lead to Biased Outcomes. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW), Nov. 2019. ISSN: 2573-0142. DOI: 10.1145/3359301.

[5] M. Ali, P. Sapiezynski, A. Korolova, A. Mislove, and A. Rieke. Ad Delivery Algorithms: The Hidden Arbiters of Political Messaging. In *Proceedings of the 14th ACM International Conference on Web Search and Data Mining*, WSDM '21. Association for Computing Machinery, Mar. 2021. DOI: 10.1145/3437963.3441801.

[6] Alphabet. Alphabet Announces Second Quarter 2023 Results. 2023. URL: https://abc.xyz/assets/20/ef/844a05b84b6f9dbf2c3592e7d9c7/2023q2-alphabet-earnings-release.pdf.

[7] A. Andreou, G. Venkatadri, O. Goga, K. P. Gummadi, P. Loiseau, and A. Mislove. Investigating Ad Transparency Mechanisms in Social Media: A Case Study of Facebook's Explanations. In *Proceedings 2018 Network and Distributed System Security Symposium*. Internet Society, 2018. DOI: 10.14722/ndss.2018.23191.

[8] J. Asplund, M. Eslami, H. Sundaram, C. Sandvig, and K. Karahalios. Auditing Race and Gender Discrimination in Online Housing Markets. *Proceedings of the International AAAI Conference on Web and Social Media*, 14, May 2020. ISSN: 2334-0770. DOI: 10.1609/icwsm.v14i1.7276.

[9] C. Ballard, I. Goldstein, P. Mehta, G. Smothers, K. Take, V. Zhong, R. Greenstadt, T. Lauinger, and D. McCoy. Conspiracy Brokers: Understanding the Monetization of YouTube Conspiracy Theories. In *Proceedings of the ACM Web Conference 2022*, 2022. DOI: 10.1145/3485447.3512142.

[10] J. Bandy. Problematic Machine Behavior: A Systematic Literature Review of Algorithm Audits. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW1), Apr. 2021. ISSN: 2573-0142. DOI: 10.1145/3449148.

[11] M. A. Bashir, U. Farooq, M. Shahid, M. F. Zaffar, and C. Wilson. Quantity vs. Quality: Evaluating User Interest Profiles Using Ad Preference Managers. In *Proceedings 2019 Network and Distributed System Security Symposium*. Internet Society, 2019. DOI: 10.14722/ndss.2019.23392.

[12] Y. Beugin and P. McDaniel. Interest-disclosing Mechanisms for Advertising are Privacy-Exposing (not Preserving). *Proceedings on Privacy Enhancing Technologies*, 2024(1), 2024. ISSN: 2299-0984. DOI: 10.56553/popets-2024-0004.

[13] J. G. Cabañas, Á. Cuevas, and R. Cuevas. Unveiling and Quantifying Facebook Exploitation of Sensitive Personal Data for Advertising Purposes. In *27th USENIX Security Symposium*, 2018. ISBN: 978-1-939133-04-5. URL: https://www.usenix.org/system/files/conference/usenixsecurity18/sec18-cabanas.pdf.

[14] C. Castelluccia, L. Olejnik, and T. Minh-Dung. Selling Off Privacy at Auction. In *Network and Distributed System Security Symposium (NDSS)*. Internet Society, 2014. ISBN: 978-1-891562-35-8. URL: https://inria.hal.science/hal-01087557/.

[15] S. Chandio, M. D. P. Dar, and R. Nithyanand. How Audit Methods Impact Our Understanding of YouTube's Recommendation Systems. *Proceedings of the International AAAI Conference on Web and Social Media*, 18, May 2024. ISSN: 2334-0770. DOI: 10.1609/icwsm.v18i1.31311.

[16] M. Chau and E. K. Clemons. Individual Privacy and Online Services. In *2011 44th Hawaii International Conference on System Sciences*, Jan. 2011. DOI: 10.1109/HICSS.2011.242.

[17] J. W. Clark and D. McCoy. There Are No Free iPads: An Analysis of Survey Scams as a Business. In *6th USENIX Workshop on Large-Scale Exploits and Emergent Threats (LEET 13)*, 2013. URL: https://www.usenix.org/system/files/conference/leet13/leet13-paper_clark.pdf.

[18] J. Cohen. *Statistical Power Analysis for the Behavioral Sciences*. Routledge, 2013. DOI: 10.4324/9780203771587.

[19] P. Covington, J. Adams, and E. Sargin. Deep Neural Networks for YouTube Recommendations. In *Proceedings of the 10th ACM Conference on Recommender Systems*. ACM, Sept. 2016. DOI: 10.1145/2959100.2959190.

[20] A. Datta, M. C. Tschantz, and A. Datta. Automated Experiments on Ad Privacy Settings. *Proceedings on Privacy Enhancing Technologies*, 2015. ISSN: 2299-0984. URL: https://petsymposium.org/popets/2015/popets-2015-0007.php.

[21] M. Degeling and J. Nierhoff. Tracking and Tricking a Profiler: Automated Measuring and Influencing of Bluekai's Interest Profiling. In *Proceedings of the 2018 Workshop on Privacy in the Electronic Society*, WPES'18. Association for Computing Machinery, Jan. 2018. DOI: 10.1145/3267323.3268955.

[22] S. J. Dixon. Most popular social networks worldwide as of July 2023, ranked by number of monthly active users. 2023. URL: https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/.

[23] M. Faddoul, G. Chaslot, and H. Farid. A Longitudinal Analysis of YouTube's Promotion of Conspiracy Videos, Mar. 2020. URL: http://arxiv.org/abs/2003.03318.

[24] F. M. Farke, D. G. Balash, M. Golla, M. Dürmuth, and A. J. Aviv. Are Privacy Dashboards Good for End Users? Evaluating User Perceptions and Reactions to Google's My Activity. In *30th USENIX Security Symposium*, 2021. ISBN: 978-1-939133-24-3. URL: https://www.usenix.org/system/files/sec21-farke.pdf.

[25] R. A. Ford. Data Scams. *Houston Law Review*, 57(1), 2019. URL: https://houstonlawreview.org/article/10856-data-scams.

[26] L. Gak, S. Olojo, and N. Salehi. The Distressing Ads That Persist: Uncovering The Harms of Targeted Weight-Loss Ads Among Users with Histories of Disordered Eating. *Proceedings of the ACM on Human-Computer Interaction*, 6(CSCW2), Nov. 2022. DOI: 10.1145/3555102.

[27] Global Disinformation Index. The Quarter Billion Dollar Question: How is Disinformation Gaming Ad Tech? 2019. URL: https://www.disinformationindex.org/research/2019-9-1-the-quarter-billion-dollar-question-how-is-disinformation-gaming-ad-tech/.

