MY BUSINESS NOT THEIRS: A PILOT ANALYSIS OF THE SENSITIVITY OF AD CATEGORIES

Ashwin Chetty, Panya Gupta, Su Yilmaz
The University of Chicago
{achetty, panyagupta, suyilmaz}@uchicago.edu

ABSTRACT

Online behavioral advertising (OBA) refers to a broad class of methods which present advertisements only to users who find them the most interesting or relevant. A good deal of research has been previously conducted to determine algorithmic ways in which OBA can be carried out without sending too much data to a central server. However, relatively little research has explored targetted advertisements' social dimension: how do users feel about receiving certain types of ads? More importantly, do users see their personalized advertisements as reflections of themselves? What might others think about the ads that they receive? We conduct a study to find the particular advertisement categories which cause users discomfort, and to try to determine the factors involved in creating these feelings of discomfort.

INTRODUCTION

In the past decade, advertising on the web has emerged as a sophisticated machinery of advertisers, web browsers and ad networks, publishers, and third-party intermediaries working together to gather data on individual users and use data-intensive algorithms to provide these users with targeted and individualized online ads. This is a marked departure from the approach towards advertising on more traditional forms of media, such as newspapers or magazines, which, while seeking to provide ads tailored to their intended audiences, are not able to do so with the high degree of specificity that online advertising and algorithmic advertising allows for. This former mechanism is known as *Online Behavioral Advertising (OBA)* and has become the prevailing paradigm in web advertising, an industry sized at $49.5 billion in 2014 and projected to exceed $100 billion within two years time by 2019 [15].

Given that this type of personalized advertising on an individual user basis depends upon the amalgamation of large amounts of user data by third-party tracking services that are not only able and willing to sell this data to advertisers, but also may have a fiduciary responsibility to shareholders to do so, even as the users affected may lack the awareness or knowledge of this occurring, more research in this realm is necessary. Prior efforts have shown that while some users do find this service useful at times, more generally the public finds tracking of their online behavior concerning and invasive [16]. However, it is not yet clear what exactly people find concerning or invasive -- at what point does online tracking change from useful to creepy? Is it the accuracy or inaccuracy of profiling, the context in which users become aware of their profiling, the subject matter being tracked, or the extrapolations drawn from user profiling that result in specific ads being shown? In this paper, we aim to explore these questions further and provide a pilot study about the sensitivity of Google AdSense's ad-interest categories from which further research in this arena can be informed. In this, we aim to explore three main questions:

1. Are the categories or topics that Google AdSense classifies as sensitive actually perceived as sensitive by participants? That is, how accurate is the definition of sensitive categories?
2. Do people react differently to being accurately profiled versus inaccurately profiled? The dynamic we aim to explore here is whether people perceive privacy violations due to online tracking and OBA or due to inaccurate representations of themselves as a result of this tracking.
3. Does expected user behavior align with actual user behavior on a privacy dashboard? If users claim that they find a particular topic sensitive or uncomfortable, will they actually change their ad settings to reflect this?

As a pilot study, we do not aim to find statistically robust results for these questions. Rather, we aim to begin a limited discussion about a topic that requires much further research - privacy and sensitivity of ad-interest categories. Given how large the digital advertising industry is, this form of advertising and media is not going to decrease in significance; due to the potential privacy violations and ethical considerations present, research must be done to mitigate these concerns and guide how digital advertising and OBA evolves.

BACKGROUND AND RELATED WORK

It comes as no surprise that internet users tend to value convenience of action over security and privacy, and it is thus prudent to expect users to approach security and privacy from the perspective of practicality. This maxim is found to hold in particular within the realm of advertising, both on- and offline. While users see the benefits of personalized advertising, they are nonetheless wary of the amount and type of data collected by advertising agencies. Prior research, in examination of consumers' conceptions of online behavioral advertising, has shown that the magnitude of consumers' privacy concerns is related to the sensitivity of information collected [9]. Furthermore, consumers are generally more willing to make sacrifices to privacy for tailored advertisements that are relevant to their purposes, again highlighting how privacy and security tends to be a secondary concern [9]. Fitting with this theme of practicality over privacy, Agarwal et al. found that study participants were more concerned about being sent embarrassing advertisements (which others might see) than they were about the privacy concerns inherent in being tracked online [2]. Still, there seems to be a real danger with the information gained by the actions that people take while online. We see that "in 2008, 8% of U.S. companies employing 1000 workers or more had reported firing an employee because of information released on online social networks" [13]. Our research aims to examine this dynamic of embarrassment further, identifying what types of advertisements users find sensitive and the reasons they hold for this. Furthermore, we seek to quantitatively show participants' willingness to actively receive or not receive ads for topics they perceive as sensitive, thereby better understanding whether sensitivity to ads affects participants' behavior. However, the question of whether consumers feel that tailored advertisements are a reflection of their online presence, a notion that arose in Agarwal et al., remains an unanswered and very interesting question.

It appears that the tradeoffs consumers are willing to make with regards to the current state of online tracking and OBA are due to the knowledge, or lack thereof, that consumers have about OBA. Research shows that Internet users in the Netherlands do not have sufficient amounts of knowledge about OBA and cookies (both first and third-party) but do display some levels of concerns about potential misuse of private data and violations of privacy [10]. Similar results are found in the United States about knowledge gaps about the technologies and persistent identifiers advertisers can use to build comprehensive user profiles, such as cookies [1]. As a result, it is possible that participants would change their levels of concern about persistent identifiers for advertising given greater knowledge of the risks or privacy concerns that exist. Behavioral advertising, as it currently exists, does not conform to consumer expectations, resulting in greater privacy violations due to misconceptions consumers hold about the technologies used to gather data and about the usage of that personal data [6]. Consumers lack the knowledge to make informed decisions about behavioral advertising, and thus, research about the relationship between privacy concerns and online advertising is potentially misleading as even participants do not fully understand the tradeoffs they are making [2, 16].

User profiling methods generally fall into three types: social-based personalization, which includes potentially-sensitive information such as real name, demographic information, and interpersonal communication; behavioral profiling, which tracks the activity of a particular user over the long term (for which sensitive information may include information about sites visited and products purchased); and location-based personalization (where a user's location itself may be sensitive information) [13]. Ad personalization may either take place on the side of the server (such as is the case with Amazon and Google, who are able to build detailed histories of a user's browsing and purchasing habits), or on the side of the client, where the choice of which of several ads to display is determined based on local cookie data [5, 14]. Systems such as those use an aggregate of many users' browsing habits to predict advertisements for newer users, with success in limiting the number of false predictions while simultaneously reducing the amount of data which must be collected about particular users [18].

