Analyzing the Impact and Accuracy of Facebook Activity on Facebook's Ad-Interest Inference Process

AAFAQ SABIR[*], North Carolina State University, USA
EVAN LAFONTAINE, North Carolina State University, USA
ANUPAM DAS, North Carolina State University, USA

Social media platforms like Facebook have become increasingly popular for serving targeted ads to their users. This has led to increased privacy concerns due to the lack of transparency regarding how ads are matched against each user profile. Facebook infers user interests through their activities and targets ads based on those interests. Although Facebook provides explanations for why a particular interest is inferred about a user, there is still a gap in understanding what activities lead to interest inferences and the extent to which the sentiment or context of activities is considered in inferring interests.

To obtain insights into how Facebook generates interests from a user's Facebook activities, we performed controlled experiments by creating new accounts and systematically executing numerous planned activities. This enabled us to make *causal inferences* about activities that lead to generating specific interests, many of which were not representative of actual user preferences. We also evaluated which activities resulted in interests and found that very naive activities, such as only viewing/scrolling through a page, lead to an interest inference. We found 33.22% of the inferred interests were inaccurate or irrelevant. We further evaluated the interest inference explanations provided by Facebook and found that these explanations were too generalized and, at times, misleading. To understand if our findings hold for a large and diverse sample, we conducted a user study where we recruited 146 participants (through Amazon Mechanical Turk) from different regions of the world to evaluate the accuracy of interests inferred by Facebook. We developed a browser extension to extract data from their own Facebook accounts and ask questions based on such data. Our participants reported a similar range (29%) of inaccuracy as observed in our controlled experiments. We also found that most of our participants were unaware of the availability of Facebook's ad preference manager, interest inference process, and even interest explanations.

CCS Concepts: • **Security and privacy** → *Social aspects of security and privacy*.

Additional Key Words and Phrases: Facebook Ads; Targeted Ads; Causal Inference

ACM Reference Format:
Aafaq Sabir, Evan Lafontaine, and Anupam Das. 2022. Analyzing the Impact and Accuracy of Facebook Activity on Facebook's Ad-Interest Inference Process. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW1, Article 76 (April 2022), 34 pages. https://doi.org/10.1145/3512923

[*]Part of this work was done when lead author was at Lahore University of Management Sciences.

Authors' addresses: Aafaq Sabir, asabir2@ncsu.edu, North Carolina State University, Raleigh, North Carolina, USA; Evan Lafontaine, elafont@ncsu.edu, North Carolina State University, Raleigh, North Carolina, USA; Anupam Das, anupam.das@ncsu.edu, North Carolina State University, Raleigh, North Carolina, USA.

2573-0142/2022/4-ART76 $15.00
https://doi.org/10.1145/3512923

1 INTRODUCTION

Online advertising has become a significant source of revenue for many social media platforms. For instance, Facebook generates nearly 98.5% of its revenue by displaying ads on its users' homepage [33]. These ads are targeted based on interest profiles that Facebook creates for each user [9]. Given Facebook's extremely large user base of approximately 2.6 billion active users [22] (along with 854.5 million active Instagram users [23]), targeted ads provide a broader reach for advertisers [28]. Furthermore, over 10 million advertisers utilize Facebook's targeted advertising platform [10], many of which are small and medium-sized businesses that are directly dependent on advertising. However, behavioral-based targeted advertising raises many privacy concerns. Several studies have found that users are uncomfortable with pervasive tracking and are concerned about the lack of transparency surrounding these practices [18, 36, 42]. For instance, a study by Pew found that 50% of Internet users are concerned about the amount of personal information that is publicly available online [37].

To ease users' privacy concerns, Facebook provides an Ad Preference Manager (APM) [8] that allows users to view and edit their interest profiles, as well as obtain an explanation for each displayed sponsored ad. In recent years, researchers have compared the accuracy of interests inferred by social media platforms, such as Facebook and Google [20], using data obtained through platform-specific APMs [8, 11] (also known as ad settings). Venkatadri et al. [45] analyzed the coverage and accuracy of offline data brokers, including Acxiom and Experian, that posted ads on Facebook utilizing the "Partner Categories" feature. However, Facebook removed the "Partner Categories" option in March 2018 [4] and now serves ads based on interests inferred by Facebook itself. The details of such inference algorithms, however, remain a black box. The implications of inferring inaccurate interests on one of the largest social media platforms have both economic (effectiveness of paid ads) and privacy (inaccurate data sharing across platforms) ramifications.

In this paper, we focus on answering the following research questions: **RQ1:** ***How does Facebook infer user interests?*** Existing literature has not focused on obtaining insights into how user interests are captured through their activities on Facebook. We perform *controlled experiments*, where we create new Facebook accounts and perform planned activities (e.g., liking and commenting on posts or pages) to make *causal inferences* between inferred interests and different online activities. **RQ2:** ***How accurately does Facebook infer interests from user activities?*** Through controlled experiments, we determine the extent to which inferred interests relate to our preplanned activities. To understand if our findings hold for geographically diverse users, we perform a user study where we recruit participants from different parts of the world through Amazon Mechanical Turk (MTurk). The user study utilizes a browser extension to extract relevant data from each participant's ad profile and dynamically generates survey questions. **RQ3:** ***Does Facebook accurately explain how inferred interests are derived?*** We analyze the explanations provided by Facebook's APM regarding the interests inferred about users. We use data from both our controlled accounts and user-study participants to determine the accuracy and effectiveness of explanations.

In summary, we make the following key contributions:

- We conducted controlled experiments by creating new Facebook accounts to uncover how Facebook infers ad interests for users based on their Facebook activities. We performed different activities for each new account, such as liking pages or posts, reacting to posts, posting comments, and scrolling through posts. We further divided these activities into positive and negative *sentiments* to discover if the interest inference process accounts for sentiment or context. We found that Facebook does not take sentiment into account while inferring interests — something that can potentially lead to incorrect inferences. We believe that this is the first attempt to shed

light on the reasons behind the inaccuracies found in interest profiles for Facebook, as prior studies have highlighted such inaccuracies, but none have attempted to determine its root cause.
- We performed a user study regarding the accuracy (i.e., relevancy) of Facebook ads and interest profiles. We recruited 146 participants from the US (52), Europe (45), and India (49) using Amazon Mechanical Turk. We developed a Chrome browser extension to collect data from participants' own Facebook accounts and posed questions about the accuracy/relevancy of ads and inferred interests. We also highlight cross-regional differences.
- We found that the explanations describing why a particular interest was inferred about a user are vague and at times misleading. We evaluated the accuracy of interest explanations by performing controlled activities and mapping the inferred interests to an activity performed. We then compared the provided explanations for every interest with our ground-truth interest explanations.
- We evaluated participants' awareness with Facebook advertisements and ad explanations. We found that 52.7% of the participants were unaware that they could view the reason for every ad shown on Facebook and that 65.8% of the participants had never visited their interest profile. Only 38.3% of the participants were satisfied with their experience with Facebook ads. 58.8% of the participants who were not satisfied with ads were unaware that they could view ad reasons, suggesting that ad transparency features are not easily noticeable.

The remainder of this paper proceeds as follows. Section 2 provides background information about Facebook APM and ad publishing dashboard (Ad Manager). Section 3 describes related work. Section 4 describes our data collection and analysis methodology. Section 5 contains the analysis about how Facebook infers user interests from activities. Section 6 analyzes the accuracy of the inferred interests. Section 7 evaluates the completeness of explanations of inferred interests by Facebook. Section 8 provides some recommendations to improve the transparency of the interest inference process. Section 9 discusses limitations of our work. We conclude in Section 10.

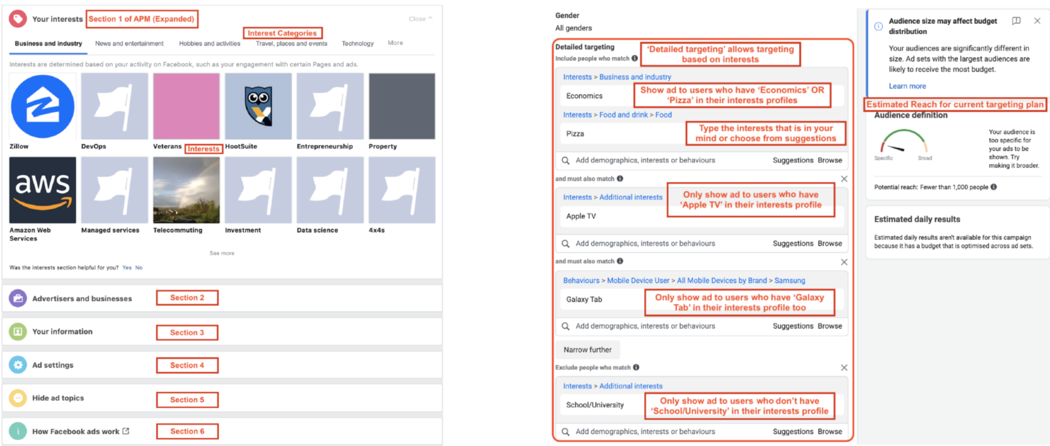
2 BACKGROUND

Facebook allows users to review their interests through the Ad Preference Manager (APM). The APM is a dashboard containing information typically utilized for targeted ads (as shown in Figure 1a). This dashboard enables users to view any interest that Facebook has inferred. The APM contains the following sections:[1]

(1) **Your Interests:** This section includes all of the interests that Facebook has inferred about a user, divided into different categories. Each interest contains an explanation about why a given interest has been inferred. The user can also remove wrongly-inferred interests from the list.
(2) **Advertisers and Business:** This section contains advertisers who have uploaded the user's contact information, such as phone number or email address, to target ads. It also includes a list of sites visited, ads clicked, and ads hidden by a user.
(3) **Your Information:** This section allows the user to choose if Facebook can target ads based on "Relationship status," "Employer," "Job Title," or "Education." It also contains another section that lists behavior categories, such as "Frequent traveler" or "Owns iPhone X."
(4) **Ad Settings:** The user may allow or block targeted ads based on Facebook data partners. This section includes information gathered from partnering websites and apps, as well as specific offline interactions. For instance, certain purchases or ads based on a user's activity from visiting websites or using apps (not including Facebook products) may appear in this information.

[1]Facebook's APM UI changed throughout our study, but we found no significant difference through our controlled experiments.

(5) **Hide Ad Topics:** Allows users to view fewer ads about sensitive topics, such as "Alcohol" or "Politics."
(6) **How Facebook ads work:** This segment provides details about Facebook's targeted advertisements.



(a) Snapshot of Facebook APM

(b) Snapshot of Facebook Ad Publishing Dashboard

Fig. 1. Overview of different information available under Ad Preference Manager and Ad publishing dashboard on Facebook.

Advertisers can use the Ad Publishing dashboard provided by Facebook to serve targeted ads to users with specific interests (as shown in Figure 1b). The advertiser can select specific demographics, such as age group, location, or gender. Facebook also provides advertisers with a "Detailed Targeting" option, where advertisers can target ads based on user interests selected through a pre-populated list. Furthermore, this dashboard provides an estimated audience size, allowing the advertiser to expand or narrow the interests to reach the desired audience pool. The Ad Publishing dashboard regulates ads for all Facebook-owned products, such as Instagram and Facebook Messenger.

3 RELATED WORK

Social networks are an important subject of investigation due to their large user base and rich ad ecosystem. Many studies have found inconsistencies between the information disclosed to advertisers and the data social media platforms would collect about users. For example, Datta et al. [24], and Wills et al. [46] have found Google ad profiles to be incomplete because the information provided in targeted ads was not listed in users' ad profiles. These omissions also make it challenging to analyze APMs [19, 29, 34]. Similarly, a study by Andreou et al. [17] revealed a discrepancy between what advertisers could observe and what users could view in their ad profiles. To provide improved transparency for targeted ads, Venkatadri et al. [44] proposed a transparency mechanism that can force online advertising platforms to disclose information to targeted users fully.

Researchers have also investigated if sensitive information is used for targeted ads. Andreou et al. [16] provide insights into how advertisers are using the Facebook ad ecosystem. Their results reveal that a wide range of advertisers targets users (e.g., from popular to niche advertisers) and that a non-negligible fraction of advertisers target potentially sensitive information. Cabañas et al. [21] also found that Facebook was targeting ads based on attributes that were deemed sensitive. Recently, Venkatadri et al. [43] found that Facebook uses Personally Identifiable Information (PIIs) obtained through two-factor authentication for advertising.

Furthermore, studies have highlighted the accuracy of social media platforms capturing user interests. For instance, Venkatadri et al. [45] investigated offline data brokers (known as "Partner Categories") that Facebook previously used for advertising. They found that a surprisingly large percentage of Facebook accounts (e.g., above 90% in the US) are linked to data broker information. Moreover, by running controlled ads to 183 crowd-sourced US-based volunteers, they found that at least 40% of the user attributes sourced from data brokers were not accurate. It is important to note that Facebook stopped using "Partner Category" attributes since October 2018 due to rising privacy concerns [4]. Bashir et al. [20] compared four different APMs (Facebook, Google, Blukai, and Exelate) in terms of both quantity and quality, and found that ad profiles do not have significant overlaps; however, each service successfully captured many of the participants' interests. Others have looked at the efficacy of explanations provided by different platforms. Eslami et al. [27] evaluated ad explanations by manually showing 32 participants their ad explanations, and then interviewed them about the explanations provided. They found that users preferred interpretable, non-invasive explanations as well as a recognizable link to their ad profile page.