[28] J. Goldman. Google's robust Q2 earnings buoyed by significant growth in advertising revenues across search and YouTube. 2024. URL: https://www.emarketer.com/content/google-s-robust-q2-earnings-buoyed-by-significant-growth-advertising-revenues-across-search-youtube.

[29] Google. Google Safety Center: Ads and Data. 2023. URL: https://safety.google/privacy/ads-and-data/.

[30] Google. My Ad Center. URL: https://myadcenter.google.com/. Accessed June 2025.

[31] Google. Personalized and non-personalized ads. URL: https://support.google.com/adsense/answer/9007336. Accessed June 2025.

[32] Google. What happens when your content is set as made for kids. URL: https://support.google.com/youtube/answer/9527654?sjid=12816564088846718219-NA#what_happens. Accessed June 2025.

[33] D. W. Hosmer and S. Lemeshow. *Applied Logistic Regression*. Applied Logistic Regression. Wiley, 2004, page 167. ISBN:

9780471654025. URL: https://books.google.com/books?id=Po0RLQ7USIMC.

[34] E. Hussein, P. Juneja, and T. Mitra. Measuring Misinformation in Video Search Platforms: An Audit Study on YouTube. *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW1), May 2020. ISSN: 2573-0142. DOI: 10.1145/3392854.

[35] B. Imana, A. Korolova, and J. Heidemann. Auditing for Discrimination in Algorithms Delivering Job Ads. In *Proceedings of the Web Conference 2021*, WWW '21, June 2021. DOI: 10.1145/3442381.3450077.

[36] L. Kaplan, N. Gerzon, A. Mislove, and P. Sapiezynski. Measurement and Analysis of Implied Identity in Ad Delivery Optimization. In *Proceedings of the 22nd ACM Internet Measurement Conference*. ACM, Oct. 2022. DOI: 10.1145/3517745.3561450.

[37] A. Kharraz, W. Robertson, and E. Kirda. Surveylance: Automatically Detecting Online Survey Scams. In *2018 IEEE Symposium on Security and Privacy (SP)*. IEEE, May 2018. DOI: 10.1109/SP.2018.00044.

[38] S. Kingsley, C. Wang, A. Mikhalenko, P. Sinha, and C. Kulkarni. Auditing Digital Platforms for Discrimination in Economic Opportunity Advertising. *arXiv*, 2020. URL: https://arxiv.org/pdf/2008.09656.

[39] A. Lambrecht and C. E. Tucker. Algorithmic Bias? An Empirical Study into Apparent Gender-Based Discrimination in the Display of STEM Career Ads. *Management Science*, 65(7), 2019. DOI: 10.1287/mnsc.2018.3093.

[40] V. Le Pochat, L. Edelson, T. V. Goethem, W. Joosen, D. McCoy, and T. Lauinger. An Audit of Facebook's Political Ad Policy Enforcement. In *31st USENIX Security Symposium*, 2022. ISBN: 978-1-939133-31-1. URL: https://www.usenix.org/system/files/sec22-lepochat.pdf.

[41] M. Ledwich and A. Zaitsev. Algorithmic Extremism: Examining YouTube's Rabbit Hole of Radicalization, Dec. 2019. URL: http://arxiv.org/abs/1912.11211.

[42] O. Lesota, A. Melchiorre, N. Rekabsaz, S. Brandl, D. Kowald, E. Lex, and M. Schedl. Analyzing Item Popularity Bias of Music Recommender Systems: Are Different Genders Equally Affected? In *Fifteenth ACM Conference on Recommender Systems*. ACM, Sept. 2021. DOI: 10.1145/3460231.3478843.

[43] X. Li, A. Yepuri, and N. Nikiforakis. Double and Nothing: Understanding and Detecting Cryptocurrency Giveaway Scams. In *Proceedings 2023 Network and Distributed System Security Symposium*. Internet Society, 2023. DOI: 10.14722/ndss.2023.24584.

[44] Z. Li, K. Zhang, Y. Xie, F. Yu, and X. Wang. Knowing Your Enemy: Understanding and Detecting Malicious Web Advertising. In *Proceedings of the 2012 ACM conference on Computer and communications security*, CCS '12. Association for Computing Machinery, Oct. 2012. DOI: 10.1145/2382196.2382267.

[45] McKinsey. The value of getting personalization right—or wrong—is multiplying. 2021. URL: https://www.mckinsey.com/capabilities/growth-marketing-and-sales/our-insights/the-value-of-getting-personalization-right-or-wrong-is-multiplying.

[46] Media Bias/Fact Check. URL: https://mediabiasfactcheck.com/.

[47] T. Medjkoune, O. Goga, and J. Senechal. Marketing to Children Through Online Targeted Advertising: Targeting Mechanisms and Legal Aspects. In *Proceedings of the 2023 ACM SIGSAC Conference on Computer and Communications Security*, CCS '23. Association for Computing Machinery, Nov. 2023. DOI: 10.1145/3576915.3623172.

[48] Meta. Personalized Advertising and Privacy Are Not at Odds. 2020. URL: https://about.fb.com/news/2020/12/personalized-advertising-and-privacy-are-not-at-odds/.

[49] N. Miramirkhani, O. Starov, and N. Nikiforakis. Dial One for Scam: A Large-Scale Analysis of Technical Support Scams. In *Proceedings 2017 Network and Distributed System Security Symposium*. Internet Society, 2017. DOI: 10.14722/ndss.2017.23163.

[50] K. Munger, M. Luca, J. Nagler, and J. Tucker. The (Null) Effects of Clickbait Headlines on Polarization, Trust, and Learning. *Public Opinion Quarterly*, 84(1), Mar. 2020. ISSN: 0033-362X. DOI: 10.1093/poq/nfaa008.

[51] MyTop100Videos. Most Viewed Videos of All Time. URL: https://www.youtube.com/@MyTop10Videos/playlists.

[52] T. Nelms, R. Perdisci, M. Antonakakis, and M. Ahamad. Towards Measuring and Mitigating Social Engineering Software Download Attacks. In *25th USENIX Security Symposium*, 2016. ISBN: 978-1-931971-32-4. URL: https://www.usenix.org/system/files/conference/usenixsecurity16/sec16_paper_nelms.pdf.

[53] C. Oh, C. Kanich, D. McCoy, and P. Pearce. Cart-ology: Intercepting Targeted Advertising via Ad Network Identity Entanglement. In *Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security*. ACM, Nov. 2022. DOI: 10.1145/3548606.3560641.

[54] M. Pachilakis, P. Papadopoulos, N. Laoutaris, E. P. Markatos, and N. Kourtellis. YourAdvalue: Measuring Advertising Price Dynamics without Bankrupting User Privacy. *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, 5(3), Dec. 2021. ISSN: 2476-1249. DOI: 10.1145/3491044.