There appears to be a tradeoff between accuracy and privacy: what a system gains in term of privacy (a proxy for which might be the *amount* of information gathered about an individual), it must lose in terms of accuracy, resulting in research on systems that attempt to preserve both accuracy and privacy [8]. This is particularly true for client-side solutions [11]. However, there is only limited research done on users' perceptions on this tradeoff as well as the particular factors which lead to users' amiability toward online behavioral advertising. Even if there might be ways to protect user privacy and accurately profile users, would users want to be put into clusters based on their profile similarity to other users? This is a questions that requires further research.

User profiling based on preferences and search habits arose as online databases grew larger and more popular among the general public, resulting in an inundation of data in information retrieval systems [4, 7, 18]. Some existing recommender systems involve the collection and aggregation of individual users' data, followed by measuring correlations between each users to form similarity measures to discover 'like-minded' users and predict interests [7]. However, systems which generate user models must necessarily be aware of the information which they collect and the distinction between implicit features, such as those derived from keystrokes and elapsed time viewing documents, and explicit ones, such as survey results [18]. Additional thought must be placed in determining the nature of the information which is retrieved from users. While some services such as Amazon perform user-profiling on the server, many ad-distribution services instead send a complete list of ads to the client, which then locally determines those which are the most relevant [4]. However, while ample effort is placed into the task of developing privacy-aware recommendation systems, it seems that users still remain largely distrustful with the methods used in advertising, and are generally hesitant to condone such user-aware advertisement/recommendation systems. In our paper, we propose a design to elucidate the effect which user-knowledge of a system has on their attitude toward user experience personalization.

In fact, it appears that users react very negatively upon inaccurate profiling, as an article in the WSJ titled "If TiVo Thinks You are Gay, Here's How to Set It Straight, describes using anecdotes of TiVo users' changing TV viewing behavior to change TiVo's profile of them [17]. This again raises the question of sensitivity of information gathered, as sexual orientation is commonly accepted as deeply personal information; sensitivity and accuracy of ad categories is a main focus of our research in this paper. While finding online profiling, such as OBA, useful when done right, if not, it merely appears strange, concerning, and creepy [16]. However, what counts as scary or creepy is not clearly defined, and OBA is not going away. What is needed is a more nuanced understanding of what users find useful and what users find invasive -- where does this line exist and how can advertisers build systems that will ensure that they remain alongside the useful side of it?

METHODOLOGY

We conducted an online survey to gauge participants' awareness of and sensitivity to various ad-interest categories, taken from Google's list of topics. We then tasked participants to speculate on their ad preferences for each category in a modeled real-life scenario order to compare their expected and actual behavior and perception of both sensitive and nonsensitive ad-interest categories.

Study Structure

This study undertook a 4-part survey in order to investigate participants' sensitivities to various ad-interest categories. In the first part of the survey, participants were shown a series of statements questioning their basic knowledge of online behavioural advertising to establish a baseline level of knowledge and awareness. In the second part of the survey, each participant shown 5 ad-interest categories, including 3 non-sensitive and 2 sensitive categories, and asked to speculate on their expected behavior given each category. In the third part of the survey, participants were shown a privacy dashboard and tasked with making active decisions on their ad preferences based on the ad-interest categories shown. The final section of the survey was a basic demographics section.

In the first part, titled *baseline knowledge*, participants were provided with a series of statements about basic knowledge of computer usage and advertising on the web, and they were asked to rate the truthfulness of each statement on a 5-point Likert scale. The purpose of this section of the survey was to establish a baseline of participants' average awareness of online behavioural advertising, prior to any detailed questions. By conducting this section of the survey prior to the latter two sections, we sought to prevent priming the participants to consider OBA in any particular bias while still understanding what definitions and assumptions of OBA participants held. We also included an attention-check question in this section, which was a basic question about internet usage, "I am currently using a web browser to access the Internet," that all participants by necessity needed to agree with to answer truthfully.

In the second part of the survey, titled *sensitivity of ad categories*, participants were provided with a basic definition of online behavioural advertising, as taken from Google's definition of advertising. This definition was stripped down to avoid including references to any particular company. Participants were then shown 5 ad-interest categories selected from Google's list of ad-interest categories available online. The chosen categories deliberately included certain categories generally thought to be sensitive, such as those related to online dating, personal finance, and adult websites, as well as more "innocuous" categories. Each participant was shown 3 randomly selected non-sensitive categories out of 10 randomly selected from Google's broader list, 1 sensitive category out of 3 categories that we randomly selected from the larger list, and 1 restricted sensitive category out of the 2 restricted sensitive categories. This 3:1:1 ratio of categories was used in order to differentiate between non-sensitive, sensitive, and restricted sensitive ad categories. Because Google AdSense treats these ad categories differently, which affects how users receive ads from Google's ad services, we sought to emulate the Google standards as much as possible. However, it is important to note that the main difference between sensitive and restricted sensitive ad categories is that Google defines restricted sensitive ad categories as non-family safe due to content-based age restrictions on a per-country basis, and thus blocks them by default. Users can actively change their ad preferences to be shown ads in these categories, but for the purpose of this study, we decided to assume that most users have not done so. Furthermore, due to the small scale of this pilot study, we only used 10 standard ad categories selected at random out of the larger list of topics Google maintains, so results will vary depending on the standard categories used.

For each category, participants were asked to provide an example of a product or service that would fit within that category; this was to help us understand what participants believe fit within each category and see if that affects participants' understanding of the sensitive of each category. They were then asked whether participants would be interested in receiving ads from such a category in order to measure whether a participant falls within a particular category or not and how the accuracy of categorization affects participants' responses. The participant was also asked various questions to probe at the participant's sense of privacy from this ad without actually using the word privacy based on a 5-point Likert scale. For example, each of these questions was along the lines of: "How would you feel if others saw that you received this ad?" If a participant had initially answered that they would be interested in receiving ads from a particular category, they were then asked to assume that they were not interested and then answer the same questions, and vice-versa for participants who had initially answered that they were not interested in receiving ads from that category. This was to measure whether participants did or did not perceive potential privacy violations due to the content of the ad or due to their user profiling that resulted in them receiving that particular ad. The order in which participants received these privacy-based questions was randomized.