Ali et al. [15] found that Facebook's targeted ad delivery process can "skew" ad delivery in ways that the advertisers may not intend. They demonstrate that such skewed delivery occurs on Facebook due to market and financial optimization effects and the platform's own predictions about the "relevance" of ads to different groups of users. Such discoveries motivate the need to investigate the interest inference process, which becomes the driving force behind the ad delivery process targeting specific users.

Studies also found discrepancies such as gender bias in Facebook's ad delivery process [15, 31, 40]. Ali and Imana et al. focused on finding Facebook ad delivery bias when ads are targeted to *general audience*, such as audience based on only location or age, rather than specific attributes such as interests, and found that Facebook ad delivery is skewed even when the chosen audience by advertisers contains no bias [15, 31]. Speicher et al. found that even though Facebook does not allow gender selection when running certain ads such as financial and employment-related ads, advertisers can still target specific gender by abusing free-form attributes (interests) [40].

Distinction from prior works. This work contributes to the current field by studying the causal inferences of activities that lead to generating specific user interests on Facebook. Previous literature has researched this process to provide details on how sensitive information is collected [21], yet our work is the first to assess the algorithm itself and provide suggestions on improving transparency and usability. Furthermore, the clarity of the explanations provided for a given interest appearing in one's profile remains unstudied (others have studied the "Why am I seeing this ad?" option shown with a displayed ad, but not the interests that have been inferred). While Bashir et al. [20] have analyzed the accuracy of inferred interests, our work is different as it uses controlled accounts and performs planned activities to derive causal relationship between inferred interests and the planned activities. Existing studies have also showcased inaccuracies of interests inferred by Facebook, yet none attempt to find the root cause of it. We believe this is the first attempt to shed light on the reasons behind such inaccuracies. We also collect data from participants from different geographical regions to contrast any difference in their perception of accuracy of the inferred interests. Again, while Andreou et al. [17] partially analyzed the explanations provided for inferred interests, they failed to point out that explanations are overly generalized and at times misleading. They leave comprehensive controlled experiments as future work. While prior works have found that Facebook ad delivery is biased even when the advertiser intentionally does not target any specific population [15, 31] or that interests can be maliciously exploited to target a certain gender [40], our study focuses on the inference of interests and not the ad delivery process itself. We show that there is no difference in the inferred interests by Facebook among groups with different demographics. However, the way ads are delivered may have a bias that is orthogonal to our scope. To the best of our knowledge, we are the first to rigorously evaluate Facebook's interest inference process by studying the effect of controlled activities and sentiments on inferred interests.

4 DATA AND METHODOLOGY

In this section, we describe our data sources and the methods we used to analyze the data.

4.1 Data Sources

To gauge how Facebook infers user interests from behavioral data and the extent to which such inferences are accurate, we collect data through two methods: 1) controlled experiments creating customized user accounts in Facebook, and 2) recruiting participants through a user study. Details of our data collection methodology follow.

4.1.1 Controlled Facebook Accounts. We created Facebook accounts to analyze how an interest profile grows as we interact with different Facebook content. Our account setup process and the types of activities conducted on different accounts are described below.

Account Setup. We performed two iterations of controlled experiments. One iteration was conducted in January 2020, and the second iteration was conducted in May 2021. The first iteration consisted of four accounts with the same demographic information: 22 years old, male, and based in Pakistan. This iteration allowed us to compare interest inferences between accounts with similar demographics to correctly map which interests were formed from what actions without worrying about any other factors. The second iteration consisted of ten new Facebook accounts. This iteration was performed to identify whether differences in interest inferences occurred based on the account's demographic information. Eight of these accounts were created solely for positive activities, four in the United States and four in Pakistan. We then differed the demographic information for each account (as shown in Table 5). The last two accounts were used to perform a mixture of positive and negative activities to determine if the type of actions would have an effect on the interest inference. We used separate virtual machines (VMs), and different newly-created Gmail accounts for each controlled account to reduce the chance of interference across the Facebook accounts. We also provided valid phone numbers with each account; new phone numbers were purchased to ensure that those numbers were clean and were never previously used on Facebook or any online platform. Additionally, the VM sessions were only used to visit Facebook.

Account Activities. For single-activity accounts (the first iteration) we perform one type of action, yet for the second iteration combined different activities. To test how specific types of activities influence one's interest profile, we restricted activities to the following four categories:

- **Baseline:** No activities were performed in baseline, as it served as a control. This allowed us to determine differences in interest profiles when compared to other accounts.
- **Positive interaction:** This account was used to determine if Facebook inferred interests based on positive interactions with existing Facebook content. We performed four types of positive interactions (summarized in Table 1) in separate phases and recorded the interest profile for each phase — no two types of activities were performed simultaneously on a single account.
- **Negative interaction:** This account tested if Facebook considers the semantics of any interaction, i.e., distinguishing between positive versus negative reactions. Two types of negative interactions were performed (details in Table 1) in separate phases.

Table 1. Descriptions of activities per control account.

Account	Activities	Description of activities
Baseline	None	No specific activities were performed
Positive interaction	Page like	Liking a page without interacting with its posts
	Post Like/Love react	Like or Love react on posts without liking page
	Comments	Posting positive comments without liking the page
	Scroll only	Just scrolling page posts without liking page or its posts
Negative interaction	Angry react	Posting 'Angry' react on some page posts without liking the page
	Comments	Posting negative comments on posts without liking the page
Interaction with friend	None	Being friends with the Positive Interaction account; no other activity was performed
	Send messages	Sending messages on messenger to the friend

- **Interaction with friend:** This account analyzed whether being friends with a person affects one's interest profile. We tested two types of interactions (summarized in Table 1) in separate phases.

Specific Facebook content (e.g., a page) was visited to perform the activities. We grouped the visited content into very specific categories (representing a topic) and ensured that each category was unique and unrelated to others. All of the daily activity was solely related to content from the chosen categories. For instance, if we chose the category "Clothing," we would visit pages relating to clothing, such as "Breakout Clothing." Furthermore, the visited content was chosen prior to performing activities on any controlled accounts; a separate Facebook account was used to search for content for the daily interactions. This ensured that the controlled accounts were not contaminated with extra interests and remained consistent with visited pages between all accounts.

We discovered the relevant Facebook pages by searching for predefined *topics* on Facebook; we then picked the top five results that had relevant Facebook pages. Since the content of the pages we interact with can affect the inference algorithm, we attempted to minimize this in our experiments by choosing the set of pages that did not have a mixture of positive and negative vibe about a given topic. For example, we interacted with a 'Dog Lovers' page containing all positive content about dogs. Furthermore, activities were performed on the same set of pages for both the positive and negative interaction accounts. The same amount of activities were also performed on each page from both accounts. For example, if the positive account liked N posts on page 'X', the negative account performed angry reactions on the same N posts on page 'X'. This methodology was also utilized when writing positive and negative comments. In order to map a performed activity to an interest, each type of activity was done on a separate set of pages (i.e., different content). For example, we performed page likes on certain pages and post likes (i.e., liking posted comments) on entirely unrelated pages. This allowed us to determine the causal effect of each activity. The full list of Facebook content and categories visited is provided in Tables 2 and 3. We continued our controlled activities for about 40 days; we visited 2 or 3 pages daily while performing 8 to 10 activities on each page. To collect the changes in the interest profile, we used a JavaScript crawler to scrape the interests from the controlled accounts on a *daily basis*.

4.1.2 Browser-Extension-Based User Study. We conducted a user study to determine whether our findings from controlled experiments were valid for a larger, diverse population. We collected data from our participants' Facebook account and asked questions based on the ads they viewed and

Table 2. Positive interaction activity table

Activities*	Topic	Pages interacted †	Related Interests †	Unrelated interests †
Page like	Baking	My Baking Addiction, Dessert Recipes	My Baking Addiction, Baker, Recipes, Baking, Desserts, Flour	
	City	Islamabad, Pakistan	Islamabad, Pakistan	Dubai, Lahore, Karachi,India, Multan, Redmond Washington, United States National language, Man, Languages of Pakistan, Gross domestic product, Punjab
	Clothing Brands	Diners Clothing, Outfitters Uniworth Shop, Breakout Clothing	Diners, OUTFITTERS Outfitter	Breakout clone, Boutiques
	Electronics	Apple, Samsung	Samsung, Apple Inc., Electronics, Multinational corporation Conglomerate, Consumer electronics, Computer hardware	Insurance, NASDAQ-100, Apple (food) Dow Jones Industrial Average
	Food	Food Directory Pakistan, Pakistani Food,	Cooking, Food, Pakistani cuisine	
	Hotels	Marriott Hotel, Avari Hotel Pearl-Continental, Ramada Hotel	Marriott International, Marriott Hotels and Resorts, Pearl-Continental Hotels and Resorts	
	News Channel	Express News	Express News	
	Pizza	Domino's	Pizza, Domino's Pizza	
Like/love react	Car	BMW, Mercedes, Ferrari	BMW Z1, V4 engine	
	Cricket	PTV Sports Cricket, Sky Cricket		
	Culture	BBC Culture, World Culture Forum		UNICEF, Member states of the United , Nations, South Asian Association for Regional Cooperation, World
	Decor	Decor by Ihsan, Him and her wedding decor		
	Gym	UFC Gym, Gym feed, Yoga.com		
	Perfume	Perfume.com, Fragrance Direct	online shopping	Retail
	Personality	Mian Nawaz Shareef		imran khan official,Jeff Merkley, Rupert Grint,Prime Minister of Pakistan, President of Pakistan, Ingvar Kamprad, Jorge Sanz, Mike Newell (director),Federal government of United States,Maribel Verdú, Democratic Party (United States), Arif Alvi
	Shoe brands	Nike, Borjan	Shoes, Shopping	
	University	LUMS, FAST NUCES, NUST	Academic degree, National University of Computer and Emerging Sciences	
	Watches	Rado, Rolex, Blancpain	Rolex, Watch, Blancpain	
Comments	AI Assistants	Google Assistant, Amazon Alexa, Siri, Cortana	Amazon.com, cortana, Siri, Artificial Intelligence, Amazon Echo	
	Animals	Cats and Kittens, Dogs Lovers, The Rabbit heaven Hamster lovers	Rabbits, Dogs, Cats, Kitten, Hamster, Mammal, Dog Lovers,Cat And Kittens	Vertebrate, Hunting
	Bikes	Harley Davidson, Ducati, Kawasaki	Ducati, Ducati Multistrada, Ducati Monster, Ducati Pantah, Ducati Desmosedici, V-twin engine, KAWASAKI, Motorcycles, Ducati Desmosedici RR, Ducati Apollo, Types of motorcycles Grand Prix motorcycle racing, Kawasaki Heavy Industries, Sport bike, Ducati 851, Ducati MH900e	
	Deodorant	Axe, Degree, Dove	Deodorant,Dove Men+Care	Dove (chocolate)
	Furniture	IKEA, Liberty Furniture, Stanley Furniture	Furniture, IKEA, Ready-to-assemble furniture	
	Medicine	Panadol, Medicine, Sleeping Pills	Medicine	
	Novels	Harry Potter, Game of Thrones,	Daniel Radcliffe, Harry Potter, Game of Thrones Harry Potter (film series), Fantasy film	Haven (TV series), Mama (2013 film)
	Scientists	Albert Einstein, Issac Newton, Marie Curie	Marie Curie, Albert Einstein, Isaac Newton, Physics	Philosophy, Ethics
	Singer	AKON, Shakira, Rihanna, Inna	Rihanna, Shakira, Akon, Inna	
	Studio	Coke Studio, Nescafe Basement, Pepsi battles of bands, Music	Soft drinks, Pepsi, PepsiCo, pepsi pakistan, Coke Studio (Pakistan), Pop music	Studio Ghibli, TV reality shows, Hip Hop music, Electronic music, Entertainment Weekly, Coming-of-age story, Funny or Die, L.T.D (band), TVLine, Televisions, TV, Eurodance, Lovers (1991 film)
	Tech (Software)	Google, Facebook, Microsoft	Cloud computing, Microsoft, Facebook, Online, Google, Social network, List of Google products, Alphabet Inc.	Andriod (operating system) Online
Scroll only	Airline	PIA, Emirates, American Airlines, Airport, Airline	Emirates (airline), Dubai international Airport The Emirates Group, Airline	
	Buildings	Burj khalifa, World Trade Center	One World Trade Center, World Trade Center site, Burj Khalifa	
	Board game	Ludo, Chess		
	Construction	Vinci, Power China, Strabag	Strabag	
	Guitar	Yamaha, Fender, Gibson Guitars	Fender Musical Instruments Corporation, Guitar amplifier, Yamaha Corporation, Base guitar	
	Online transit	Airlift, SWVL	Airlift, Transportation network company	
	Ride services	Uber, Careem, Lyft	Lyft, Careem	
	Soap	Dettol, Lifebouy	Lifebuoy	
	Space exploration	NASA, Space X	NASA, SpaceX, Aerospace	
	Superhero	Superman, Spiderman	DC Comics, Spider-Man, Superman, American comic book Superhero, Kryptonian	
	Tea	Lipton, tapal, Vital Tea	Unilever, Tea, Lipton	Tapal (Travel and places)
	Tech news (web)	The Next Web, The Verge, Engadget		
	Weapons	Weapons World, Future Weapons, Weapon Lovers	Future Weapons	
	Wall paint	Nippon Paints, Berger Paints	Nippon Paint	

* Except for liking a page, all other activities were performed without liking the respective page. † Blank entries means no interests were inferred.

the interest profiles that Facebook generated. All participants were recruited through Amazon Mechanical Turk (MTurk) [1], where participants were required to be at least 18 years old and have a 95% HIT (Human Intelligence Task) approval rate spanning over at least 100 HITs. The survey questions were rendered through a Chrome browser extension that we developed and made publicly available through the Chrome Web store. Since the extension was only built for Chrome, we required participants to download the extension from the Chrome web store and log into their Facebook account. Our study went through IRB approval. Moreover, we contacted Amazon Mechanical Turk to verify that we were not violating their 'Terms of Agreement' by asking

Table 3. Negative interaction activity table.