[55] Panoptykon Foundation. Algorithms of trauma: new case study shows that Facebook doesn't give users real control over disturbing surveillance ads. 2021. URL: https://en.panoptykon.org/algorithms-of-trauma.

[56] K. Papadamou, S. Zannettou, J. Blackburn, E. De Cristofaro, G. Stringhini, and M. Sirivianos. "It is just a flu": Assessing the Effect of Watch History on YouTube's Pseudoscientific Video Recommendations, 2022. URL: https://ojs.aaai.org/index.php/ICWSM/article/download/19329/19101.

[57] M. Paquet-Clouston, M. Romiti, B. Haslhofer, and T. Charvat. Spams meet Cryptocurrencies: Sextortion in the Bitcoin Ecosystem. In *Proceedings of the 1st ACM Conference on Advances in Financial Technologies*, AFT '19. Association for Computing Machinery, Oct. 2019. DOI: 10.1145/3318041.3355466.

[58] V. Rastogi, R. Shao, Y. Chen, X. Pan, S. Zou, and R. Riley. Are these Ads Safe: Detecting Hidden Attacks through the Mobile App-Web Interfaces. In *Proceedings 2016 Network and Distributed System Security Symposium*. Internet Society, 2016. DOI: 10.14722/ndss.2016.23234.

[59] N. Reitinger, B. Wen, M. L. Mazurek, and B. Ur. Analysis of Google Ads Settings Over Time: Updated, Individualized, Accurate, and Filtered. In *Proceedings of the 22nd Workshop on Privacy in the Electronic Society*. ACM, Nov. 2023. DOI: 10.1145/3603216.3624968.

[60] P. Sampson, R. Encarnacion, and D. Metaxa. Representation, Self-Determination, and Refusal: Queer People's Experiences with Targeted Advertising. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*, FAccT '23. Association for Computing Machinery, June 2023. DOI: 10.1145/3593013.3594110.

[61] L. Spinelli and M. Crovella. How YouTube Leads Privacy-Seeking Users Away from Reliable Information. In *Adjunct Publication of the 28th ACM Conference on User Modeling, Adaptation and Personalization*. ACM, July 2020. DOI: 10.1145/3386392.3399566.

[62] I. Srba, R. Moro, M. Tomlein, B. Pecher, J. Simko, E. Stefancova, M. Kompan, A. Hrckova, J. Podrouzek, A. Gavornik, and M. Bielikova. Auditing YouTube's Recommendation Algorithm for Misinformation Filter Bubbles. *ACM Transactions on Recommender Systems*, 1(1), Mar. 2023. ISSN: 2770-6699. DOI: 10.1145/3568392.

[63] K. Thomas, E. Bursztein, C. Grier, G. Ho, N. Jagpal, A. Kapravelos, D. Mccoy, A. Nappa, V. Paxson, P. Pearce, N. Provos, and M. A. Rajab. Ad Injection at Scale: Assessing Deceptive Advertisement Modifications. In *2015 IEEE Symposium on Security and Privacy*, May 2015. DOI: 10.1109/SP.2015.17.

[64] K. Thomas, J. A. E. Crespo, R. Rasti, J.-M. Picod, C. Phillips, M.-A. Decoste, C. Sharp, F. Tirelo, A. Tofigh, M.-A. Courteau, L. Ballard, R. Shield, N. Jagpal, M. A. Rajab, P. Mavrommatis, N. Provos, E. Bursztein, and D. McCoy. Investigating Commercial Pay-Per-Install and the Distribution of Unwanted Software. In *25th USENIX Security Symposium*, 2016. ISBN: 978-1-931971-32-4. URL: https://www.usenix.org/system/files/conference/usenixsecurity16/sec16_paper_thomas.pdf.

[65] Top Draw Inc. Is online advertising expensive? Online Advertising Costs In 2025. 2024. URL: https://www.topdraw.com/insights/is-online-advertising-expensive/.

[66] Tubics. How Does YouTube Count Views? 2020. URL: https://www.tubics.com/blog/what-counts-as-a-view-on-youtube/.

[67] J. Turow, M. Hennessy, and N. Draper. The Tradeoff Fallacy: How Marketers are Misrepresenting American Consumers and Opening Them Up to Exploitation, SSRN Scholarly Paper, June 2015. DOI: 10.2139/ssrn.2820060.

[68] G. Venkatadri, A. Andreou, Y. Liu, A. Mislove, K. P. Gummadi, P. Loiseau, and O. Goga. Privacy Risks with Facebook's PII-Based Targeting: Auditing a Data Broker's Advertising Interface. In *2018 IEEE Symposium on Security and Privacy (SP)*, May 2018. DOI: 10.1109/SP.2018.00014.

[69] G. Venkatadri, P. Sapiezynski, E. M. Redmiles, A. Mislove, O. Goga, M. Mazurek, and K. P. Gummadi. Auditing Offline Data Brokers via Facebook's Advertising Platform. In *The World Wide Web Conference*, WWW '19, May 2019. DOI: 10.1145/3308558.3313666.

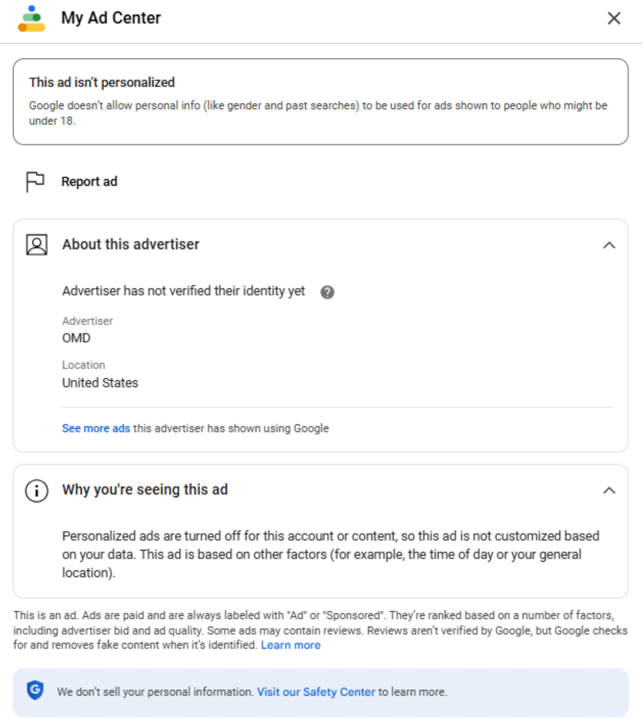
[70] M. Wei, M. Stamos, S. Veys, N. Reitinger, J. Goodman, M. Herman, D. Filipczuk, B. Weinshel, M. L. Mazurek, and B. Ur. What Twitter Knows: Characterizing Ad Targeting Practices, User Perceptions, and Ad Explanations Through Users' Own Twitter Data. In *29th USENIX Security Symposium*, 2020. URL: https://www.usenix.org/system/files/sec20-wei.pdf.