In the third part of the survey, titled *privacy dashboard*, participants were shown a privacy dashboard with all 15 of the categories we had selected for the prior section of the survey, including 10 categories that they did not answer detailed questions about. The categories were presented alphabetically to emulate other, more formal, privacy dashboards. Participants were told that their answers to the privacy dashboard might be sent to advertisers in

Table 2 - Residuals, Category vs Dashboard Result

	Show	Don't Care	Don't Show
Alcohol	-0.54	-0.16	0.41
Apartments & Residential Rentals	0.62	0.24	-0.51
Bars, Clubs & Nightlife	-0.54	0.24	0.1
Birth Control	-0.54	-0.55	0.71
Boats & Watercraft	-1.71	0.63	0.41
Dating	-0.54	-1.34	**1.33**
Fashion & Style	2.37	0.24	-1.43
Gambling and Betting	-1.13	-0.16	0.71
Hiking & Camping	1.21	-0.16	-0.51
Job Listings	0.62	-0.55	0.1
Legal Services	-1.71	0.63	0.41
Mobile & Wireless Accessories	0.62	1.42	-1.43
Pet Food & Supplies	2.96	0.24	-1.74
Scientific Equipment	0.04	0.24	-0.2
Sexual and Reproductive Health Clinics and Medication	-1.71	-0.95	**1.63**

order to better inform the ads they, as individuals, were shown. It was decided that advertisers would be used in a general sense, rather than specific advertising companies or Google, in order to prevent participants' prior biases about individual companies from biasing their answers. For instance, due to the ubiquitousness of Google's services, people may trust Google more with personal data than

specific advertising companies, so their answers to the questions, had they been framed to represent either Google or a specific

advertising company, may have reflected these biases. The purpose of this deception was to inculcate a greater sense of realism within the study in order to try to obtain more accurate results. As participants' data was not actually being sent to advertisers, the slight deception in the survey did not cause any harm to participants.

For each category, participants were tasked with making one decision out of the following three choices: "Show me more ads like this," "I don't care," and "Don't show me ads like this." This was to measure a participant's actual behavior in response to certain ad-interest categories. That way, we can compare the participant's believed behavior from Part 2 with their actual behavior in Part 3.

The final section of the survey, titled *demographics*, asked participants for basic demographic information, including age, gender, and technical background.

Recruitment

Participants were recruited for this survey via Amazon's Mechanical Turk (MTurk) platform for a survey about "online tracking and ad categories." As MTurk has been shown to provide data of similar quality as other, more traditional recruitment methods, our user of MTurk as a recruitment platform is acceptable. Participants were limited to be MTurk users age 18 and older who live in the United States, and participants were compensated $5 USD for the 30-minute study. Additionally, to ensure quality of the MTurk data, an attention check question, as described in the Study Structure section, was included.

PILOT RESULTS

Dashboard Behavior and Sensitivity of Ad-Interest Categories

One of the questions we wanted to answer in the beginning of our study was about the relationship between user preferences of ad-interest categories and the sensitivity of these categories. To test whether there is a relationship between users' dashboard preferences and Google's pre-prescribed sensitivity ratings, we related each category's sensitivity type (ie. general, sensitive, or restricted) to each user's dashboard responses. Tables 1 and 3 depict these results, $\chi^2 = 17.2, p = 0.001$.

We ran another Chi-squared test between category and user-preference to determine whether there were significant differences in dashboard responses between categories (Table 2).

Table 1 – Contingency Table

	Show	D_Care	D_Show	Total
General	37	72	91	200
Sensitive	4	12	44	60
Restricted	3	12	25	40
Total	44	96	160	300

Table 3 - X^2 Statistics

	Show	D_Care	D_Show
General	1.42	1	-1.52
Sensitive	-1.62	-1.64	2.12
Restricted	-1.18	-0.22	0.79

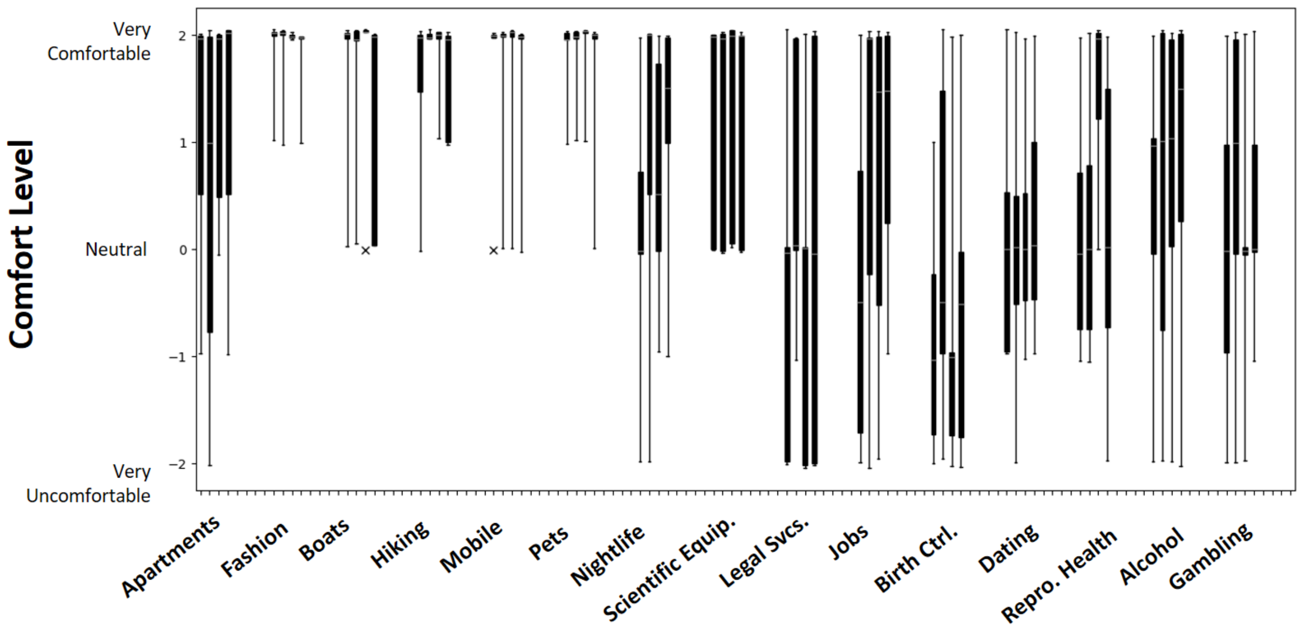


Figure 4 – A plot of user responses to the "how comfortable or uncomfortable would you be if [X] saw you receive ads in [Y]" questions asked in the survey. Each question asks about user comfort with respect to a particular category and observer. Categories are listed horizontally. Observers within each category are, from left to right, "Stranger", "Employer", "Friend, Family Member or Significant other", and "Computer Algorithm or Database". The right-most 5 categories are sensitive. Black bars represent 25-75th percentile of responses.