Activities *	Topic	Pages interacted †	Related Interests †	Unrelated Interests †
Angry React	Baking	My Baking Addiction, Dessert Recipes		
	Cars	BMW, Mercedes, Ferrari	BMW, Automobiles, Ferrari, BMW Z1, V4 engine	
	City	Islamabad, Pakistan	Pakistan	Lahore, karachi, Punjab, India, Redmond Washington
	Clothing Brands	Diners Clothing, Outfitters, Uniworth Shop Breakout Clothing	Diners	
	Cricket	PTV Sports Cricket, Sky Sports Cricket		
	Culture	BBC Culture, World Culture Forum		
	Decor	Decor by Ihsan, Him and her wedding decor		
	Electronics	Apple, Samsung	Apple, Samsung	NASDAQ-100, Dow Jones, Industrial Average
	Food	Food Directory Pakistan, Pakistani Food,		
	Gym	UFC Gym, Gym feed, Yoga.com		
	Hotels	Marriott Hotel, Avari Hotel, Pearl-Continental Hotel, Ramada Hotel	Ramada Hotel	
	News Channel	Express News		
	Perfume	Perfume.com, Fragrance Direct		Retail
	Personality	Mian Nawaz Shareef		
	Pizza	Domino's		
	Shoe brand	Nike, Borjan	Shopping	
	University	LUMS, FAST NUCES, NUST	Academic degree, LUMS	
	Watches	Rado, Rolex, Blancpain	Rado	
Comments	AI Assistants	Google Assistant, Amazon Alexa, Siri, Cortana	Amazon.com, Artificial intelligence, Cortana, Siri, Amazon Echo	
	Animals	Cats and Kittens, Dogs Lovers, The Rabbit heaven, Hamster lovers	Dogs, Cats, Kitten, Hamster, Mammal Rabbits, Dog Lovers, Cats and Kittens	Vertebrate
	Bikes	Harley Davidson, Ducati, Kawasaki	Ducati, Ducati Monster, Kawasaki Heavy Industries, Ducati Desmosecdici,Ducati Pantah, V-twin engine, Ducati 851, Types of motorcycles, Ducati Desmosedici RR, Ducati Apollo, Ducati MH900e Ducati Multistrada, Sports bike Grand Prix motorcycle racing, Motorcycles	
	Deodorant	Axe, Degree, Dove	Deodorant, Dove Men+Care	Dove (chocolate)
	Furniture	IKEA, Liberty Furniture, Stanley Furniture	Furniture, IKEA,Ready-to-assemble furniture	
	Medicine	Panadol, Medicine, Sleeping Pills	Medicine	
	Novels	Harry Potter, Game of Thrones	Daniel Radcliffe, Harry Potter, Game of Thrones, Harry Potter (film series), Fantasy film	
	Scientists	Albert Einstein, Issac Newton, Marie Curie	Physics, Nobel Prize, World Science Festival, Marie Curie, Albert Einstein, Isaac Newton	Child, Human, Man
	Singer	AKON, Shakira, Rihanna, Inna	Rihanna, Shakira, Akon	George R.R Martin, DJ Whoo Kid, Rupert Grint, Gulshan Kumar, Maribel Verdu, Andy Moor (musician)
	Studio	Coke Studio, Nescafe Basement, Pepsi battles of bands, Music	coke studio, Pepsi, pepsi, Soft Drinks, Pop music	Cable television, T-Series, Televisions, Hip hop music, TV reality, Amanda Cerny, shows, Electronic music, TV, Popular music, Contemporary R&B, Latin pop, Entertainment, Weekly, Funny or Die, Coming-of-age story, Cinemax, Haven(TV series), Lovers (1991 film)
	Tech (Software)	Google, Facebook, Microsoft	Microsoft, Facebook, Cloud computing, Google, Social network, List of Google products, Alphabet inc.	Android (operating system), Online

* Except for liking a page, all other activities were performed without liking the respective page. † Blank entries means no interests were inferred.

MTurkers to install our browser extension and provide data; we were assured that we could proceed with our study with the IRB in place.

Types of Data Collected. Once the extension was installed, it first verified that participants were logged into their Facebook account. If not, participants were redirected to the Facebook sign-in page. Afterward, participants were shown a list of data that we would collect and their consent was requested before proceeding. If participants provided consent, the extension proceeded to collect *only* the following data from their Facebook account:

- The first 10 ads appearing in the timeline along with reasons for displaying such ads (i.e., we collect the "Why am I seeing this ad?" information).
- Names and links of the most recent 100 page likes and the total number of pages liked by the participant.
- Number of activities performed since account creation.
- Number of years since account creation.
- Complete ad profile of participant, i.e., data scraped from the participant's APM page.

We also imposed a number of constraints upon the extension to improve the quality of data collected. We ensured that Facebook accounts were at least three years old to limit fake accounts. Additionally, we collected the total number of activities performed on an account to understand how frequently the account was used.

Survey Generation. Once the data was collected from a Facebook account, the extension rendered a survey on a different page. The extension generated personalized questions based on the ads and interests collected from each participant's Facebook account. For instance, the survey would ask, "On a scale of 1 (not at all interested) to 5 (very interested), how much are you interested in <a specific interest>?" Along with the personalized questions, the extension also prompted participants for their demographic information and Internet usage. Furthermore, we placed an *attention check* question to ensure that participants were accurately completing the survey. Once a participant completed the survey, the responses and the data collected from their Facebook account were sent to our dedicated server. We also presented participants with a copy of their collected data. On average, it took participants 15 minutes to complete the entire process. We paid $4 per participant to complete the survey, and the extension automatically uninstalled itself once the survey ended.

Table 4. Demographics of 146 MTurk participants.

Attribute	Value (count)
Age	18-24 (13), 25-34 (70), 35-44 (40), 45-54 (17), 55-64 (5), 65 or older (1)
Gender	Male (95), Female (49), Prefer not to answer (2)
Country	United States (52), India (49), United Kingdom (25) Germany (8), Italy (4), Spain (3), France (2), Finland (2), Ireland (1)

User Data Summary. We recruited 146 participants from Amazon Mechanical Turk [1] during February 2020. There were 52 participants from the United States, 49 from India, and 45 from different European countries (mainly UK and Germany). Table 4 summarizes the demographics.

To ensure that our participants were experienced Facebook users, we added a condition that the user could only continue the survey if their Facebook account was at least three years old. We verified this by scraping data from their "Activity Log" page. On average, our participants had used their Facebook accounts for approximately 12 years. Furthermore, we also checked whether the accounts were active by collecting the number of activities performed on their account from the "Activity Log" page. The majority of the participants (~60%) logged more than 200 activities.[2] Note that number of activities is a lower-bound measure of account usage, as not all activities are logged. For instance, visiting pages or scrolling through content without interacting with the post will not create an entry in the "Activity Log" page. More details on the characteristics of our participants' Facebook account is provided in Appendix A.

4.2 Data Analysis Methods

We now explain the different methods used to analyze data from our controlled experiments as well as the user study.

4.2.1 Analysis on Controlled Experiment Data. We found that the majority of the inferred interests appeared in the ad profiles of the controlled accounts approximately 24 hours after the performed activity. We then manually mapped the inferred interests in each account to a specific performed activity. As the content (i.e. pages) that we interacted with were divided into different categories based on an overarching topic (see Tables 2 and 3), we were able to match the interests inferred to a specific category of content. Furthermore, all of the activities performed each day were completely

[2]We limited our scrapping up to 1050 activities to reduce the wait time for participants.

unrelated, thus reducing the possibility of incorrectly mapping interests to activities. Additionally, we recorded the interest profile daily, allowing us to verify that a certain interest was inferred after interacting with a specific page. Two independent reviewers manually mapped activities to interests for each iteration and came to a consensus to resolve conflicts. The calculated Cohen's kappa (κ) ranged from 0.7 to 0.91 across the two iterations of controlled experiments. The independent reviewers then set out stricter guidelines on the accuracy of an interest inference and reached a consensus on all unresolved interests during the next round of coding.

Moreover, we compared all ad profiles with the baseline account to demonstrate the causal effect of activities on inferred interests. While the methodology is relatively straightforward, performing casual inference requires very careful and precise experiments. We took great measures to isolate all account activities to prevent polluting any interactions (e.g., using separate VMs for each account and only opening a Facebook page on the browser).

To calculated the percentage of correctly (i.e., relevant) inferred interests for various accounts, we accumulated the interests that were not relevant to any of the visited content. Lastly, we analyze if there is any statistical difference in the number of interests inferred across positive and negative interaction accounts. We use Pearson's Chi-Square test or Fisher's exact test when conditions for minimum expectancy counts are not met to compute statistical difference between the two accounts [39]. We consider $\alpha = 0.05$ as an indicator of statistical significance.

4.2.2 Analysis on User Study Data. We performed Pearson's Chi-Square test to compute statistical significance between different groups. For comparing five-point Likert scale responses (for interests rating question), we used Mann-Whitney U test [25]. We consider $\alpha = 0.05$ as an indicator of statistical significance. The null hypothesis H_0 represents no statistical difference or relationship between the tested factors, whereas the alternate hypothesis H_a indicates a statistical difference between the factors. If the $p - value$ is less than 0.05, we reject the H_0; otherwise, we do not reject the H_0. We mention $p - value$, χ^2 statistic, and df (degree of freedom) for each respective analysis. We use Fisher's exact test when conditions for minimum expectancy counts of variables are not met for the Chi-Square test [39]. Bonferroni's correction was used for all post-hoc analysis to adjust for the risk of a Type I error [2].

5 INFERENCE OF INTERESTS

A recent study by Bashir et al. [20] reported that around 38% of their participants listed more than 500 interests in their Facebook accounts. We found similar proportions in our user study, where 26.7% of our participants' profiles contained more than 500 interests.[3] Such high numbers raise the question of how Facebook infers such interests. In this section, we seek to answer **RQ1:** ***How does Facebook infer user interests?*** We performed controlled experiments to understand how Facebook infers user interests. Providing insights into the interest inference process helps determine why ads may or may not be relevant to a user's actual interests. Note that our control experiments are not an attempt to reverse engineer Facebook's inference process but rather study the causal effects of activities on the interests inferred. Furthermore, we solely focus on different interactions as the current 'generic' explanation only discusses user interaction with Facebook content.

5.1 Causal Inference of Interests from Different Types of Activities

We now study what types of activities cause the ad profile to grow. We observed that different types of activities, such as commenting on posts, are more sensitive to inferring interests than

[3]On average, each participant's account listed 327 interests, except for two outlier participants, whose profile included 1684 and 2097 interests, respectively. The two outliers were excluded from the mean calculation.

other activities, such as liking posts. Following are details of observations for all four accounts in the first iteration and the different activities performed per account:

Baseline. The ad profile of the baseline account remained empty except for two interests: "City, Country" (location) [4] and "T-series" (a music company). These interests were inferred within the first few days of account creation, but there were no changes in the interest profile afterward.

Positive interaction. Interests were inferred based on the pages where all types of activities (i.e., page like, post like/love react, scroll post, and post comment) were performed. This account's interest profile was constantly updated throughout our controlled experiment — many of the pages we interacted with resulted in a direct match with an entry in the ad profile.