[71] A. Zarras, A. Kapravelos, G. Stringhini, T. Holz, C. Kruegel, and G. Vigna. The Dark Alleys of Madison Avenue: Understanding Malicious Advertisements. In *Proceedings of the 2014 Conference on Internet Measurement Conference*. ACM, Nov. 2014. DOI: 10.1145/2663716.2663719.

[72] E. Zeng, T. Kohno, and F. Roesner. What Makes a "Bad" Ad? User Perceptions of Problematic Online Advertising. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, May 2021. DOI: 10.1145/3411764.3445459.

[73] E. Zeng, R. McAmis, T. Kohno, and F. Roesner. What Factors Affect Targeting and Bids in Online Advertising? A Field Measurement Study. In *Proceedings of the 22nd ACM Internet Measurement Conference*, IMC '22. Association for Computing Machinery, Oct. 2022. DOI: 10.1145/3517745.3561460.

[74] E. Zeng, M. Wei, T. Gregersen, T. Kohno, and F. Roesner. Polls, Clickbait, and Commemorative $2 Bills: Problematic Political Advertising on News and Media Websites Around the 2020 U.S. Elections. In *Proceedings of the 21st ACM Internet Measurement Conference*. ACM, Nov. 2021. DOI: 10.1145/3487552.3487850.



This ad isn't personalized

Google doesn't allow personal info (like gender and past searches) to be used for ads shown to people who might be under 18.

About this advertiser

Advertiser has not verified their identity yet

Advertiser
OMD

Location
United States

See more ads this advertiser has shown using Google

Why you're seeing this ad

Personalized ads are turned off for this account or content, so this ad is not customized based on your data. This ad is based on other factors (for example, the time of day or your general location).

This is an ad. Ads are paid and are always labeled with "Ad" or "Sponsored". They're ranked based on a number of factors, including advertiser bid and ad quality. Some ads may contain reviews. Reviews aren't verified by Google, but Google checks for and removes fake content when it's identified. Learn more

We don't sell your personal information. Visit our Safety Center to learn more.

Figure 3: An example of the targeting disclosure when the ad was not personalized because Google suspects the viewer might be under 18 (even though we age-verified all accounts).

A Labeling Practices

Raters were provided with a link to the YouTube video ad, any display text, and when available, the domain of the landing page (or when available, also the landing page), along with the advertiser's name and location. The video ad was unavailable at labeling time, raters relied on other available information. The video description often included links to the business websites, which we visited as needed. In cases where the video was unavailable and there was no other information, we labeled the ad as non-predatory (we did not encounter this very often). We only closely examined ads that exhibited signs of potential scams, which are listed in the codebook below. Besides watching the video ad, raters were instructed to examine user reviews on reputable sites such as Yelp, Better Business Bureau, Trustpilot, or sites with known authentic user reviews such as Reddit or Facebook Reviews. If the ad or the homepage were in a language that the raters could not understand, we used Google Translate; if the content remained unclear, we labeled the ad as non-predatory. It is important to note that due to time and effort constraints, we did not interact with the advertisers (e.g. make a purchase, download software, sign up for free products etc.) beyond visiting their websites, nor did we spend more than 10 minutes researching a single business. Thus we made reasonable assumptions about the predatoriness of the ads but also exercised caution and erred on the side of being conservative in our labeling.

B Codebook

We include indicators of predatory ads seen in our experiments. These indicators serve as red flags, with a predatory ad potentially having one or more of these signs. Not all ads that exhibited these indicators were labeled as predatory, either because after researching the advertiser, we found no compelling evidence for predatory practices, or because of the distinctions that we have listed under each indicator. Under each indicator, we list examples from ads that did and did not receive the predatory label, either in quotes from their websites/ads or specific examples that pertain to the category. It is not within the scope of our study to be exhaustive about categorizing all possible predatory ads. Ads that did not contain any of the following indicators were labeled as non-predatory.

Ads in sectors that are controversial or prone to be predatory, such as subprime credits and loans, speculative or unregulated industries, e.g. cryptocurrency, foreign exchange market (forex), fortune-telling, alternative medicine etc.

- We label the ads as predatory if they make strongly worded guarantees or results that are highly unlikely to be met, or make claims about their services/products or expertise that cannot be verified. We also use user reviews to corroborate the labeling, e.g. when the review scores are very low. Examples include:
  - Guarantees of overnight, low-effort success in business, trading profits, weight loss
  - Sketchy dating sites or matching services
  - Cure-all products for health problems without scientific backing
  - "Quantum Healing Hypnosis practitioners have cured various types of cancer"
  - "I promise you'd make another $10 million a year"
  - "This natural 10-second-a-day method, discovered by Harvard researchers, will help your hearing problems"
- We label the ads as non-predatory if they use exaggerated language but are transparent about the advertised product or the variability of outcomes, or there is no strong evidence for predatory practices beyond the controversial nature of the business.
  - Legitimate subprime lenders or creditors
  - For-profit universities
  - Legitimate crypto and other currencies trading platforms
  - Aggregator sites (affiliated product reviews, news sites)
  - "Our panel of experts will make your financial independence possible"
  - "The Most Effective Nurse Coach Training on the Planet"

Ads advertising businesses with strong evidence of malicious or deceptive practices that we can verify through user reviews or Google searches of the business.

- We label the ads as predatory if the business or the ad contains, including but not limited to:
  - Businesses with (class-action) lawsuits for their practices (e.g., Lendio,[7] My Forex Funds[8])
  - Promises of free products but does not deliver [17]
  - Evidence of delivering counterfeit or very low quality products, based on user reviews
  - Pyramid scheme, multi-level marketing scheme
  - Hidden recurring/non-cancellable billing
  - Websites that contain or distribute malware, e.g. search hijackers, survey scams
  - Other types of scams evidenced through Google searches
- Ads in this category are never labeled as non-predatory.

Ads advertising activities that are illegal or promote dishonest behaviors that we can verify through Google searches.