Due to the small sample size (n=20), these should be considered pilot results. However, the presence of significant p-values indicates that there may be a warrant to conduct future research.

Ad-Category Interest and Dashboard Behavior

We next attempted to find correlations between user interest in a category and their advertisement dashboard responses. We found that, with the exception of *Sexual and Reproductive Health Clinics and Medication*, all categories admitted nonnegative correlations between reported user interest and user willingness to view more ads in the category. In *birth control*, all users were either *uninterested* or *not at all interested*, and all dashboard responses were marked *don't show*. All participants responded *not at all interested* for Legal Services, and again all marked their dashboard with either *don't care* or *don't show*. Interestingly, *Sexual and Reproductive Health, Clinics and Medication* displayed a slight negative correlation. One participant, despite being *somewhat interested*, still refused to allow ads within the category. The relationships between interest and dashboard response for each category can be seen in the Appendix.

Participant Comfort by Category

As another measure of category sensitivity, we determined aggregate reported user comfort/discomfort felt by users as they received ads by each of a stranger, employer, friend, or algorithm. These results were taken from just the Our results show that, while people are in fact comfortable with being observed as receiving ads for many of the general categories, categories such as Apartments, Nightlife, Legal Services, and Jobs elicit more mixed reactions. In particular, our participants generally did not feel comfortable when strangers, friends, or computer algorithms observed their receiving ads for legal services.

Ads for Alcohol and Gambling seemed to be comfortable to receive among most respondents, and, despite their position in Google's list of restricted sensitive categories, were remarkably more welcome than those in legal services.

Comfort based on Profiling Accuracy

To understand the effect of *profiling accuracy* of advertisements on users' comfort reactions, we plotted the difference in reported comfort that users felt when they received ads for categories in which they *were* and *were not* interested. For each category, observer, and user, we recorded the difference in comfort between observation by algorithm; observation by friend; observation by employer; observation by stranger; the maximum comfort byany observer; the minimum comfort (ie. maximum discomfort) by any observer; and the average comfort over each observer. The *average* scores give us insight into the sensitivity of categories as wholes. The *minimum* scores give us insight into the *worst-case* discomfort experienced by any of the four observers. These results are shown in Figure 5. The values listed for each category and observer with

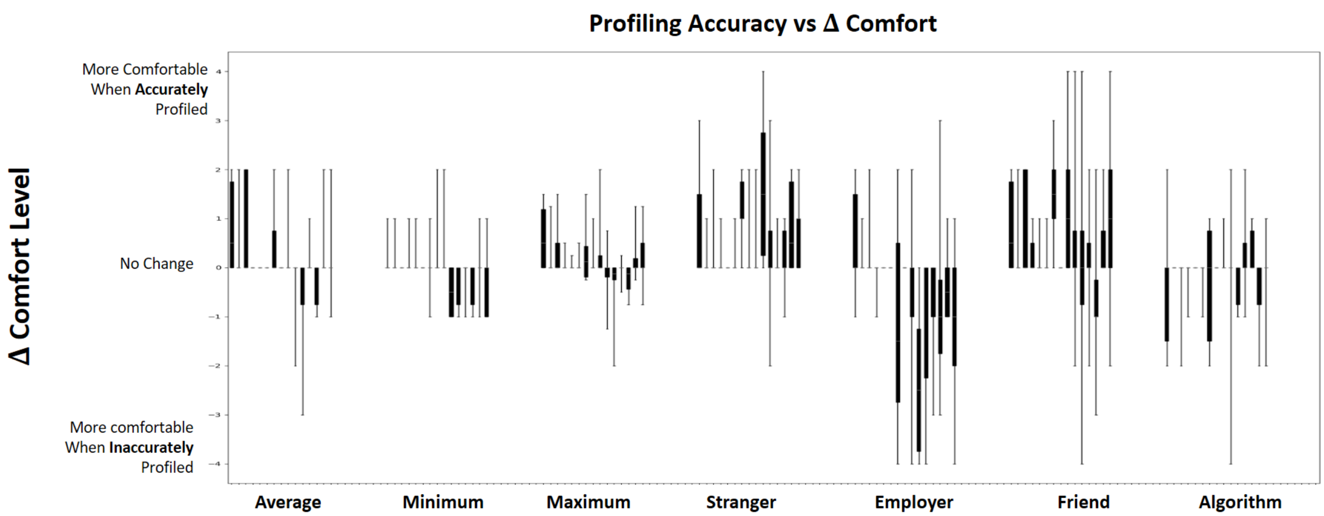


Figure 5 – A plot of changes in comfort level by both category and observer. Categories are represented from left-to-right in the canonical order (c.f. Figure 4), with the 10 general categories to the left, 3 sensitive categories in the middle, and 2 restricted categories on the right of each series.

$p < 0.15$ are presented in Table 6. Entries with $p < 0.05$ are nolded. Each general category contained 6 responses.

Each sensitive category contained either 6 or 7 responses. Each restricted-sensitive category contained 10 responses. *t-tests* were used to determine significance for each category and observer.

In general, our participants seemed to feel more comfortable when they were accurately represented by advertisements that were seen by strangers and friends. However, our participants were more uncomfortable when *correctly* receiving ads when they were observed by their employers. Of particular note is the fact that participants reported that they would feel far more uncomfortable when they accurately received ads for job listings while observed by their employer. One participant, who was not actively looking for a job, answered *somewhat uncomfortable* because it "may make my boss think I'm looking for a new job". His response changed to *extremely uncomfortable*, with the rationale that "They'd know I'm searching for a new job and I really would be. I'd not want them to know that"—seeming to imply that the *truth* of the implication that he is looking for a job makes the advertisement have a more deleterious effect.