- We found that liking a page always resulted in an interest inference. We found at least one inferred-interest per page like performed on the content in Table 2.
- Liking or loving a post on a page, without liking the page itself, resulted in a relatively low number of interest inferences. 60.86% (14/23) of the pages where we liked or loved a post resulted in an interest inference.
- Commenting positively on a post (without liking the page or post) consistently resulted in an interest inference. 94.11% (32/34) of the positive comments resulted in an interest inference, suggesting that Facebook's inference algorithm considers posting comments as an essential indicator of interest — even more, important than liking pages or posts.
- Scroll-only activities were also found to be a significant source of interest inferences. We only viewed and scrolled through posts available on pages; we did not perform any interaction with these pages or posts. 72.97% (27/37) of the page visits in this category resulted in at least one interest. This is an aggressive form of interest inference strategy that can lead to potentially false interests, resulting in ads on uninterested topics. For instance, we only viewed the posts by "Uber," "Careem," and "Lyft" (Ride-hailing service was the topic) and received interests for "Careem" and "Lyft" without even liking their page or posts.

Negative interaction. Similar to the *positive interaction* account, we found a significant amount of interests derived from negative interactions. We performed angry reactions and negative comments on this account.

- We found that fewer interests were inferred from angry reactions compared to post likes and love reacts in the positive interaction account. 43.47% (10/23) pages in this activity resulted in one or more interests. This demonstrates that angry reactions could lead to interests that a user does not positively associate themselves with.
- We found that negative comments also result in aggressive interest inferences. We found at least one interest inference for each page that we interacted with. This demonstrates that Facebook's interest inference algorithm does not differentiate between the *sentiments* of interactions, thus inferring interests that the user may dislike or even hate. For instance, we commented negatively on a "Harry Potter" page and received interests in "Harry Potter" and "Daniel Radcliffe" (the lead actor in the Harry Potter movies).

Friend Account. We found that being friends with a person who actively and positively interacts with different Facebook content does not result in any additional interests being inferred. Moreover, we also observed that sending or receiving personal messages (through Facebook Messenger) does not cause any additional interest to be inferred. This possibly indicates that Facebook may not mine Messenger content to infer interests.

[4]We removed the city and country information for the purpose of double-blinded review.

5.2 Comparing Positive and Negative Accounts

We found that all types of activities performed in the positive and negative accounts result in interest inferences. Whether the user scrolls over a post or actively likes/dislikes a post, an interest is likely to appear in their interest profile. Facebook aggressively, and at times, falsely, identifies any type of interaction as a real interest.

In order to compare the positive and negative accounts, we compared the positive account's "Post Like/Love react" activity against the negative account's "Angry react" activity. Similarly, we compared "Positive comments" against "Negative comments." We did not use the "Post like" or "Scroll only" activities during our comparisons as these activities did not have a negative counterpart. Despite the fact that the "Angry react" activity is performed on pages also used for "Page like" activities on the positive account, we only consider the activity on similar pages/topics when comparing with the "Post Like/Love react" activity.

The total number of interests inferred on the positive and negative controlled accounts for counterpart actions were 110 and 86, respectively. The negative account contained 21.81% less inferred interests compared to the positive interaction account (considering only the counterpart activities). However, when testing the number of interests inferred across the "Positive Like/Love react" and the "Angry react" activities, we found no statistically significant difference ($\chi^2(df = 1) = 0.43, p = 0.5$). Furthermore, we also tested for differences between the "Positive comments" and "Negative comments" activities. Again, we found no statistically significant difference between these activities ($p = 0.49$, Fisher's exact test). Thus, although the total number of interests inferred across the two account type was different, there were no statistically significant differences across the various comparable positive and negative activities. We, therefore believe Facebook does not account for the type of sentiment associated with an activity when inferring interests. We were able to find negative activities resulting in interest inferences which can potentially lead to unwanted ads. However, failure to reject a null hypothesis does not demonstrate that the null hypothesis is true — only that the test did not prove it to be false. It is possible that a much larger scale experiment might find statistical significance across the various activities.

Finding 1: *A potential reason for overly large interest profiles stems from the fact that Facebook does not differentiate between sentiments (negative or positive) of user's performed activities.* Rather, Facebook considers almost every type of interaction with a post/page as an interest. Taking sentiment into account can possibly improve the quality of interest profiles and user experience.

5.3 Comparing Demographic Differences

To analyze the impact of different types of activities, we intentionally kept the demographics of different accounts the same (Location: Pakistan, Gender: Male, Age: 22). This ensured that the only difference in all the accounts is based on differences in the nature of activities and not demographics. However, performing control experiments on accounts with the same demographic information can potentially lead to biased results that cannot be generalized over a larger Facebook audience. To test if the results of our control experiments stay valid across multiple demographics, we repeated the same set of activities on newly created controlled accounts having different demographics (iteration 2). We created eight new control accounts, with each account having a specific set of demographics (i.e., location, age, and gender). The details of the demographics of the newly created controlled accounts are presented in Table 5.

We completed positive activities on the same pages as before and then compared the interest profiles of these accounts against each other. We performed statistical significance testing using Chi-Square tests (or Fisher's exact test where conditions for Chi-square tests are not met) to determine differences. To test for location-based differences, we compared accounts of the same

Table 5. Demographics of new control accounts

Account	Age	Gender	Location
Account 1	22	Male	Pakistan
Account 2	22	Male	US
Account 3	45	Male	Pakistan
Account 4	45	Male	US
Account 5	22	Female	Pakistan
Account 6	22	Female	US
Account 7	45	Female	Pakistan
Account 8	45	Female	US

age and gender from the US and Pakistan, i.e., we compared account 1 with account 2, account 3 with account 4, etc. We compared accounts of different ages but the same gender and location to determine the impact of age, i.e., we compared account 1 with account 3, account 2 with account 4, etc. Similarly, for gender, we compared account 1 with 5, account 2 with account 6, etc.

We found that Facebook's inference algorithm performed very consistently across all accounts of various demographics. None of the accounts had interest profiles significantly different from another account for all the demographics we tested. The test statistics are provided in the Tables 6, 7, and 8. The completed interests tables 13, 14, 15, 16 are present in appendix for account 1 to 8.

Table 6. Statistical tests statistics for *location* comparison. $df = 1$ for all tests

Activity type	Acc. 1 vs. 2	Acc. 3 vs. 4	Acc. 5 vs. 6	Acc. 7 vs. 8
Page Like	$\chi^2(n = 57) = 0.2586$ $p = 0.611062$	$\chi^2(n = 57) = 0.2586$ $p = 0.611062$	$\chi^2(n = 57) = 0.2586$ $p = 0.611062$	$\chi^2(n = 56) = 0.0288$ $p = 0.865272$
Post Like/Love react	Fisher exact $p(n = 35) = 0.5361$	Fisher exact $p(n = 35) = 0.5361$	Fisher exact $p(n = 35) = 1$	Fisher exact $p(n = 35) = 1$
Comment	$\chi^2(n = 146) = 0.121$ $p = 0.7279$	$\chi^2(n = 146) = 0.121$ $p = 0.7279$	$\chi^2(n = 145) = 0.0016$ $p = 0.9685$	$\chi^2(n = 145) = 0.0016$ $p = 0.9685$
Scroll Only	Fisher exact $p(n = 4) = 1$	Fisher exact $p(n = 5) = 1$	Fisher exact $p(n = 3) = 1$	Fisher exact $p(n = 15) = 1$

Table 7. Statistical tests statistics for *age* comparison. $df = 1$ for all tests

Activity type	M20 - M45 (Pakistan)	F20 - F45 (Pakistan)	M20 - M45 (US)	F20 - F45 (US)
Page Like	$\chi^2(n = 61) = 0.0424$ $p = 0.83678$	$\chi^2(n = 60) = 0.1835$ $p = 0.668351$	$\chi^2(n = 52) = 0$ $p = 1$	$\chi^2(n = 53) = 0.0055$ $p = 0.940751$
Post Like	$\chi^2(n = 66) = 0$ $p = 1$	$\chi^2(n = 52) = 0.0095$ $p = 0.922487$	Fisher exact $p(n = 3) = 1$	Fisher exact $p(n = 2) = 1$
Comment	$\chi^2(n = 151) = 0.2311$ $p = 0.630692$	$\chi^2(n = 152) = 0$ $p = 1$	$\chi^2(n = 139) = 0$ $p = 0.763518$	$\chi^2(n = 138) = 0$ $p = 1$
Scroll Only	Fisher exact $p(n = 9) = 1$	Fisher exact $p(n = 18) = 1$	Fisher exact $p(n = 0) = 1$	Fisher exact $p(n = 0) = 1$

5.4 Mixing Negative and Positive Interactions

Thus far, the positive and negative activities were conducted on separate dedicated accounts and interests were recorded to see how the inference algorithm functions for both types of actions. However, this isolated behavior testing does not capture real-world user behavior, i.e., a real user may perform both positive and negative interactions on a single account and do so in a mixed fashion. Thus, the inference algorithm might act differently when all kinds of actions are done collectively. For example, we suspected that the interest inference algorithm for the purely negative

Table 8. Statistical tests statistics for *gender* comparison. $df = 1$ for all tests

Activity type	M20 - F20 (Pakistan)	M45 - F45 (Pakistan)	M20 - F20 (US)	M45 - F45 (US)
Page Like	$\chi^2(n = 62) = 0$ $p = 0.83678$	$\chi^2(n = 59) = 0.0497$ $p = 0.668351$	$\chi^2(n = 52) = 0$ $p = 1$	$\chi^2(n = 53) = 0.0055$ $p = 0.940751$
Post Like	$\chi^2(n = 56) = 0.2876$ $p = 1$	$\chi^2(n = 62) = 0.4595$ $p = 0.922487$	Fisher exact $p(n = 3) = 1$	Fisher exact $p(n = 2) = 1$
Comment	$\chi^2(n = 152) = 0$ $p = 1$	$\chi^2(n = 151) = 0.2311$ $p = 0.630692$	$\chi^2(n = 139) = 0.0905$ $p = 0.763518$	$\chi^2(n = 138) = 0$ $p = 1$
Scroll Only	Fisher exact $p(n = 7) = 1$	Fisher exact $p(n = 20) = 1$	Fisher exact $p(n = 0) = 1$	Fisher exact $p(n = 0) = 1$

account might utilize all negative actions, as there was no positive activity. When a real user performs both negative and positive actions together, positive actions might suppress negative ones. To test whether intermixing positive and negative activities affects the inference algorithm, we performed positive and negative activities together on two other control accounts, one created in Pakistan (account 9) and the other in the US (account 10). We call them mixed-activity accounts. We compared the interests obtained from the mixed account in Pakistan with the other single-activity type accounts in Pakistan (ones that perform only one type of activity, either positive or negative). We also compare the interests captured across the mixed accounts from Pakistan and the US.

The mixed-activity accounts consisted of both positive and negative interactions to best mimic real user behavior. We performed page like, positive comment, angry react, and negative comment activities on these accounts. All positive and negative activities were performed on different pages to differentiate what activity and sentiment caused an interest inference. For instance, we would perform 'like' and 'angry' reactions on two different pages consecutively, or positively and negatively comment on different pages one after the other. While a real user can perform positive and negative activities on a single page, this is not feasible in our control experiment as we would not be able to determine if the interests were inferred due to negative or positive actions. To test whether the behavior of the inference algorithm is different under a mixed setting, we compared the interest profiles of the mixed-activity accounts with single-activity accounts. We compared the interests generated from specific positive and negative pages on the mixed account with their corresponding pages from the purely positive and purely negative accounts, respectively. For instance, interests from 'post-like' and 'positive comments' activities in the mixed account are compared to interests derived from 'post-like' and 'positive comments' in the isolated positive account. Similarly, interests derived from 'angry react' and 'negative comments' activities in the mixed accounts are compared with 'angry react' and 'negative comment' in the negative isolated account. This allowed us to determine if the interest inference algorithm behaved differently with activities performed in a mixed setting versus activities of a single sentiment. We performed Chi-Square tests to measure the difference between the interest profiles. We found that the behavior of the inference algorithm under isolated and mixed settings is similar; there is no statistically significant difference ($p < .05$) among the interest profiles for all types of activities. Therefore, none of the given actions suppresses interest inferences of other actions. This suggests that the results drawn in the isolated setting are valid for the mixed activities as well. The details of the statistical test comparing isolated and mixed accounts are presented in Table 9. The completed interests table 17 is present in appendix for account 9 and 10.

5.5 Unliking Pages and Hiding Posts

Previous work by Andreou et al. [17] found that user interests may be removed over time, yet did not attempt to uncover why. Thus, we investigated which actions could cause an interest to be

Table 9. Statistical tests statistics for mixed vs. isolated account comparisons. $df = 1$ for all tests

Activity type	Mixed (Pakistan) - Isolated	Mixed (Pakistan) - Mixed (US)
Page Like	$\chi^2(n = 71) = 1.6466$ $p = 0.199424$	$\chi^2(n = 52) = 0.2311$ $p = 0.630701$
Positive Comments	$\chi^2(n = 119) = 1.9961$ $p = 0.157704$	$\chi^2(n = 61) = 0.4416$ $p = 0.506371$
Angry React	$\chi^2(n = 34) = 3.2343$ $p = 0.072113$	Fisher Exact $p(n = 11) = 1$
Negative Comments	$\chi^2(n = 124) = 0.7877$ $p = 0.37480$	$\chi^2(n = 87) = 0.0098$ $p(n = 0) = 0.921208$

removed from a user profile. While there are likely many reasons for the inference algorithm to remove an interest, we hypothesize that two main actions can lead to an interest removal: unliking a page and hiding a post. These actions were performed on the mixed-activity accounts once pages were already liked and the news feed was populated with posts or ads based on the inferred interests.