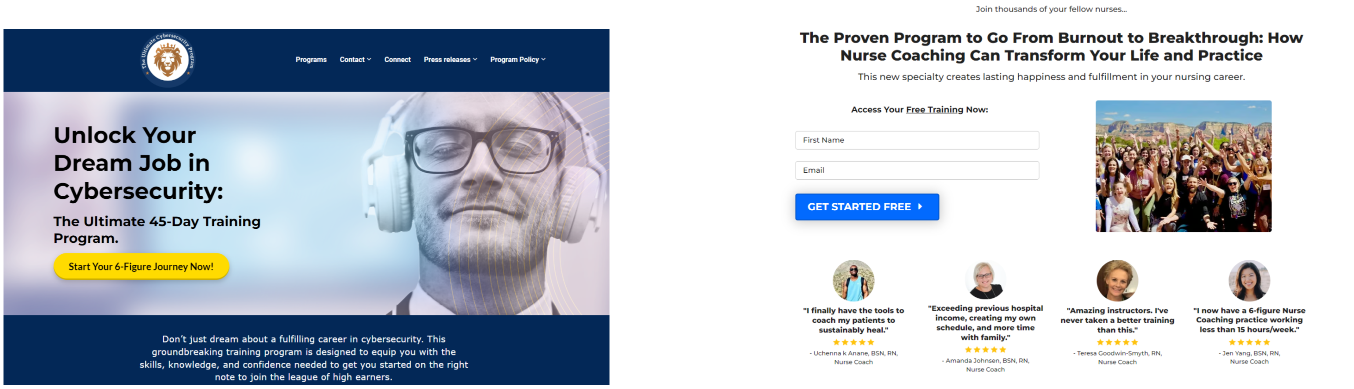
- We label the ads as predatory if they promote, for instance:
  - Hacking services
  - Selling fraudulent user engagement (fake likes or reviews)
  - Academic dishonesty services, such as paper-writing or exam-taking
- Ads in this category are never labeled as non-predatory.

Ads that use problematic or manipulative patterns such as clickbait (which can be promises of free products, attention-grabbing yet misleading thumbnails), misinformation, exaggerated language.

- We label the ads as predatory if we can find evidence of predatoriness of the business, using the indicators listed above. We label the ads as predatory if they use a bait-and-switch tactic for products or services that users have to pay money for.
- We label the ads as non-predatory if the ads are problematic but there is no substantial evidence of the business itself engaging in predatory practices that we could find. We also label the ads as non-predatory if the products or services are offered free of charge (without data theft, malware, or other harmful caveats).

[7] https://www.ftc.gov/news-events/news/press-releases/2020/05/ftc-sba-warn-operator-sbacom-lead-generator-lendio-stop-potentially-misleading-coronavirus-relief

[8] https://www.cftc.gov/PressRoom/PressReleases/8771-23


The Ultimate Cybersecurity Program
Programs
Contact
Connect
Press releases
Program Policy
Start Your 6-Figure Journey Now!
Access Your Free Training Now:
First Name
Email
GET STARTED FREE

Unlock Your Dream Job in Cybersecurity:

The Ultimate 45-Day Training Program.

Don't just dream about a fulfilling career in cybersecurity. This groundbreaking training program is designed to equip you with the skills, knowledge, and confidence needed to get you started on the right note to join the league of high earners.

Join thousands of your fellow nurses...

The Proven Program to Go From Burnout to Breakthrough: How Nurse Coaching Can Transform Your Life and Practice

This new specialty creates lasting happiness and fulfillment in your nursing career.

"I finally have the tools to coach my patients to sustainably heal." ★★★★★ - Uchenna k Anane, BSN, RN, Nurse Coach

"Exceeding previous hospital income, creating my own schedule, and more time with family." ★★★★★ - Amanda Johnsen, BSN, RN, Nurse Coach

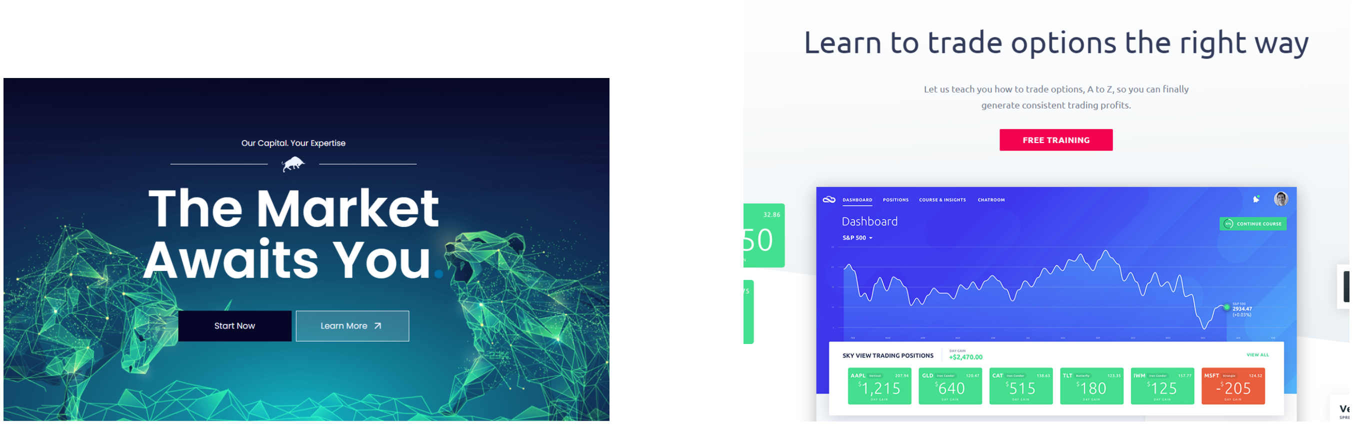
"Amazing instructors. I've never taken a better training than this." ★★★★★ - Teresa Goodwin-Smyth, RN, Nurse Coach

"I now have a 6-figure Nurse Coaching practice working less than 15 hours/week." ★★★★★ - Jen Yang, BSN, RN, Nurse Coach

(a) Homepage for ExcelMindCyber. The site advertises 45-day training course at $1,999 to break into cybersecurity with no experience required, with guaranteed job offers. We labeled this ad as predatory due to the highly unlikely nature of the advertised claim.

(b) Homepage for The Nurse Coach Collective. The site boasts a "proven" program to career transformation. We labeled this as non-predatory due to the neutral reviews of users, and evidence of success in the program (through actual health board certifications).

Figure 4: Two examples of career coaching ads' homepages, one ad labeled as predatory and the other not.


Our Capital. Your Expertise
The Market Awaits You.
Start Now
Learn More
FREE TRAINING
DASHBOARD
POSITIONS
COURSE & INSIGHTS
CHATROOM
Dashboard
CONTINUE COURSE
S&P 500
S&P 500
2934.47
(+0.03%)
SKY VIEW TRADING POSITIONS
DAY GAIN
+$2,470.00
VIEW ALL
AAPL Vertical 207.94 $1,215 DAY GAIN
GLD Iron Condor 120.47 $640 DAY GAIN
CAT Iron Condor 138.63 $515 DAY GAIN
TLT Butterfly 123.35 $180 DAY GAIN
IWM Iron Condor 157.77 $125 DAY GAIN
MSFT Strangle 124.52 -$205 DAY GAIN

Learn to trade options the right way

Let us teach you how to trade options, A to Z, so you can finally generate consistent trading profits.