Table 6

Category	**Observer**	**P-Value**	Δ Comfort
apartments	average	0.093	0.83
apartments	maximum	0.081	0.62
apartments	friend	0.093	0.83
bars	stranger	0.013	1.17
bars	friend	0.017	1.50
legal	friend	0.135	1.40
jobs	minimum	0.076	-0.50
jobs	stranger	0.055	1.67
jobs	employer	0.017	-2.33
dating	employer	0.140	-0.71
sexual	maximum	0.111	-0.25
alcohol	stranger	0.022	0.80
alcohol	employer	0.104	-0.40
alcohol	friend	0.096	0.50
alcohol	computer	0.104	-0.40
gambling	stranger	0.023	0.78
gambling	employer	0.035	-1.33
gambling	friend	0.108	1.00

In contrast, our participants seemed to be more comfortable when correctly represented in front of friends and strangers. None of our results regarding computer algorithms were significant.

DICUSSION

Answers to Free Response Questions

In the second part of the survey, participants were asked to respond to statements of type "Rate how comfortable or uncomfortable you would be if [type of observer] saw you receive an ad in [category name]." After each of these questions, we also asked participants to give a reason for their rating of comfort. As shown in Comfort Based on Profiling Accuracy in the Pilot Results section, for some categories such as "Alcohol," "Gambling," and "Job Listings," participants indicated different levels of comfort with significant p-values when they were accurately profiled as opposed when they were inaccurately profiled. These changes, however, were only sometimes reflected in free response answers. They were reflected in some responses, for example, when responding to category “Gambling” and observer type “employer,” P18 was “somewhat interested” in gambling and said “I wouldn't want them to think I have a gambling problem” when they were accurately profiled, whereas they said “my employer won't care about a random advertisement that i show no interest in” when they were inaccurately profiled. However, in “Job Listings” category with observer type “employer,” most free response answers of the participants stayed the same or very similar even though they changed their comfort ratings when profiled accurately as opposed to inaccurately. Most participants responded that they would not want their employers to be around when they received an ad in this category no matter how interested they were in “Job Listings.” For example P7 said that they were “very interested” in this category, picked “extremely uncomfortable” when the observer was an employer when they were accurately profiled, but picked “somewhat uncomfortable” when they were inaccurately profiled. However, the reason they presented was “I wouldn't want my employer to know I was job hunting” when accurately profiled and “I wouldn't want my employer to think I am looking for other jobs” when inaccurately profiled. In a case like this, we were not able to learn the reason behind the slight change in the participant’s decision. Thus, in a future survey, instead of asking a free response question after every rating question, asking a free response question only if the participant rating changed across the accuracy of profiling would be better. This would also make the survey less dry as some of the participants put the exact same answers under some of the free response questions.

Listed Products and Participant Reactions

For each of five categories participants were introduced in the second part of the survey, they were asked to list a product or a service from that category. Variation in listed products by participants for some of the categories is considerable. For instance in Gambling & Betting, P2 listed “lottery” as a product while some other respondents listed “online casinos”. When asked about how comfortable they would feel if an employer saw them receiving an ad in this category, P2 responded “I wouldn't feel bad or uncomfortable as a lot of people gamble” whereas P6, who listed “casinos” as a product, responded “I wouldn't want my employer to see this” to the same question. We had similar responses under the Sexual and Reproductive Clinics and Medication, where P5 listed “viagra” as a product and when asked about how comfortable they would feel if a family, close friend, or significant other saw them receive an ad in this category, they responded “neither comfortable nor uncomfortable” and presented the following reason: “they would avoid the conversation”. However, P13, who listed “planned parenthood” under the same category responded “somewhat uncomfortable” and said “They might wonder why such ads pop up on my screen, so it might raise questions I was obliged to answer.” Thus, to get more accurate results of participant reactions to the personalized ads, a next step could be either using more specific ad interest categories, especially for sensitive categories, or presenting participants with product names instead of category names; however the latter option could also make it harder to generalize the results to broader ad-interest categories.

Limitations

As this study is meant to be a pilot study to showcase the need for further research in this arena and introduce potential methodologies for further research, the scope of the study was limited in size. As such, we only conducted the study on 20 participants. This small sample size means that our pilot results and analysis are not statistically robust and should not be used definitively, but rather, as a starting point to inform future research. Furthermore, again due to the nature of a pilot study, we were not able to question participants on the entirety of Google AdSense’s list of topics. A more expansive study on this topic should examine participants beliefs on the entire list of ad categories in order to measure the accuracy of Google’s definition of sensitive and standard categories. Furthermore, because Google is one of the leading digital advertising publishers, we used only used Google’s list of categories to inform the categories tested in this study. Further research should examine the ad categories used by other advertising agencies and publishers, as there may be differences in what categories are used to describe users and which categories are considered sensitive or nonsensitive.

There are also several methodology choices made during this study that we would recommend changing or nuancing in future research. We would recommend randomizing the order in which sections 2 (sensitivity of ad categories) and 3 (privacy dashboard) were shown to participants to prevent priming participants to think of the ad categories from a privacy perspective. We would recommend altering the wording of the privacy dashboard to match the exact wording of Google or other advertisers’ privacy dashboards to create an even greater sense of realism.

CONCLUSION

With the results and limitations of our study in mind, we are slightly better-equipped to approach the classification of sensitivities of advertisement categories.

Using Google’s existing AdSense categories as a model, we found that the types of products which are contained in even the most specific subcategories are still experienced with varying degrees of user comfort. While hypothetical advertisements for Viagra and Planned Parenthood elicited different responses within our participant pool, they both fall under the same sensitive category. Perhaps this is a sign that advertisement categories should be classified more granularly.

We found that not all sensitive categories are created equal. In general, our participants did not seem to be uncomfortable when presented with advertisements with alcohol and gambling—even though these are categories which Google AdSense hides by default. Rather, participants felt more discomfort when being presented with advertisements within conventionally-nonadult categories, such as *legal services* and *apartments*. This would point toward a need to reclassify other categories as sensitive, or to be more precise in the meaning and handling of *sensitivity*.

It would also be prudent for advertisers to be informed of a certain "contextual sensitivity", where advertisements may be experienced as uncomfortable only within certain settings. Our participants tended to have much stronger reactions to a wider range of ad categories when observable by their employers, for example. Perhaps advertisements should have a "safe-for-work" option, which temporarily hides advertisements for categories such as *Nightlife*, *Legal Services* and *Jobs*; and perhaps, even further, advertisers should allow preference control panels that change based on user location. A *do not show me ads like this on this computer* setting might have some use.