5.5.1 Unliking Pages. This action consisted of unliking a group of pages 10 days after the same group of pages was originally liked. This time frame allowed ample time for interest inferences to generate interests for all pages. During the page-like phase for the mixed accounts, pages were grouped together by topic, and all of the pages within that topic were then liked consecutively. Page unlike activities were performed similarly; the group of pages that were originally liked was all unliked consecutively. For instance, the topic "Technology" contained the pages for Apple and Samsung. Both of these pages were liked in the first day of activities, and interests were quickly generated. Ten days later, we unliked all pages relating to the topic "Technology," or the pages for Apple and Samsung.

We found that unliking a page can remove the interests that were inferred by liking that particular page. This is the first instance where we observed interests being removed from ad profiles. We saw that only a significant portion, not all of the interests inferred by liking pages, were removed by unliking them. 80% and 69.56% of interests were removed by page-unlike activity for Pakistan and US-based accounts, respectively. This demonstrates that the effect on interest profiles is not entirely reversible through unliking pages.

5.5.2 Hiding Posts. We also wanted to determine if hiding a post on the news feed (derived from an action on a page) causes the removal of interests that were related to those pages. Before beginning the post-hide activity, we paused for two days to allow new interests to be generated in the ad profile. We wanted to limit the number of posts hidden in a day, as hiding posts too frequently is not a natural user behavior and might alert Facebook of unusual activity. Thus, we hid the first two posts that appeared on the news feed belonging to the pages we had previously liked, and we limited ourselves to hiding at most four posts (two posts per relevant page we had liked). For example, if we see a post from pages in the following order (X_1 means the first post relevant to page X we had previously liked): $Hulu_1$, $Hulu_2$, CNN_1, $Amazon_1$, $Hulu_3$, CNN_2, $Netflix_1$, then we would hide $Hulu_1$, $Hulu_2$, CNN_1, and CNN_2 on a single day. This activity was performed for seven days, and posts were hidden regardless of whether they were previously hidden. This caused posts from several pages to be hidden multiple times throughout many days. This activity was performed to mimic real user behavior where users may hide irrelevant posts on their timelines. However, we observed no interest removal from the post-hide activity.

6 ACCURACY OF INTEREST INFERENCE

In this section, we seek to answer **RQ2: *How accurately does Facebook infer interests from user activities?*** Facebook builds interest profiles for users, which include the keywords of products, topics, or places that a person is interested in. Ads are targeted based on the interests available in a user's profile. Therefore, the relevancy of displayed ads ultimately depends upon the accuracy of inferred interests. For this reason, we analyze the accuracy of user interest profiles.

6.1 Controlled Experiment

We computed the accuracy of the inferred interests for both positive and negative interaction accounts. This was accomplished by calculating the percentage of correctly inferred interests out of the total interests inferred in each controlled account. Since we mapped the interests to the relevant pages, we also noted the incorrectly inferred interests for each topic. Note, two independent reviewers manually mapped activities to interests and came to a consensus to resolve conflicts (Cohen's kappa ranged from 0.7 to 0.91). We consider interests irrelevant if they are not semantically related to the pages on which activities were performed on a given day.

We found a total of 187 inferred interests for the positive interaction account. Around 68.98% (129/187) of such interests were relevant to our planned activities. The negative interaction account also had a comparable accuracy rate of 62.60% (i.e., 62 out of 99 total interests were accurately inferred). Note that when we are stating accuracy, we imply that the inferred interests match the topic of the content visited, irrespective of whether the user is interested or not (i.e., ignoring the sentiment of the performed activity). The relevance of both positive and negative accounts combined is 66.78% (191/286).

In order to find how interest relevance varies across *activities*, we computed the relevancy for all activities across both positive and negative accounts. We observed 61.36% of interests inferred from the "Page like" activity to be relevant. For the "Post Like/Love react" activity, only 37.03% interests were relevant. This demonstrates that Facebook more accurately infers interests based on page likes than post likes. Additionally, commenting positively on posts led to 73.49% relevant interests, depicting that comments are very important interest indicators. Lastly, and surprisingly, the "Scroll only" activity resulted in 96.87% of the inferred interests being accurate. This activity was performed without any other action on the page. We also observed that in some cases, Facebook inferred the wrong interest based on word similarity rather than the underline context. For example, upon visiting the Apple (Tech company) page, Apple (fruit) was inferred as an interest. For the 'Angry reacts' activity, we observed 65.21% relevant interests. For the 'Negative comments' activity, 61.84% of interests were relevant. When comparing the relevancy of interests inferred based on positive and negative comments, we found no significant statistical difference ($\chi^2(df = 1) = 2.47, p = 0.11$).
Finding 2: *Approximately 33.22% (95/286) of interests inferred across both positive and negative accounts were found to be irrelevant.* This suggests that Facebook may not accurately infer interests regardless of an action's sentiment.

6.2 User Study

In this section, we analyze whether our findings hold for a larger population by conducting a user study, where we collect data from participants' own Facebook accounts and then ask questions about their interests inferred by Facebook. Then we analyze the accuracy of interests inferred about them based on their belief about their preferences.

6.2.1 Relevancy of Interests in Ad Profiles. In this section, we measure the accuracy of the ad profiles for our user study participants. To accomplish this, we asked each participant a set of questions about the interests listed in their ad profile.

Five interests were randomly selected from the participant's ad profile; the participant was asked for their interest level on a scale of 1 to 5 (1 being "not at all interested" and 5 being "very interested").We selected random, yet diverse, interests using the *Datamuse API* [3] to ensure that a selected interest was not related to any of the other selected interests. The Datamuse API takes a word and returns related, relevant words on a pre-trained language model. We retrieved 100 related words for each interest and verified that the returned words did not match any previously selected interest. If no match was found, then the interest under consideration was used; otherwise, we randomly selected another interest and repeated the whole process. If a participant's ad profile did not contain five diverse interests, their profile was excluded from this analysis. 26 participants were eliminated from this analysis as their interest profile contained less than five unrelated interests. The remaining 120 participants' profiles contained five distinct and unrelated interests that were used for this analysis.

Figure 2 depicts the results of the ratings for the various interests. We found that 29.3% of all interests (out of 600 total interests) were not interesting to users, obtaining a rating of either "not at all interested" or "somewhat uninterested." Among all interests rated, 17% of the interests were rated as "not at all interested", indicating that a substantial portion of interests inferred by Facebook is not relevant to users' actual interests. We found that 65.8% (79 out of 120) of the participants consider at least one interest listed in their ad profiles as irrelevant.

We also looked at whether this trend varied across different geographic regions. We found that 37.7%, 14.1%, and 36.6% of the interests were flagged as not interesting by participants from the US, India, and Europe, respectively.[5] Figure 2 shows the location-wise breakdown for interest ratings. We performed pairwise Mann-Whitney U tests to find any statistically significant differences in the ratings across different locations. Table 10 highlights the statistical results. We found a statistically significant difference in the interest ratings obtained from Indian participants to that of US participants ($p < 0.001$) and European participants ($p < 0.001$).

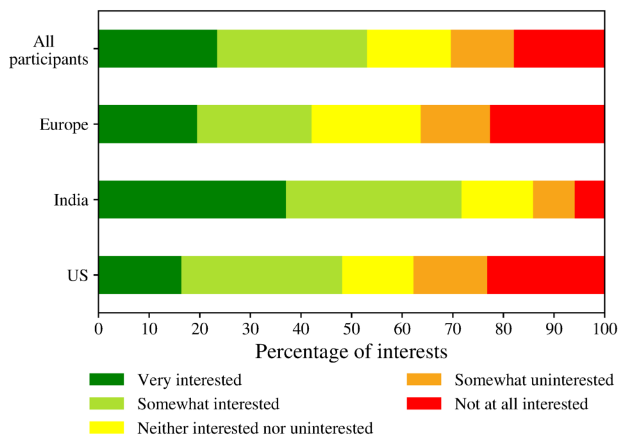


Fig. 2. Interests rating on a scale of 1 to 5 for five random interests found on each participant's Facebook profile. $p - values$ for different location pairs, US-EU=0.136, US-India: < 0.001, EU-India: < 0.001.

Finding 3: *Considerable proportions of the participants indicated that many of the interests listed in their ad profiles were inaccurate (i.e., irrelevant) based on what participants considered as relevant to*

[5]"Not interested" encompasses both "not at all interested" or "somewhat uninterested."

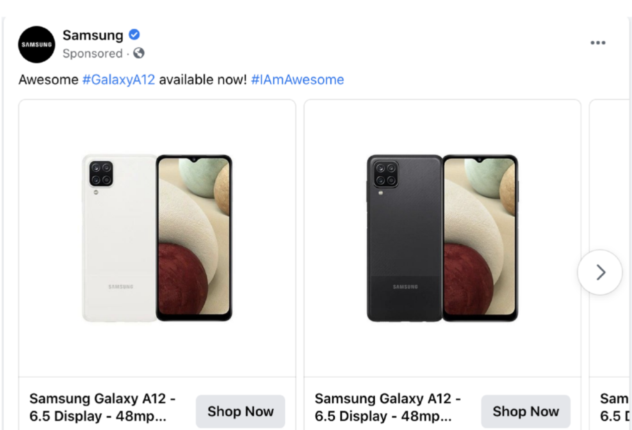
Table 10. Pairwise Mann-Whitney U test on relevancy of interests among participants from different geographic regions. Median value for US and Europe is 3 (Neither interested nor uninterested), for India median is 4 (Somewhat interested).

Country Pair	Mann-Whitney U stat	Corrected p-value
US - Europe	20758	1
US - India	12326	$< 0.001^{***}$
Europe - India	10590	$< 0.001^{***}$

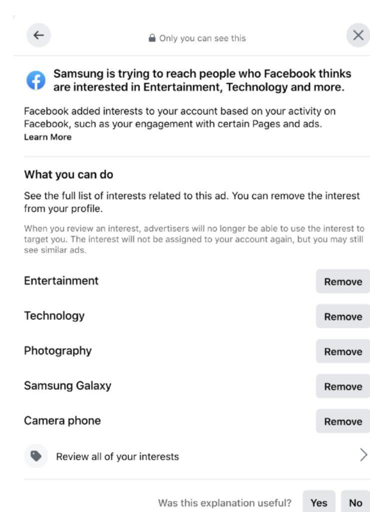
$^{*} = p < 0.05$, $^{**} = p < 0.01$, $^{***} = p < 0.001$

their perceived interests.. This is particularly true for participants from the US and Europe, where around 37.7% and 36.3% of the interests are marked irrelevant by the participants, respectively.

6.2.2 Relevancy of Interests Listed in Sponsored Ads. To further analyze the relevancy of interests listed in ad profiles, participants were asked questions about sponsored ads appearing on their home page. Facebook enables users to view the reasons for seeing any sponsored ad — this often includes interests used to find a matching ad (see Figure 3). We asked participants questions about sponsored ads they viewed after ensuring that the keywords in the advertisement matched one or more interests in their ad profile. As a result, participants saw varying amounts of questions in this section. On average, 3 out of the 10 ads collected were based on interests (standard deviation was 1.7). We asked participants to rate whether each interest mentioned in an ad explanation felt relevant to their true interest or not. In other words, we asked whether users explicitly believe the interests are relevant, irrespective of whether in reality they would interact with ads related to such interests.



(a) Targeted ad by Samsung



(b) Ad targeted based on interests

Fig. 3. Example of a targeted ad and reasons for seeing such an ad. We can see ad interests listed in the explanation.

Figure 4 depicts the distribution of interest-based ads based on one or more interests. The majority of the ads were targeted based on one interest. Around 43.3% of such ads were perceived to contain an irrelevant interest. Among all participants who received an interest-based ad (37

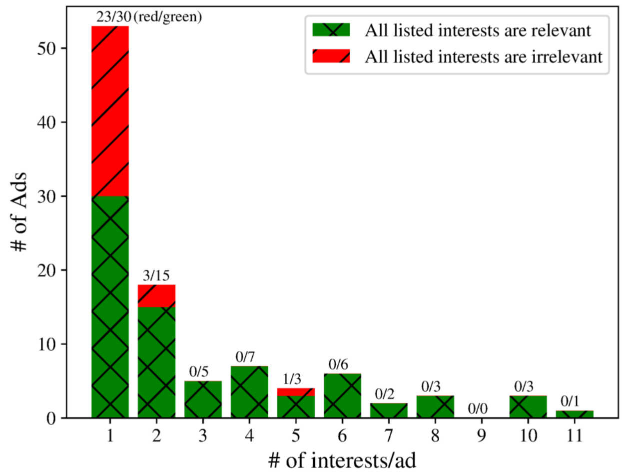


Fig. 4. The distribution of relevant and irrelevant interests across different interest-based ads.

such participants)[6], 45.2% of them reported seeing at least one ad containing at least one irrelevant interest. 66.7% of all interest-based ads (102 such ads) were also flagged to contain at least one irrelevant interest. Moreover, 26.47% of ads contain were flagged to contain all irrelevant interests. Figure 8e (in appendix) shows the CDF of the percentage of interests marked as irrelevant per ad. We can see that for 50% of the ads over 60% interests are irrelevant.