(a) Homepage for My Forex Funds. The prop trading firm is being sued by The Commodity Futures Trading Commission for fraudulently taking customer money. We labeled the ad as predatory.

(b) Homepage for Skyview Trading. It is an options-trading educational website that sells courses and other financial tools, along with providing free-of-charge content. We labeled the ad as non-predatory as there are mixed user reviews.

Figure 5: Two examples of financial ads' homepages, one ad labeled as predatory and the other not.

C Test Results & Data

This appendix section lists the statistical test results and raw data.

	Default: Pers. On/Activity On	Pers. On/Activity Off	Pers. Off/Activity Off	*n*
News	13.0, [5.0, 35.0], $\sigma = 5.1$	15.9, [7.5, 23.8], $\sigma = 4.4$	16.0, [6.5, 35.0], $\sigma = 5.8$	30
Popular	15.9, [11.5, 19.5], $\sigma = 2.0$	19.6, [10.5, 25.8], $\sigma = 3.5$	19.6, [13.5, 24.2], $\sigma = 2.5$	30
Science	16.2, [13.5, 20.5], $\sigma = 1.4$	21.1, [15.5, 27.8], $\sigma = 2.3$	21.5, [12.8, 26.5], $\sigma = 2.4$	30
Conspiracy	16.9, [11.8, 21.0], $\sigma = 1.9$	22.3, [18.8, 28.0], $\sigma = 2.0$	23.0, [20.2, 27.0], $\sigma = 1.9$	30
Kids	19.3, [16.8, 26.0], $\sigma = 1.8$	24.4, [13.8, 29.0], $\sigma = 3.4$	24.8, [12.0, 30.8], $\sigma = 4.0$	30
Australia	16.6, [11.5, 20.5], $\sigma = 2.2$	21.2, [13.8, 28.7], $\sigma = 3.1$	21.4, [7.0, 27.5], $\sigma = 4.1$	30
Canada	15.4, [11.0, 19.0], $\sigma = 2.2$	20.0, [12.0, 25.8], $\sigma = 3.2$	20.4, [12.0, 27.0], $\sigma = 3.3$	30
Ireland	15.3, [6.2, 26.0], $\sigma = 4.2$	17.5, [7.5, 24.5], $\sigma = 5.5$	17.9, [6.5, 24.5], $\sigma = 5.3$	30
United Kingdom	16.2, [5.0, 20.8], $\sigma = 3.1$	21.5, [16.8, 28.0], $\sigma = 2.8$	21.6, [12.5, 27.3], $\sigma = 3.6$	30
United States	17.9, [11.2, 35.0], $\sigma = 4.2$	23.1, [13.0, 29.0], $\sigma = 4.3$	23.5, [12.8, 35.0], $\sigma = 5.0$	30
All	16.3, [5.0, 35.0], $\sigma = 3.4$	20.6, [7.5, 29.0], $\sigma = 4.3$	21.0, [6.5, 35.0], $\sigma = 4.7$	150

Table 7: Pre-roll Ad Load: The ratio of ads to videos, shown as percentages (i.e., the number of pre-roll ads per 100 videos). Entries are: *mean, [min, max], standard deviation*; they are aggregated over *n* watch sequences. Our sock puppets with the default setting received 16.3 ads per 100 videos on average; with ad personalization disabled, the number increases to 21.0 ads.

	Ad Load (One-way r-ANOVA)			Predatory Ad Rate (Friedman Test)		
	Test statistic	**Original p-value**	**Corrected p-value**	**Test statistic**	**Original p-value**	**Corrected p-value**
All	232.262 475	0.000000	0.000000	68.474 383	0.000000	0.000000
Country USA	106.033 036	0.000000	0.000000	18.740 000	0.000000	0.000597
Country Canada	59.516 871	0.000000	0.000000	3.500 000	0.173774	0.347548
Country UK	81.833 281	0.000000	0.000000	27.160 714	0.000001	0.000011
Country Australia	38.310 915	0.000000	0.000000	18.686 869	0.000088	0.000597
Country Ireland	12.622 289	0.000028	0.000028	12.134 615	0.002317	0.011587
Video Set Popular	25.937 917	0.000000	0.000000	10.473 118	0.005319	0.015956
Video Set Conspiracy	202.646 033	0.000000	0.000000	41.058 824	0.000000	0.000000
Video Set Kids	39.361 418	0.000000	0.000000	0.285 714	0.866878	0.866878
Video Set News	17.602 907	0.000001	0.000002	25.351 351	0.000003	0.000025
Video Set Science	161.408 635	0.000000	0.000000	11.274 336	0.003563	0.014252

Table 8: Test statistics and p-values for omnibus tests for pre-roll ad load and predatory ad rates. Holm-Bonferroni correction is used to correct p-values. All p-values are significant at the 0.01 level.

		Ad Load (Tukey's HSD Test)		Predatory Ad Rate (Conover Test)	
		Original p-value	**Corrected p-value**	**Original p-value**	**Corrected p-value**
All	False_False vs True_False	0.799900	1.000000	0.705134	1.000000
	False_False vs True_True	0.000000	0.000000	0.000000	0.000000
	True_False vs True_True	0.000000	0.000000	0.000000	0.000000
Country USA	False_False vs True_False	0.970900	1.000000	0.173043	1.000000
	False_False vs True_True	0.000000	0.000000	0.039907	0.638519
	True_False vs True_True	0.000000	0.000000	0.000840	0.017633
Country Canada	False_False vs True_False	0.786700	1.000000	0.617826	1.000000
	False_False vs True_True	0.000000	0.000000	0.065978	0.989670
	True_False vs True_True	0.000000	0.000000	0.176941	1.000000
Country UK	False_False vs True_False	0.988900	1.000000	0.795927	1.000000
	False_False vs True_True	0.000000	0.000000	0.000199	0.005361
	True_False vs True_True	0.000000	0.000000	0.000484	0.011610
Country Australia	False_False vs True_False	0.998600	1.000000	0.870161	1.000000
	False_False vs True_True	0.000000	0.000000	0.000679	0.014932
	True_False vs True_True	0.000000	0.000000	0.000392	0.009802
Country Ireland	False_False vs True_False	0.900700	1.000000	0.906421	1.000000
	False_False vs True_True	0.106200	1.000000	0.005409	0.102769
	True_False vs True_True	0.244600	1.000000	0.007554	0.135975
Video set Popular	False_False vs True_False	0.990800	1.000000	0.159475	1.000000
	False_False vs True_True	0.000000	0.000000	0.031258	0.531392
	True_False vs True_True	0.000000	0.000000	0.000514	0.011825
Video set Conspiracy	False_False vs True_False	0.385300	1.000000	0.293537	1.000000
	False_False vs True_True	0.000000	0.000000	0.000000	0.000000
	True_False vs True_True	0.000000	0.000000	0.000000	0.000000
Video set Kids	False_False vs True_False	0.909300	1.000000	0.746542	1.000000
	False_False vs True_True	0.000000	0.000000	0.598719	1.000000
	True_False vs True_True	0.000000	0.000000	0.396327	1.000000
Video set News	False_False vs True_False	0.907500	1.000000	0.456310	1.000000
	False_False vs True_True	0.010900	0.152600	0.000008	0.000230
	True_False vs True_True	0.003100	0.046500	0.000132	0.003686
Video set Science	False_False vs True_False	0.188900	1.000000	0.749085	1.000000
	False_False vs True_True	0.000000	0.000000	0.000904	0.018083
	True_False vs True_True	0.000000	0.000000	0.000309	0.008046