REFERENCES

[1] Mark S. Ackerman, Lorrie F. Cranor, and Joseph Reagle. 1999. Privacy in E-commerce: Examining user scenarios and privacy preferences. In *Proceedings of the 1st ACM Conference on Electronic Commerce, EC 1999* (pp. 1-8). DOI: 10.1145/336992.336995

[2] Lalit Agarwal, Nisheeth Shrivastava, Sharad Jaiswal, and Saurabh Panjwani. 2013. Do Not Embarrass: Re-Examining User Concerns for Online Tracking and Advertising. In *Symposium On Usable Privacy and Security 2013*. DOI: 10.1145/2501604.2501612

[3] Keith Bradley, Rachael Rafter, and Barry Smyth B. 2000. Case-Based User Profiling for Content Personalisation. I*n Adaptive Hypermedia and Adaptive Web-Based Systems* and *Lecture Notes in Computer Science, vol 1892*.

[4] Joaquin Delgado, Richard Davidson, and Triplehop. 2002. Knowledge Bases and User Profiling in Travel and Hospitality Recommender Systems. In *ENTER 2002*.

[5] Ari Juels. Targeted Advertising ... and Privacy Too. 2001. In *CT-RSA 2001 Proceedings of the 2001 Conference on Topics in Cryptology: The Cryptographer's Track at RSA pp. 408-424.*

[6] Aleecia McDonald and Lorrie F. Cranor. 2010. Beliefs and Behaviors: Internet Users' Understanding of Behavioral Advertising (August 16, 2010). In *TPRC 2010*. Available at SSRN: https://ssrn.com/abstract=198909

[7] Stuart E. Middleton, Nigel R. Shadbolt, and David C. De Roure. 2003. Capturing interest through inference and visualization: ontological user profiling in recommender systems. In *Proceedings of the 2nd international conference on Knowledge capture*.

[8] Andres Moreno, Harold Castro, and Michel Riveill. Decentralized Recommender Systems for Mobile Advertisement. 2011. In *Workshop on Personalization in Mobile Applications, ACM New York.*

[9] Joseph Phelps, Glen Nowak, and Elizabeth Ferrell. 2000. Privacy Concerns and Consumer Willingness to Provide Personal Information. J*ournal of Public Policy & Marketing: Spring 2000, Vol. 19, No. 1, pp. 27-41.*

[10] Edith G. Smit, Guda Van Noort, and Hilde A.M. Voorveld. 2014. Understanding online behavioural advertising: User knowledge, privacy concerns and online coping behaviour in Europe. *Computers in Human Behavior, Volume 32, March 2014, Pages 15-22*. ISSN 0747-5632, http://doi.org/10.1016/j.chb.2013.11.008.

[11] William E. Spangler, Kathleen S. Hartzel, and Mordechai Gal-or. 2006. Exploring the Privacy Implications of Addressable Advertising and Viewer Profiling. In *Communications of the ACM May 2006 , Vol. 49 Issue 5, pp. 119-123.* DOI:10.1145/1125944.1125951

[12] Mirco Speretta and Susan Gauch. 2005. Personalized Search Based on User Search Histories. In *IEEE/WIC/ACM International Conference on Web Intelligence*.

[13] Eran Toch, Yang Wang, and Lorrie Faith Cranor. Personalization and Privacy: A Survey of Privacy Risks and Remedies in Personalization-Based Systems. 2010. *User Model User-Adap Inter (2012) 22: 203*. DOI: 10.1007/s11257-011-9110-z

[14] Vincent Toubiana, Arvind Narayanan, Dan Boneh, Helen Nissenbaum, and Solon Barocas. 2010. Adnostic: Privacy Preserving Targeted Advertising. In *NDSS 2010*. https://crypto.stanford.edu/adnostic/adnostic.pdf.

[15] Trefis Team contributors. U.S. Digital Advertising Landscape And Key Players (Part 2). *Forbes*. https://www.forbes.com/sites/greatspeculations/2015/07/09/u-s-digital-advertising-landscape-and-key-players-part-2/#152e32bb9ef7. (09 July 2015).

[16] Blase Ur, Pedro G. Leon, Lorrie F. Cranor, Richard Shay, and Yang Wang. Smart, useful, scary, creepy: Perceptions of online behavioral advertising. In *SOUPS 2012*.

[17] Jeffrey Zaslow. 2002. If TiVo Thinks You Are Gay, Here's How to Set It Straight. *Wall Street Journal.* https://www.wsj.com/articles/SB1038261936872356908. (November 26, 2002).

[18] Philip Zigoris and Yi Zhang. 2006. Bayesian Adaptive User profiling with Explicit & Implicit Feedback. In *CIKM'06*.

APPENDIX

Category	Products
Alcohol	Jagermeister; coors; Vodka; Beer; Budweiser; beer; Maker's Mark bourbon; Beer; Miller lite; Beer
Apartments & Residential Rentals	zillow; community park; House for rent; an apartment finding service, such as apartmentfinder.com; apartment blocks; Apartment
Bars, Clubs & Nightlife	O'Reily's; Nightclub; Budweiser; Any given number of local bars, clubs, or restaurants in the area near me -- Ruth Chris is one that specifically comes to mind; Discount on bottle service; Happy Hours at a Bar
Birth Control	Depo Provera; trojan; condom; Abortion Clinics, Planned Parenthood, Various Pharmaceutical Companies that manufacture birth control or morning-after pills; condoms; This product is not a representation of my character.
Boats & Watercraft	Life Vests; jet ski; boston waterways; jet ski; Ski-Doo jet skis; Speedboats
Dating	tinder; Tinder; OkCupid; dating website such as plenty of fish or tinder; matchmaking services; Dating service (Match.come, etc); Clothing; match.com
Fashion & Style	Shoes; pants; ad for a department store, such as Macy's; Clothes; nike shoes
Gambling & Betting	The Lottery; fantasy sports; foxwoods casino; Casinos; Casinos, Indian Reservations, Gambling Websites; Ad for an online or physical casino. I don't know a specific example since I don't pay attention to these services; poker websites; Online Poker; pokerstars; Poker game online
Hiking & Camping	a tent; rei; Sleeping bag; Sequoia National Park Campground; Tents; hiking boots; Tents
Job Listings	monster; Monster job listings; indeed.com; Indeed.com, Glassdoor.com indeed.com; monster; Career Builder
Legal Services	lawyers; Divorce lawyer; Divorce attorney; Local Personal Injury Attorneys, Debt Relief Services, Tax Attorneys; Child Support; Divorce Lawyer
Mobile & Wireless Accessories	Apple Watch; earphones; Cellphone Charger; Otter Box cases; cell phone plans; Phone Case
Pet Food & Supplies	chewy website; Taste of the Wild dog food.; advertisements for pet stores, like PetSmart; Alpo Dog Food; Dog collars; Cat food
Scientific Equipment	a microscope; shopping; Telescope; staples.com; microscopes
Sexual and Reproductive Clinics and Medication	Viagra; viagra; Planned Parenthood; Planned Parenthood; Birth Control; prescription medication discount service