Finding 4: *A large portion of the targeted ads were based on interests that participants perceived as irrelevant.* As highlighted through our user study, 66.7% of all interest-based ads contained at least one irrelevant interest and 26.47% of ads had all irrelevant interests.

7 EXPLANATIONS PROVIDED BY FACEBOOK

To better promote transparency, Facebook provides explanations about why an interest appears on one's ad profile. In this section, we seek to answer **RQ3:** ***Does Facebook accurately explain how inferred interests are derived?*** We discuss the various interventions that Facebook takes to enhance transparency about the interests inferred and their effectiveness.

7.1 Awareness of Ad Explanation and APM

In order to promote transparency, Facebook provides explanations regarding why a user sees a sponsored ad. These explanations can be viewed from any ad through the "Why am I seeing this ad?" option. Facebook also provides the Ad Preference Manager (APM), enabling users to view all interests inferred about them. The APM also allows a user to make changes or delete information from their ad profile. This section analyzes users' knowledge about these features and how often the features are utilized.

The APM and ad explanations are helpful when users actually know how to access this information. However, privacy settings are often difficult to find and users rarely interact with them; this is evident from the participants' responses, where 65.8% of participants have never visited interest profile and 80.8% have never changed their interest profile (see Table 11). In order to understand the awareness regarding Facebook's ad transparency features, we asked participants about their knowledge in reviewing interest profiles and ad explanations. These questions consisted of asking

[6]73 users (out of 146) received ads, of which 37 users received interest-based ads.

Table 11. Survey questions and responses regarding ads explanation and ad preference manager. $^{*}= p < 0.05$, $^{**}= p < 0.01$, $^{***}= p < 0.001$, $df = 1$ for all questions.

Question	Yes	No	Pair-wise	$\chi^2(1)$	Corrected p-value
Do you know that FB created a list of products/topics/profiles that you might be interested in?	66.4%	33.6%	US - Europe	0.788	1
			US - India	4.428	0.105
			Europe - India	0.898	1
Do you know this (your interests) list is available to you on Facebook?	47.9%	52.1%	US - Europe	0.002	1
			US - India	5.32	0.063
			Europe - India	6.062	0.041*
Do you know that you can view the reason for every ad you see on your FB page?	47.3%	52.7%	US - Europe	0.609	1
			US - India	7.182	0.021*
			Europe - India	2.632	0.312
Have you ever visited Facebook Ads settings/preferences page?	34.2%	65.8%	US - Europe	0.920	1
			US - India	0.039	1
			Europe - India	0.912	1
If yes, have you ever changed the default ad settings?	19.2%	80.8%	US - Europe	0.006	1
			US - India	2.137	0.429
			Europe - India	2.435	0.354
Do you know that you can remove interests from the interests list?	24%	76%	US - Europe	0.029	1
			US - India	0.004	1
			Europe - India	0.0002	1

how many participants knew about these features or have used such features. Table 11 contains several questions asked and their corresponding responses. We see that the majority (>50% from the second and third rows of the table) of our participants were unaware of the availability of ad explanations and interest profiles. Furthermore, the majority of the participants have never visited their interest profiles, removed any interest from them, or even changed the default settings. These high numbers indicate that participants are unaware of these features due to poor accessibility. Worryingly, around one-third of our participants did not even know that Facebook builds an interest profile for them.

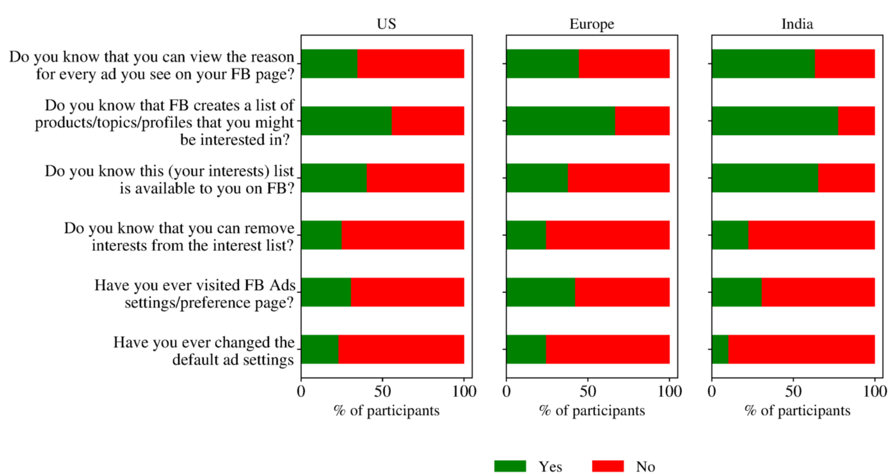


Fig. 5. Responses regarding ads explanation and ad preference manager across participants from different geographic regions.

We also analyzed whether users' knowledge of the Facebook APM and ad explanations varied between different geographic regions. Figure 5 shows the region-wise breakdown of responses. We found no significant differences in responses from participants in the US and Europe. However, we found a difference for participants from India; more than half of the Indian participants were

knowledgeable about ad explanations and ad profile availability, which is surprising because most American and European participants were not knowledgeable about these features. We performed pairwise Chi-Square tests for each question to verify the statistical difference. There was a statistical difference ($p - value < 0.05$) in the responses between participants from India and Europe as well as India and the US for questions regarding the awareness of interest profile and ad explanation.
Finding 5: *A large proportion of users are not aware of the ad explanation feature as well as the ability to view and edit their ad profiles.* This indicates that these features are not noticeable enough. Around one-third of our participants were unaware that Facebook collects interest about them, which suggests that Facebook needs to make the APM more accessible.

7.2 Explanation of Interest Inference

Facebook's APM explains why a particular interest was inferred. To evaluate the accuracy of interest inference explanations, we examined the interest explanations from our controlled accounts. As described in Section 5, the activities performed on each page were clearly planned and profiles were recorded at regular intervals, providing us with ground truth reasons for interest inferences. Upon comparing the interest explanations with the ground truth reasons, we found that the explanations were misleading and vague.

We believe that the fundamental shortcoming for interest explanations is that the explanation template is overly *generic*. For instance, let 'X' represent an interest, then the explanation related to interest 'X' would be — "You have this preference because of your activity on Facebook related to X's page, for example, liking their Page or one of their Page posts" (also shown in Figure 6). Thus, whether a user scrolls over a page or likes a post, the explanation is always the same. Such ambiguity can also lead to misunderstanding as only scrolling over a page is interpreted as liking a page or page post. Furthermore, making no distinction between positive and negative actions is somewhat misleading. For example, we negatively commented about Siri on their page, and an interest in "Siri" was inferred with an explanation that you might have liked a page or post regarding "Siri", which is not true at all. This can mislead users into believing that they acted on 'liking' the content when in reality, they never did. Therefore, Facebook explanations do not provide an accurate reason behind an inferred interest. The ambiguity and inaccuracy in interest explanations is also a privacy concern, as reading or scrolling on a sensitive page should not translate into an affirmative action, such as liking that page. While we understand that it is difficult to explain complex machine learning decisions in words, we expect Facebook to, at the very least, provide an interpretable explanation highlighting the key reasons for an interest inference, such as differentiating between liking a post, commenting on it, or just viewing it (Eslami et al. [27] also found users to prefer interpretable, non-invasive explanations). Providing such explicit reasoning for a given interest inference is feasible as such activities are distinctly recorded in Facebook's "Activity log" [7].
Finding 6: *The generic interest explanation template provided by Facebook is both misleading and ambiguous.* As the interest explanation template used by Facebook is very generic, it fails to provide any concrete reasoning as to why an interest was inferred, which as times can mislead users in believing the wrong reasons.

7.3 Notifying Inference of Sensitive Interest

Furthermore, Facebook knowingly causes additional privacy concerns by collecting data surrounding sensitive interests. Cabañas et al. [21] conducted an extensive study on the sensitivity of interests inferred by Facebook and found that Facebook collects and allows advertisers to target ads with sensitive interests, such as political leaning, religious belief, and sexuality [5]. As sensitive interests can be inferred and interest explanation lacks clarity, this quickly becomes a privacy

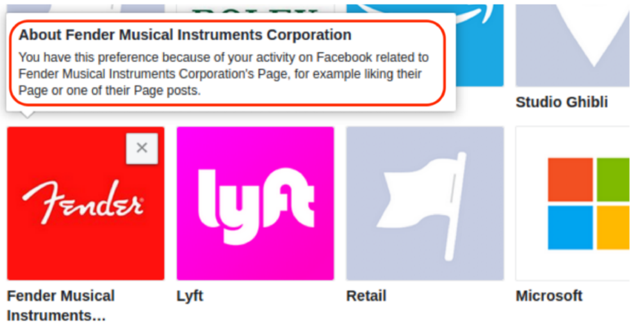


Fig. 6. Generic explanation template for any inferred interests. The ambiguity and inaccuracy in interest explanation can cause privacy concerns.

Table 12. Survey questions and responses regarding sensitive ads. *= $p < 0.05$, **= $p < 0.01$, ***= $p < 0.001$, $df = 2$ for all questions

Question	Yes	Maybe	No	Pair-wise	$\chi^2(2)$	Corrected p-value
Are you satisfied with your experience with Facebook ads?	38.3%	41.4%	20.3%	US - Europe	0.39	1
				US - India	15.15	0.0015**
				Europe - India	12.67	0.0051**
Do you see any inappropriate ads (ads the make you uncomfortable)?	37%	17.1%	45.9%	US - Europe	3.04	0.63
				US - India	3.731	0.462
				Europe - India	4.3	0.348
Have you seen any political campaign ads?	28.8%	9.6%	61.6%	US - Europe	7.76	0.0618
				US - India	5.365	0.204
				Europe - India	0.3194	1
Did you see any ads related to health care or medication?	17.1%	8.9%	74%	US - Europe	0.902	1
				US - India	3.663	0.48
				Europe - India	0.98	1

concern for users, especially since users are not properly notified. Other studies have also analyzed sensitive ads on topics surrounding political campaigns [32] and healthcare [13, 41]. For instance, a study on political ads found many "inauthentic communities" that spread political information via undeclared coordinated activity [26]. While existing studies analyzed the presence of sensitive ads, they did not focus on whether participants hid/block such sensitive ads.

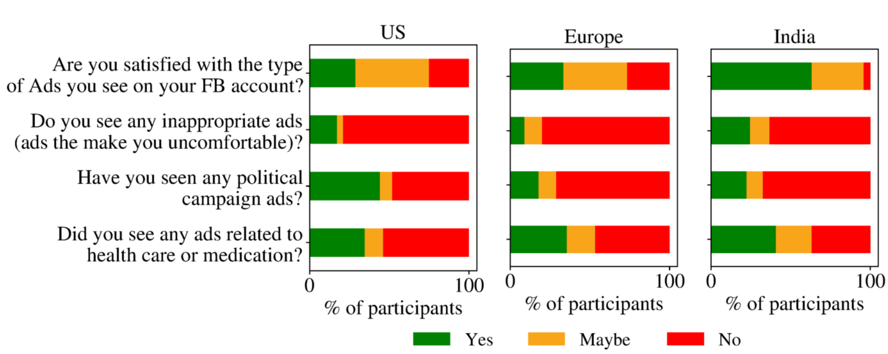


Fig. 7. Response to seeing sensitive ads across different geographic regions.

We first asked our participants whether they viewed sensitive ads, specifically related to healthcare, political campaign, or other inappropriate content. We summarized these responses among participants from different geographic regions to discover any trends. These questions and user responses are shown in Table 12, and Figure 7 shows the breakdown in terms of geographic regions. We found that the majority of participants from the US and Europe were not satisfied with their Facebook ad experience, demonstrating that targeted advertising requires improvement. However, participants from India claimed to be more satisfied. We found a statistically significant difference ($p < 0.01$) in ad satisfaction between participants from India when compared with participants from the US and Europe.

Finding 7: *A significant proportion of participants saw advertisements that are sensitive in nature, such as political campaign ads or medication/healthcare-related ads.* Political campaign ads were more common in the US than in Europe and India.