Table 9: P-values for post-hoc pairwise comparisons for pre-roll ad load and predatory ad rates. Holm-Bonferroni correction is used to correct p-values. True_True denotes the default setting, where both ad personalization and using activity for ads are enabled. True_False denotes the "middle" private setting, where ad personalization is enabled but using activity for is disabled. False_False is the setting where both ad personalization and using activity for ads are disabled.

		Default: Personalization On/Activity On		Personalization On/Activity Off		Personalization Off/Activity Off	
Location	**Run**	**Pre-roll Count**	**Pre-roll Predatory Rate**	**Pre-roll Count**	**Pre-roll Predatory Rate**	**Pre-roll Count**	**Pre-roll Predatory Rate**
Australia	1	55	0.0364	83	0.0482	83	0.0843
	2	62	0.0484	87	0.2184	70	0.1571
	3	64	0.0000	78	0.0769	86	0.1628
	4	63	0.0159	69	0.1594	61	0.1148
	5	60	0.0000	95	0.1053	86	0.1512
	6	72	0.0417	86	0.3953	28	0.0357
Canada	1	52	0.0577	66	0.0152	70	0.0857
	2	50	0.0600	65	0.0923	60	0.1667
	3	52	0.0577	51	0.0980	66	0.0758
	4	52	0.0577	67	0.0149	73	0.0411
	5	44	0.0682	69	0.1304	73	0.1096
	6	53	0.0377	70	0.1429	61	0.1311
Ireland	1	25	0.0800	30	0.1000	26	0.0385
	2	33	0.0000	35	0.0000	35	0.2571
	3	30	0.0000	31	0.0323	31	0.0000
	4	44	0.0227	38	0.0526	39	0.0513
	5	35	0.0000	32	0.0000	34	0.0000
	6	35	0.0000	34	0.0294	34	0.0000
US	1	45	0.0000	52	0.0385	60	0.0500
	2	54	0.0556	70	0.2143	74	0.1757
	3	47	0.0000	62	0.0323	71	0.0986
	4	56	0.0000	69	0.1014	69	0.0580
	5	52	0.0192	73	0.2329	77	0.1169
	6	47	0.1714	62	0.1774	70	0.1714
UK	1	51	0.0196	70	0.0286	71	0.0704
	2	50	0.0200	70	0.0429	50	0.0800
	3	69	0.0508	70	0.0714	79	0.1266
	4	49	0.0000	74	0.0000	67	0.1940
	5	55	0.0727	67	0.1940	70	0.1143
	6	20	0.0500	78	0.2179	76	0.0921

Table 10: Data for News video type, $n = 30$.

		Default: Personalization On/Activity On		Personalization On/Activity Off		Personalization Off/Activity Off	
Location	**Run**	**Pre-roll Count**	**Pre-roll Predatory Rate**	**Pre-roll Count**	**Pre-roll Predatory Rate**	**Pre-roll Count**	**Pre-roll Predatory Rate**
Australia	1	60	0.0333	80	0.0375	79	0.0253
	2	59	0.0169	73	0.0000	76	0.0132
	3	59	0.0000	78	0.0000	71	0.0282
	4	46	0.0000	66	0.0606	67	0.0000
	5	64	0.0000	80	0.0000	83	0.0000
	6	58	0.0172	76	0.0658	79	0.1013
Canada	1	48	0.0000	65	0.0000	74	0.0000
	2	60	0.0167	85	0.0353	81	0.0741
	3	55	0.0182	87	0.0230	76	0.0132
	4	63	0.0000	81	0.0247	79	0.0127
	5	66	0.0000	83	0.0602	76	0.0395
	6	55	0.0182	48	0.0000	75	0.0000
Ireland	1	72	0.0000	86	0.1047	75	0.0000
	2	69	0.0000	71	0.0000	77	0.0000
	3	59	0.0169	48	0.1250	72	0.1250
	4	63	0.0000	88	0.1477	77	0.1169
	5	57	0.0000	68	0.0294	70	0.0286
	6	70	0.0000	42	0.0000	75	0.0267
US	1	67	0.0000	83	0.0000	76	0.0132
	2	76	0.0132	92	0.0435	93	0.0000
	3	71	0.0000	90	0.0333	97	0.0000
	4	67	0.0000	80	0.0625	87	0.0000
	5	73	0.0411	95	0.0105	96	0.0104
	6	76	0.0000	103	0.0388	59	0.0000
UK	1	53	0.0000	93	0.0108	97	0.0000
	2	78	0.0128	84	0.0238	90	0.0111
	3	66	0.0000	77	0.0390	86	0.0116
	4	64	0.0156	73	0.0000	73	0.0274
	5	67	0.0299	83	0.0723	54	0.0926
	6	70	0.0000	88	0.0000	87	0.0000

Table 11: Data for Popular video type, $n = 30$.