Figure 7 – Participant-given examples for each advertisement categories

Contingency Table for Sensitivity Classes

	Show	D_Care	D_Show
General	37	72	91
Sensitive	4	12	44
Restricted	3	12	25

Pearson's Chi-Square Test

X-squared = 17.204, df = 4,
p-value = 0.001764

Expected Values

	Show	D_Care	D_Show
General	29.33	64.0	106.67
Sensitive	8.80	19.2	32.00
Restricted	5.87	12.8	21.33

Chi-Square Test Residuals

	Show	D_Care	D_Show
General	1.42	1.00	-1.52
Sensitive	-1.62	-1.64	2.12
Restricted	-1.18	-0.22	0.79

Contingency Table for Categories

	Show	D_Care	D_Show
Alcohol	2	6	12
Apartment & Residential Rentals	4	7	9
Bars, Clubs & Nightlife	2	7	11
Birth Control	2	5	13
Boats & Watercraft	0	8	12
Dating	2	3	15
Fashion & Style	7	7	6
Gambling and Betting	1	6	13
Hiking & Camping	5	6	9
Job Listings	4	5	11
Legal Services	0	8	12
Mobile & Wireless Accessories	4	10	6
Pet Food & Supplies	8	7	5
Scientific Equipment	3	7	10
Sexual & Reproductive Clinics & Medication	0	4	16

Pearson's Chi-squared test

X-squared = 48.398, df = 28, **p-value = 0.009706**
Warning message: In chisq.test(tbl) : Chi-squared approximation may be incorrect

Expected Values

	Show	D_Care	D_Show
Alcohol	2.93	6.4	10.67
Apartment & Residential Rentals	2.93	6.4	10.67
Bars, Clubs & Nightlife	2.93	6.4	10.67
Birth Control	2.93	6.4	10.67
Boats & Watercraft	2.93	6.4	10.67
Dating	2.93	6.4	10.67
Fashion & Style	2.93	6.4	10.67
Gambling and Betting	2.93	6.4	10.67
Hiking & Camping	2.93	6.4	10.67
Job Listings	2.93	6.4	10.67
Legal Services	2.93	6.4	10.67
Mobile & Wireless Accessories	2.93	6.4	10.67
Pet Food & Supplies	2.93	6.4	10.67
Scientific Equipment	2.93	6.4	10.67
Sexual & Reproductive Clinics & Medication	2.93	6.4	10.67

Chi-Square Test Residuals

	Show	D_Care	D_Show
Alcohol	-0.54	-0.16	0.41
Apartment & Residential Rentals	0.62	0.24	-0.51
Bars, Clubs & Nightlife	-0.54	0.24	0.10
Birth Control	-0.54	-0.55	0.71
Boats & Watercraft	-1.71	0.63	0.41
Dating	-0.54	-1.34	1.33
Fashion & Style	2.37	0.24	-1.43
Gambling and Betting	-1.13	-0.16	0.71
Hiking & Camping	1.21	-0.16	-0.51
Job Listings	0.62	-0.55	0.10
Legal Services	-1.71	0.63	0.41
Mobile & Wireless Accessories	0.62	1.42	-1.43
Pet Food & Supplies	2.96	0.24	-1.74
[illegible]	[illegible]	[illegible]	[illegible]