Facebook's APM also tracks any user-hidden advertisements. This section is called "Whom you've hidden," falling under "Advertisers and Businesses." We analyzed whether hidden advertisements and pages contained any sensitive interests. We found that 39 of 146 participants have hidden at least one ad on Facebook throughout their account's lifetime. Additionally, many hidden advertisements contained sensitive information, such as political ads with the keywords "Donald J. Trump" and "Hillary Clinton." Furthermore, hidden ads contained content related to prescription medicine, such as "Roman Health," "Optimal Health Knoxville," and "Hyland's Homeopathy." Our participants' interest profiles even had sensitive interests based on their personal identity. Over 40% (60 of 146) participants' profiles contained interests based on sexuality, such as "Homosexuality" and "Gay-friendly," as well as interests based on their religious beliefs, such as "Atheism" and "Irreligion." Unfortunately, Facebook does not show why an advertisement was hidden on APM. If this feature were available, it could provide deeper insight as to why users hid certain ads.

Finding 8: *Many of the advertisements hidden by participants contained sensitive interests.* Facebook serves targeted ads based on sensitive interests, which many users later hide/block. This suggests that Facebook needs to better notify users about inferring sensitive interests.

8 DISCUSSION AND RECOMMENDATIONS

We conducted controlled experiments through planned activity on new Facebook accounts to determine how Facebook infers user interests. We found that Facebook infers interests based on the pages or posts that a user interacts with through controlled experiments. This 'interaction' includes explicit activities, such as likes and comments, and passive activities, such as scrolling through posts. This points to the fairly loose criteria used by Facebook to select an interest, which may result in irrelevant ads. Furthermore, given that the semantics/context of an activity (e.g., liking or disliking a post) is not considered while inferring interests, irrelevant or even inappropriate ads may be viewed. Inaccurate interest profiles have both economic and privacy implications. From a financial perspective, advertisers should know the effectiveness of their paid ads as well as ensure that the ads are being displayed to the correct audience [35]. Facebook does not properly communicate how the interests are inferred from a privacy perspective and may produce an incorrect interest inference, which can further exacerbate if Facebook decides to share data with partners (or third parties). If the data inferred about a user is incorrect, then the third party may inadvertently misuse that data and provide inappropriate content to users [6]. Furthermore, this drastically reduces consumer transparency, as users are left unaware about what actions trigger an interest to be generated. By improving the transparency behind the interest inference process and explanations, users can become knowledgeable about what specific activities cause their interests to be generated. Thus, users would have the option to browse privately when interacting with sensitive content or other content for which they do not want any interest inference. Users would also have insights regarding what specific actions caused an interest inference, and as a result, caused a targeted advertisement. If such advertisements are negatively viewed, users would be able to remove that interest, thus hinting Facebook to not make such inferences in the future. From these findings, we provide suggestions to Facebook to improve the transparency of the algorithm and provide more control to consumers. We also offer recommendations to the community for future work in this field surrounding the development of usable user interfaces while improving user transparency.

Recommendations to Facebook. As Facebook already records every activity performed by a user, we suggest that Facebook has the opportunity to improve the interest inference algorithm to differentiate between positive and negative sentiments to add context to user actions. Sentiment analysis algorithms, such as VADER, have proven to be effective in social media contexts [30] and can be easily incorporated. This would improve the relevancy of the interests inferred.

With the new APM update, users must navigate through six links to view their interests (Settings & Privacy –> Settings –> Ads –> Ad Settings –> Categories used to reach you –> Interest Categories). The interests are then viewed in an unsorted list without any mechanisms provided to improve accessibility. Facebook should provide different filtering and grouping options to allow the user to more easily view/block their interests. Furthermore, Facebook should improve the accessibility of this page by reducing the number of links needed to access the hidden inferred interests. Future research is required in this field to determine what mechanisms should be provided to provide a more usable interface and where the general interests should be located to improve accessibility.

Future Work. Future research needs to be performed in HCI to construct interpretable explanations for machine-learning-based inference algorithms better. The current inferred-interest explanations are overly generic, and while existing studies advocate interpretable explanations [27], the right balance of information and intrusiveness has not been studied in depth.

Developing a usable interface is always challenging, and over the years, we have seen researchers focus on improving the security and privacy warning messages for browsers [14]. Similarly, we advocate the need for more user-oriented studies to determine the appropriate placement and accessibility of various ad-related indicators on the user interface. For example, over 52% of our respondents were unaware of the "Why am I seeing this ad?" functionality, highlighting that it is poorly placed on the user interface.

Facebook currently nudges users to perform a yearly privacy checkup and to learn more about their privacy settings [12]. While some users are unaware of the privacy checkup availability [38], future work needs to be conducted to determine the frequency and framing of the privacy-checkup nudge to better educate users on privacy preferences as well as the interests inferred about them.

Lastly, while capturing contextual data helps to improve inference, it also imposes privacy risks. Thus, developing techniques to capture the right amount of contextual data in a privacy-preserving manner is an important future direction. In the context of Facebook's interest inference, it would be worth thinking about ways to infer interests in a more privacy-friendly manner, e.g., not capturing interest on sensitive topics.

9 LIMITATIONS

Our study has a few limitations. Firstly, we collected data from only 146 participants — the majority of whom were male. However, we recruited a comparable number of participants (statistically significant) from three different geographic regions (US, Europe and India) to capture diverse viewpoints. Furthermore, we validated that their accounts were real accounts with regular interactions (i.e., we only considered accounts that were at least three years old and were frequently used). Next, our sample size for the controlled experiments was relatively small. The accounts created and the activities chosen could also have contained minor selection bias. We considered somewhat a diverse demographics in the controlled accounts and varied the types of pages where activities were performed. We did not find *statistically significant* difference across these factors. However, failure to reject a null hypothesis does not demonstrate that the null hypothesis is true. It is possible that a larger number of controlled accounts and higher resolution of activities may help derive more fine-grained results, but such an approach would be challenging to scale and automate. Next, we could not scrape ads from all users' news feeds as our extension was not compatible with all versions of Chrome across different platforms. As a result, we were able to collect ads from only 73 users (out of 146). Lastly, many studies have also shown that APMs may hide sensitive interests (such as religion, disease, alcohol, etc.) [17, 24, 46]. Our results, therefore, should be interpreted as a *lower bound* on the number of interests collected by APMs.

10 CONCLUSION

In this paper, we investigated how Facebook reacts to varying activities and reactions in the context of inferring user interests by performing controlled experiments. We found that Facebook does not differentiate between positive and negative interactions. As a result, many of the interests listed in the ad profile become irrelevant or inaccurate. We also confirm the inaccuracy of inferred interests through a user study, where we collect data from participants' Facebook accounts and ask questions about their interest profiles. We found that interest inference explanations are overly generalized, making them vague and often misleading, suggesting that Facebook needs to clearly state each inference's reasons. We also highlight the lack of awareness among participants regarding the availability of explanations and capability to remove interests from their profiles — suggesting Facebook needs to better promote the available transparency features.

ACKNOWLEDGEMENTS

We thank Dr. Muhammad Fareed Zaffar (Lahore University of Management Sciences) for his co-supervision in this research work and for facilitating pilot user study. We also thank our anonymous reviewers for their valuable feedback.

REFERENCES

[1] 2005. Amazon Mechanical Turk. https://www.mturk.com/
[2] 2015. Handbook of Biological Statistics. http://www.biostathandbook.com/chiind.html
[3] 2016. Datamuse API. http://www.datamuse.com/api/
[4] 2018. Facebook Help Center. https://web.archive.org/web/20180909053129/https://www.facebook.com/business/help/298717656925097
[5] 2018. Facebook lets advertisers target users based on sensitive interests. https://www.theguardian.com/technology/2018/may/16/facebook-lets-advertisers-target-users-based-on-sensitive-interests
[6] 2018. Facebook's data-sharing deals exposed. https://www.bbc.com/news/technology-46618582
[7] 2020. Facebook Activity Log. https://www.facebook.com/profile.php/allactivity
[8] 2020. Facebook Ad Preferences. https://facebook.com/ads/preferences/
[9] 2020. Facebook Advertising Targeting Options. https://web.facebook.com/business/ads/ad-targeting?_rdc=1&_rdr
[10] 2020. Facebook IQ. https://www.facebook.com/iq/insights-to-go/200m-more-than-200-million-businesses-use-our-free-tools-to-connect-with-customers-we-also-have-more-than-10-million-active-advertisers-across-our-services-the-vast-majority-of-which-are-small-and-medium-sized-businesses//
[11] 2020. Google Ads Settings. https://adssettings.google.com/
[12] 2020. Guiding You Through Your Privacy Choices. https://about.fb.com/news/2020/01/privacy-checkup/
[13] 2020. Restricted content in Facebook ads. https://m.facebook.com/policies/ads/restricted_content/pharmacies
[14] Devdatta Akhawe and Adrienne Porter Felt. 2013. Alice in warningland: A large-scale field study of browser security warning effectiveness. In *Proceedings of the 22nd USENIX Security Symposium (USENIX Security)*. 257–272.
[15] Muhammad Ali, Piotr Sapiezynski, Miranda Bogen, Aleksandra Korolova, Alan Mislove, and Aaron Rieke. 2019. Discrimination through Optimization: How Facebook's Ad Delivery Can Lead to Biased Outcomes. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 199 (Nov. 2019), 30 pages. https://doi.org/10.1145/3359301
[16] Athanasios Andreou, Marcio Silva, Fabrício Benevenuto, Oana Goga, Patrick Loiseau, and Alan Mislove. 2019. Measuring the Facebook Advertising Ecosystem. In *Proceedings of the 26th Annual Network and Distributed System Security Symposium (NDSS)*.

[17] Athanasios Andreou, Giridhari Venkatadri, Oana Goga, Krishna P Gummadi, Patrick Loiseau, and Alan Mislove. 2018. Investigating Ad Transparency Mechanisms in Social Media: A Case Study of Facebook's Explanations. In *Proceedings of the 25th Annual Network and Distributed System Security Symposium (NDSS)*. 1–15.

[18] Rebecca Balebako, Pedro Leon, Richard Shay, Blase Ur, Yang Wang, and L Cranor. 2012. Measuring the effectiveness of privacy tools for limiting behavioral advertising. In *IEEE International Workshop on Web 2.0 Workshop on Security and Privacy*.

[19] Paul Barford, Igor Canadi, Darja Krushevskaja, Qiang Ma, and Shan Muthukrishnan. 2014. Adscape: Harvesting and analyzing online display ads. In *Proceedings of the 23rd international conference on World wide web*. 597–608.

[20] Muhammad Ahmad Bashir, Umar Farooq, Maryam Shahid, Muhammad Fareed Zaffar, and Christo Wilson. 2019. Quantity vs. Quality: Evaluating User Interest Profiles Using Ad Preference Managers. In *Proceedings of the 26th Annual Network and Distributed System Security Symposium (NDSS)*.

[21] José González Cabañas, Ángel Cuevas, and Rubén Cuevas. 2018. Unveiling and Quantifying Facebook Exploitation of Sensitive Personal Data for Advertising Purposes. In *Proceedings of the 27th USENIX Security Symposium (USENIX Security)*. 479–495.

[22] J. Clement. 2020. Facebook users worldwide 2020. https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/

[23] J. Clement. 2020. Instagram users worldwide 2023. https://www.statista.com/statistics/183585/instagram-number-of-global-users/

[24] Amit Datta, Michael Carl Tschantz, and Anupam Datta. 2015. Automated Experiments on Ad Privacy Settings. *Proceedings on Privacy Enhancing Technologies* 2015, 1 (2015), 92 – 112.

[25] J. F.C. de Winter and D. Dodou. 2010. Five-Point Likert Items: t test versus Mann-Whitney-Wilcoxon (Addendum added October 2012). *Practical Assessment, Research, and Evaluation* 10, 15 (2010). https://doi.org/10.7275/bj1p-ts64

[26] Laura Edelson, Tobias Lauinger, and Damon McCoy. 2020. A Security Analysis of the Facebook Ad Library. In *Proceedings of the 41st IEEE Symposium on Security and Privacy (SP)*. 661–678.

[27] Motahhare Eslami, Sneha R. Krishna Kumaran, Christian Sandvig, and Karrie Karahalios. 2018. Communicating Algorithmic Process in Online Behavioral Advertising. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI)*. 1–13.

[28] Ayman Farahat and Michael C Bailey. 2012. How effective is targeted advertising?. In *Proceedings of the 21st international conference on World Wide Web*. 111–120.

[29] Saikat Guha, Bin Cheng, and Paul Francis. 2010. Challenges in measuring online advertising systems. In *Proceedings of the 10th ACM SIGCOMM conference on Internet measurement*. 81–87.

[30] C. Hutto and Eric Gilbert. 2014. VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. *Proceedings of the International AAAI Conference on Web and Social Media* 8, 1 (May 2014), 216–225. https://ojs.aaai.org/index.php/ICWSM/article/view/14550

[31] Basileal Imana, Aleksandra Korolova, and John Heidemann. 2021. Auditing for Discrimination in Algorithms Delivering Job Ads *(WWW '21)*. Association for Computing Machinery, New York, NY, USA, 3767–3778. https://doi.org/10.1145/3442381.3450077

[32] Mike Isaac and Cecilia Kang. 2020. Facebook Says It Won't Back Down From Allowing Lies in Political Ads. https://www.nytimes.com/2020/01/09/technology/facebook-political-ads-lies.html

[33] Matthew Johnston. 2020. How Facebook Makes Money. https://www.investopedia.com/ask/answers/120114/how-does-facebook-fb-make-money.asp

[34] Surya Mattu Julia Angwin. 2016. Facebook Doesn't Tell Users Everything It Really Knows About Them. https://www.propublica.org/article/facebook-doesnt-tell-users-everything-it-really-knows-about-them

[35] Jeremy Lee. 2017. Should advertisers be worried about Facebook's false audience figures? https://www.campaignlive.co.uk/article/advertisers-worried-facebooks-false-audience-figures/1444112

[36] Aleecia M McDonald and Lorrie Faith Cranor. 2010. Americans' attitudes about internet behavioral advertising practices. In *Proceedings of the 9th annual ACM Workshop on Privacy in the Electronic Society (WPES)*. 63–72.