		Default: Personalization On/Activity On		Personalization On/Activity Off		Personalization Off/Activity Off	
Location	**Run**	**Pre-roll Count**	**Pre-roll Predatory Rate**	**Pre-roll Count**	**Pre-roll Predatory Rate**	**Pre-roll Count**	**Pre-roll Predatory Rate**
Australia	1	64	0.0156	84	0.0000	85	0.0116
	2	66	0.0000	87	0.0575	90	0.0444
	3	70	0.0000	92	0.1087	94	0.0426
	4	60	0.0167	86	0.0581	86	0.0698
	5	64	0.0000	81	0.0000	87	0.0000
	6	57	0.0175	94	0.0638	90	0.0333
Canada	1	63	0.0952	73	0.0000	80	0.0769
	2	67	0.0149	87	0.0805	88	0.0000
	3	62	0.0000	74	0.0811	87	0.0920
	4	59	0.0169	79	0.0253	84	0.0119
	5	65	0.0154	83	0.0000	87	0.0575
	6	67	0.0299	90	0.0222	80	0.0625
Ireland	1	61	0.0164	78	0.0128	83	0.0000
	2	60	0.0167	76	0.0000	78	0.0385
	3	66	0.0303	81	0.0988	82	0.0488
	4	61	0.0164	79	0.0633	79	0.1392
	5	63	0.0317	76	0.0526	78	0.0385
	6	71	0.0000	62	0.0000	80	0.0125
US	1	68	0.0000	82	0.0426	88	0.0000
	2	71	0.0423	84	0.1190	89	0.0449
	3	54	0.0000	98	0.0204	99	0.0000
	4	70	0.0143	84	0.0119	88	0.0114
	5	73	0.0274	95	0.2421	98	0.1429
	6	82	0.0000	111	0.1351	106	0.0094
UK	1	64	0.0156	86	0.0465	85	0.0824
	2	61	0.0328	83	0.0361	85	0.0824
	3	63	0.0000	90	0.0667	91	0.0110
	4	60	0.0000	84	0.0119	96	0.0208
	5	64	0.0312	77	0.0779	81	0.1358
	6	71	0.0000	81	0.0123	91	0.0110

Table 12: Data for Science video type, $n = 30$.

		Default: Personalization On/Activity On		Personalization On/Activity Off		Personalization Off/Activity Off	
Location	**Run**	**Pre-roll Count**	**Pre-roll Predatory Rate**	**Pre-roll Count**	**Pre-roll Predatory Rate**	**Pre-roll Count**	**Pre-roll Predatory Rate**
Australia	1	65	0.0615	89	0.2921	86	0.1977
	2	73	0.0548	79	0.3165	102	0.3922
	3	67	0.0000	79	0.2025	81	0.3086
	4	68	0.0294	86	0.3953	86	0.4186
	5	59	0.0339	91	0.2198	96	0.3750
	6	75	0.0667	96	0.5312	94	0.3511
Canada	1	64	0.1250	85	0.1529	88	0.1250
	2	68	0.0882	86	0.1977	91	0.3846
	3	47	0.0851	75	0.1733	84	0.2262
	4	61	0.0492	87	0.1839	91	0.1209
	5	70	0.0857	88	0.1932	94	0.2872
	6	71	0.0845	92	0.2935	99	0.4040
Ireland	1	55	0.0545	83	0.2048	86	0.1744
	2	70	0.0714	78	0.0641	81	0.1728
	3	68	0.0441	85	0.1412	86	0.1628
	4	54	0.0926	89	0.1011	88	0.1705
	5	84	0.0595	82	0.1951	85	0.3294
	6	67	0.0299	84	0.1429	87	0.1379
US	1	60	0.0000	94	0.1809	87	0.3218
	2	71	0.0282	97	0.2577	101	0.3762
	3	78	0.0513	101	0.3267	105	0.2857
	4	75	0.1333	94	0.2447	102	0.2157
	5	74	0.0676	105	0.3524	108	0.1296
	6	80	0.0250	112	0.4375	107	0.3738
UK	1	62	0.0323	89	0.2584	85	0.2235
	2	69	0.0290	87	0.1609	89	0.1461
	3	69	0.0290	91	0.1758	95	0.1684
	4	69	0.0145	93	0.0968	90	0.1667
	5	69	0.0870	87	0.1954	90	0.3333
	6	68	0.0294	96	0.2188	96	0.1146

Table 13: Data for Conspiracy video type, $n = 30$

		Default: Personalization On/Activity On		Personalization On/Activity Off		Personalization Off/Activity Off	
Location	**Run**	**Pre-roll Count**	**Pre-roll Predatory Rate**	**Pre-roll Count**	**Pre-roll Predatory Rate**	**Pre-roll Count**	**Pre-roll Predatory Rate**
Australia	1	76	0.0000	104	0.0000	110	0.0182
	2	79	0.0000	96	0.0104	97	0.0103
	3	81	0.0000	65	0.0000	109	0.0000
	4	78	0.0000	108	0.0000	109	0.0000
	5	80	0.0000	115	0.0087	108	0.0000
	6	82	0.0000	76	0.0132	85	0.0000
Canada	1	76	0.0132	98	0.0000	97	0.0000
	2	74	0.0270	96	0.0000	89	0.0112
	3	71	0.0563	86	0.0116	108	0.0185
	4	67	0.0149	103	0.0000	48	0.0208
	5	72	0.0278	88	0.0568	106	0.0094
	6	74	0.0000	88	0.0227	91	0.0000
Ireland	1	70	0.0286	86	0.0233	57	0.0526
	2	72	0.0000	98	0.0612	96	0.0000
	3	73	0.0411	96	0.0312	90	0.0667
	4	74	0.0135	84	0.0532	98	0.0408
	5	69	0.0580	82	0.0244	78	0.0641
	6	104	0.0192	84	0.0119	97	0.0103
US	1	74	0.0000	108	0.0000	110	0.0000
	2	80	0.0125	112	0.0268	112	0.0000
	3	74	0.0135	106	0.0000	113	0.0000
	4	80	0.0000	113	0.0000	113	0.0000
	5	76	0.0132	116	0.0000	123	0.0325
	6	91	0.0000	115	0.0000	108	0.0000
UK	1	78	0.0000	97	0.0103	102	0.0098
	2	71	0.0000	104	0.0096	99	0.0101
	3	82	0.0122	112	0.0268	109	0.0000
	4	78	0.0000	100	0.0000	102	0.0098
	5	79	0.0000	95	0.0316	101	0.0495
	6	83	0.0120	102	0.0000	108	0.0000

Table 14: Data for Kids video type, $n = 30$

	Ad Load (One-way r-ANOVA)		Predatory Ad Rate (Friedman Test)	
	Test statistic	**P-value**	**Test statistic**	**P-value**
Overall	47.67	0.000000	10.8	0.004517

(a) Omnibus tests

		Ad Load (Tukey's HSD Test)		Predatory Ad Rate (Conover Test)	
		Original p-value	**Corrected p-value**	**Original p-value**	**Corrected p-value**
Overall	False_False vs True_False	0.915000	0.915000	0.00764252	0.01528505
	False_False vs True_True	0.000500	0.001500	0.00491864	0.01475593
	True_False vs True_True	0.001500	0.003000	0.75247914	0.75247914

(b) Post-hoc pairwise comparisons

Table 15: Validation experiment (female persona). Test statistics and p-values for omnibus tests for pre-roll ad load and predatory ad rates. Since all p-values are significant at 0.005, post-hoc pairwise comparisons are performed for both ad load and predatory ad rate.