Figure 8 – Data relating Category to Dashboard Interest

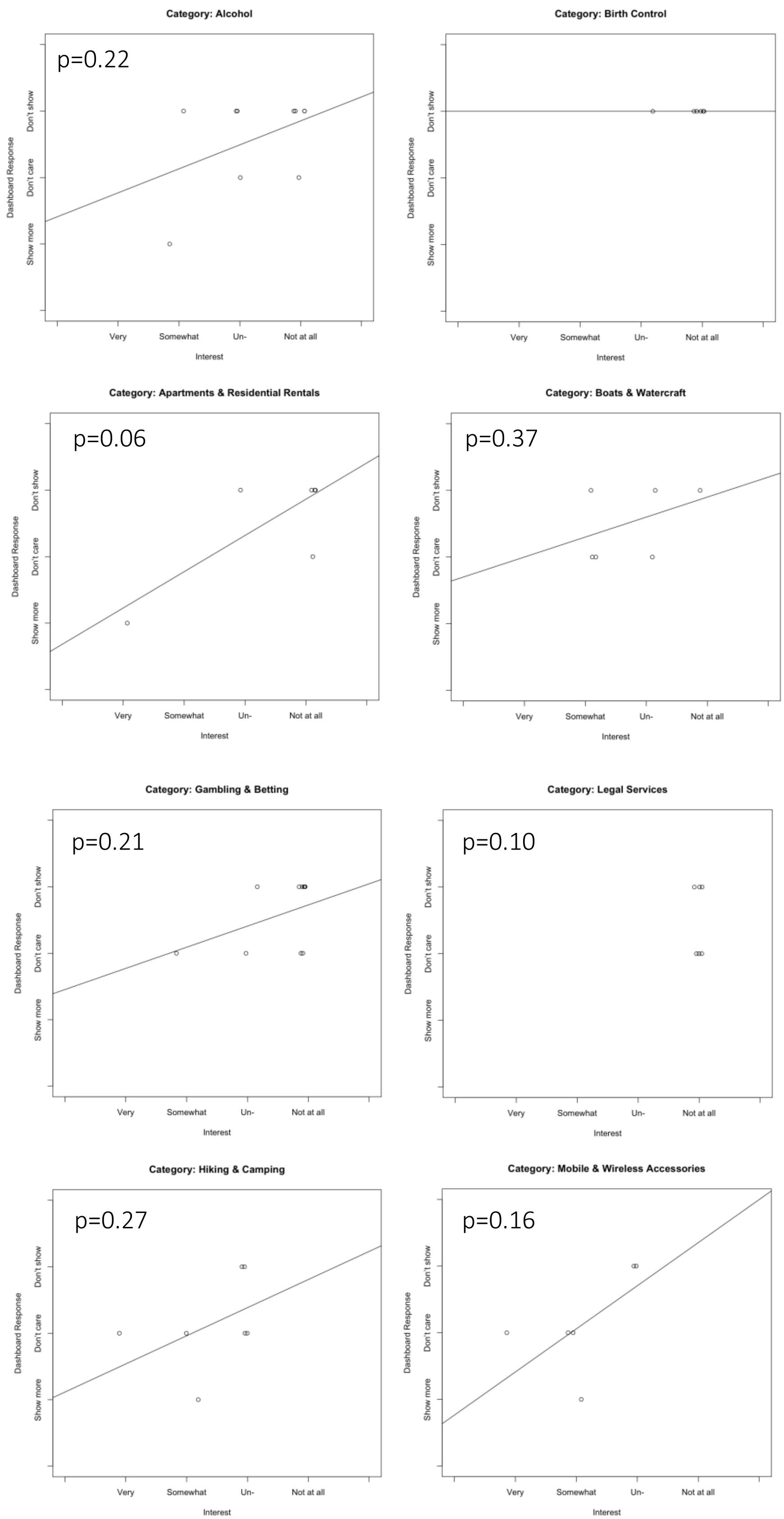
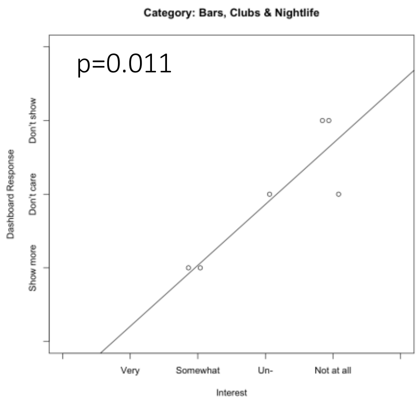
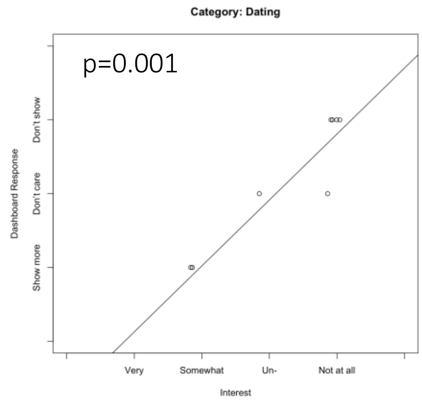
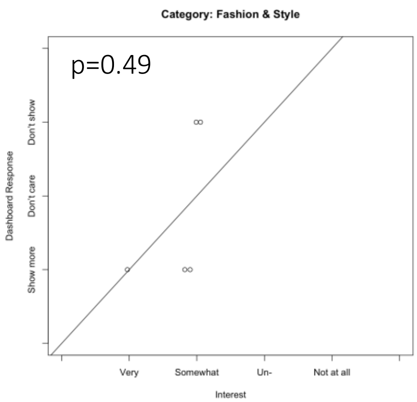
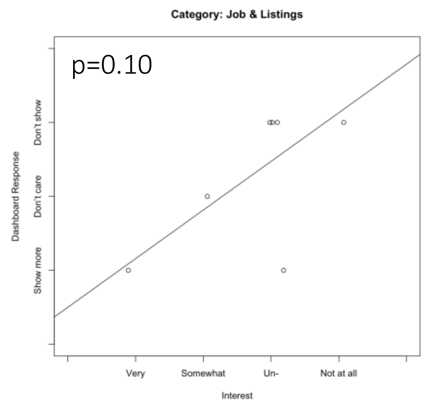


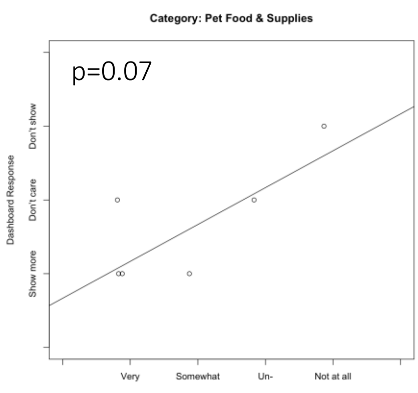
Figure 9 – Linear regressions relating user interest to dashboard response

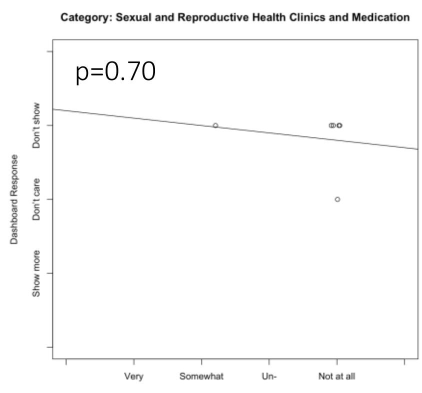

Category: Bars, Clubs & Nightlife
p=0.011
Dashboard Response
Don't show
Don't care
Show more
Very
Somewhat
Un-
Not at all
Interest

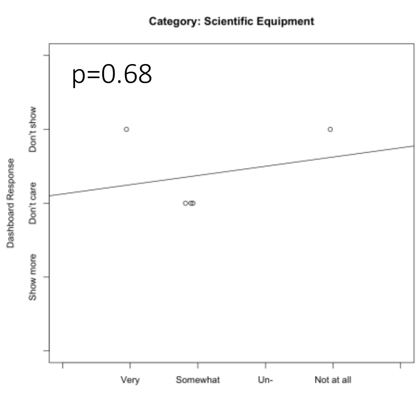

Category: Dating
p=0.001
Dashboard Response
Don't show
Don't care
Show more
Very
Somewhat
Un-
Not at all
Interest


Category: Fashion & Style
p=0.49
Dashboard Response
Don't show
Don't care
Show more
Very
Somewhat
Un-
Not at all
Interest


Category: Job & Listings
p=0.10
Dashboard Response
Don't show
Don't care
Show more
Very
Somewhat
Un-
Not at all
Interest


Category: Pet Food & Supplies
p=0.07
Dashboard Response
Don't show
Don't care
Show more
Very
Somewhat
Un-
Not at all
Interest


Category: Sexual and Reproductive Health Clinics and Medication
p=0.70
Dashboard Response
Don't show
Don't care
Show more
Very
Somewhat
Un-
Not at all
Interest


Category: Scientific Equipment
p=0.68
Dashboard Response
Don't show
Don't care
Show more
Very
Somewhat
Un-
Not at all
Interest