[37] Lee Rainie, Sara Kiesler, Ruogu Kang, Mary Madden, Maeve Duggan, Stephanie Brown, and Laura Dabbish. 2013. Anonymity, privacy, and security online. *Pew Research Center* 5 (2013).

[38] Kimberley Read and Karl van der Schyff. 2020. Modelling the intended use of Facebook privacy settings. *South African Journal of Information Management* 22 (00 2020), 1 – 9. http://www.scielo.org.za/scielo.php?script=sci_arttext&pid=S1560-683X2020000100021&nrm=iso

[39] Guogen Shan and Shawn Gerstenberger. 2017. Fisher's exact approach for post hoc analysis of a chi-squared test. *PloS one* 12, 12 (2017).

[40] Till Speicher, Muhammad Ali, Giridhari Venkatadri, Filipe Nunes Ribeiro, George Arvanitakis, Fabrício Benevenuto, Krishna P. Gummadi, Patrick Loiseau, and Alan Mislove. 2018. Potential for Discrimination in Online Targeted Advertising. In *Proceedings of the 1st Conference on Fairness, Accountability and Transparency (Proceedings of Machine*

Learning Research), Sorelle A. Friedler and Christo Wilson (Eds.), Vol. 81. PMLR, 5–19. https://proceedings.mlr.press/v81/speicher18a.html

[41] Nitasha Tiku. 2020. Facebook has a prescription: More pharmaceutical ads. https://www.washingtonpost.com/technology/2020/03/03/facebook-pharma-ads/

[42] Joseph Turow, Michael Hennessy, and Nora Draper. 2015. The tradeoff fallacy: How marketers are misrepresenting American consumers and opening them up to exploitation. *Available at SSRN 2820060* (2015).

[43] Giridhari Venkatadri, Elena Lucherini, Piotr Sapiezynski, and Alan Mislove. 2019. Investigating sources of PII used in Facebook's targeted advertising. *Proceedings on Privacy Enhancing Technologies* 2019, 1 (2019), 227 – 244.

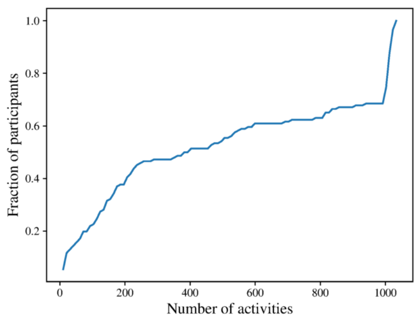
[44] Giridhari Venkatadri, Alan Mislove, and Krishna P. Gummadi. 2018. Treads: Transparency-Enhancing Ads. In *Proceedings of the 17th ACM Workshop on Hot Topics in Networks (HotNets)*. 169–175.

[45] Giridhari Venkatadri, Piotr Sapiezynski, Elissa M. Redmiles, Alan Mislove, Oana Goga, Michelle Mazurek, and Krishna P. Gummadi. 2019. Auditing Offline Data Brokers via Facebook's Advertising Platform. In *Proceedings of the 28th International Conference on World Wide Web (WWW)*. 1920–1930.

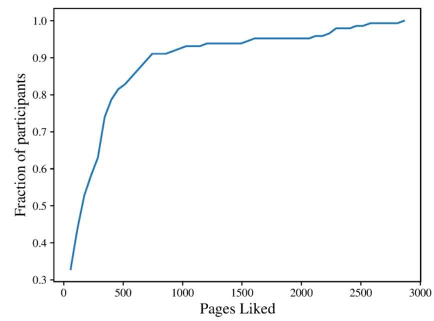
[46] Craig E. Wills and Can Tatar. 2012. Understanding What They Do with What They Know. In *Proceedings of the 2012 ACM Workshop on Privacy in the Electronic Society (WPES)*. 13–18.

APPENDIX

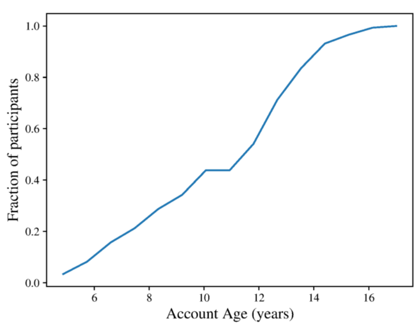
A BASIC CHARACTERISTICS OF OUR PARTICIPANTS' FACEBOOK ACCOUNT



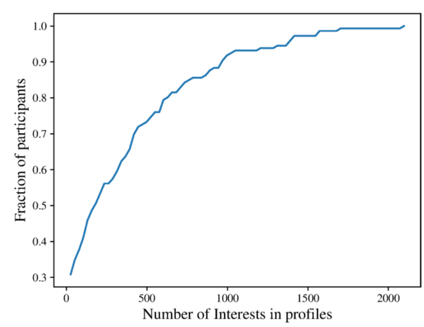
(a) Number of activities



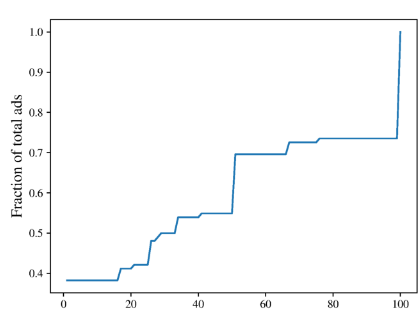
(b) Number of pages liked



(c) Account ages of participants



(d) Number of interests in ad profiles



(e) Percentage of interests marked as irrelevant per ad.

Fig. 8. CDFs of participants' account profile and percentage of irrelevant interests per ad

B LIST OF INTERACTIONS FOR ACCOUNT 1 AND 2

Table 13. Account 1 and 2 interaction table

Activities*	Topic	Pages interacted †	Related Interests (Pakistan) †	Unrelated interests (Pakistan) †	Related Interests (US) †	Unrelated interests (US) †
Page like	Baking	My Baking Addiction, Dessert Recipes	Baking, Desserts, My Baking Addiction		Baking, Desserts, Food, My Baking Addiction	
	City	Islamabad		Multan, URDU		
	Clothing Brands	Diners Clothing, Outfitters	OUTFITTERS, Outfitter, Diners		OUTFITTERS, Outfitter, Diners	
	Electronics	Apple, Samsung	Samsung, Apple Inc., Apple, Multinational corporation, Conglomerate (company), Samsung Electronics		Samsung, Apple Inc., Apple	
	Food	Food Directory Pakistan, Pakistani Food	Pakistani Cuisine		Pakistani Cuisine	
	Hotels	Marriott Hotel, PC Hotel Avari Lahore, Ramada Hotel	Marriott Hotels and Resorts	Pakistan, Lahore, Punjab Pakistan, Gross domestic product, Multan	Marriott Hotels and Resorts	Pakistan, Lahore, Punjab Pakistan, Gross domestic product, Multan
	News Channel	Express News, BBC, CNN	CNN, Current events, BBC, Express News (Pakistan)		CNN, Current events, BBC, Express News (Pakistan)	
	Pizza	Domino's	Pizza, Domino's Pizza, Pizza Hut, Restaurants	Franchising, Yum! Brands	Pizza, Domino's Pizza, Pizza Hut, Italian cuisine	cheese
Like/love react	Car	BMW, Mercedes, Ferrari	Automobiles, Mercedes-Benz, Automotive Industry, BMW, Ferrari			
	Cricket	Sky Sports Cricket, PTV Cricket	PTV Sports			
	Culture	BBC Culture, World Culture Forum	Culture,	Member states of the United Nations, Languages of Pakistan, National Language, Otaku, Tokyo, Tokyo Otaku Mode, Prefectures of japan		
	Decor	Decor by Ihsan, Everlasting Decor	Weddings, Decorative arts, Interior design	Design		
	Gym	UFC Gym, Gym feed, Yoga.com	Yoga.com			
	Perfume	Perfume.com, Fragrance Direct	Fragrances, Fragrance Direct, Perfume.com	Odor		
	Personality	Mian Nawaz Sharif		Prime minister of Pakistan	Human, Man, imran khan official	
	University	LUMS, FAST, NUST	NUST, LUMS Academic degree		Academic degree	
	Watches	Rado, Rolex, Blancpain	Blancpain			
Positive Comments	AI Assistants	Google Assistant, Amazon Alexa, Siri, cortana	cortana, Amazon Echo, Artificial intelligence, Amazon.com	Android (operating system)	Amazon Echo, Amazon.com, Siri, cortana, Artificial intelligence	Multinational corporation, Android (operating system), Shopping
	Animals	Cats and Kittens, Dogs Lovers, The Rabbit haven	Dogs, Cats and Kittens, Rabbits, Hamster	Mammal	Dogs, Kitten, Cats, Cats and Kittens, Cats, Rabbits, Hamster	Mammal
	Bikes	Harley Davidson, Ducati, Kawasaki	Ducati, Ducati Multistrada, Ducati Monster, KAWASAKI V-twin engine, Ducati Apollo, Ducati Desmosedici RR, Kawasaki Heavy Industries, Ducati Pantah, Ducati 851, Ducati MH900e, Motorcycles, Ducati Desmosedici, Types of motorcycles, Sport bike	Grand Prix motorcycle, insurance	Ducati, Ducati Multistrada, Ducati Monster, KAWASAKI V-twin engine, Ducati Apollo, Ducati Desmosedici RR, Kawasaki Heavy Industries, Ducati Pantah, Ducati 851, Ducati MH900e, Motorcycles Ducati Desmosedici, Types of motorcycles, Sport bike	Grand Prix motorcycle racing
	Deodorant	Axe, Degree, Dove	Dove Men+Care, Deodorant	Dove (chocolate)	Dove Men+Care, Deodorant, Dove (toiletries)	
	Furniture	IKEA, Liberty Furniture, Stanley Furniture	Furniture, IKEA, Ready-to-assemble furniture	Retail, Liberty (department store)	Furniture, IKEA, Ready-to-assemble furniture	Retail, Liberty (department store)
	Medicine	Panadol, Medicine, Diazepampillsuk	Medicine			
	Novels	Harry Potter, Game of Thrones, Warner Bros., Rupert Grint	Harry Potter, Game of Thrones, Daniel Radcliffe, Fantasy films, Harry Potter (film series), Television programme	London, coming-of-age story, Televisions, Funny or Die		Funny or Die
	Scientists	Albert Einstein, Issac Newton, Marie Curie	Marie Curie, Albert Einstein		Marie Curie, Albert Einstein	
	Singer	AKON, Shakira, Rihanna, Inna	Pop music, Shakira, Akon Popular music, World music	Arts and music, Eurodance	Akon, Shakira, Pop music, Popular music, World music	Eurodance
	Studio	Coke Studio, Nescafe Basement Pepsi battles of bands	Coke Studio (Pakistan)	9GAG, Soft drinks, Haven (TV series), Maribel Verdu,	Coke Studio (Pakistan)	Soft drinks, Pepsi, 9GAG, Maribel Verdu, Electronic music
	Tech (Software)	Google, Facebook, Microsoft	Microsoft, Facebook, Online, Google, Social network, List of google products, Cloud computing	NASDAQ-100, Dow Jones, Microsoft Developer, Industrial Average, Software developer, Computer hardware	Google, Facebook, Online, Social network, List of Google products, Cloud computing	NASDAQ-100
Scroll only	Airline	PIA, Emirates, American Airlines	Emirates (airline), The Emirates Group			
	Board game	Ludo, Chess				
	Buildings	Burj khalifa, World Trade Center	One World Trade Center, Burj Khalifa			
	Construction	Vinci, Power China, Strabag				
	Guitar	Yamaha, Fender, Gibson Guitars				
	Online transit	Zipcar, Urbvan				
	Ride services	Uber, Careem, Lyft				
	Soap	Dettol, Lifebouy				
	Space exploration	NASA, Space X				
	Superhero	Superman, Spiderman				
	Tea	Lipton, Tapal, Vital Tea				
	Tech news (web)	The Next Web, The Verge, Engadget				
	Wall paint	Nippon Paints, Berger Paints				
	Weapons	Weapons World, Future Weapons, Weapon Lovers				

* Except for liking a page, all other activities were performed without liking the respective page. † Blank entries means no interests were inferred.

C LIST OF INTERACTIONS FOR ACCOUNT 3 AND 4

Table 14. Account 3 and 4 interaction table

Activities*	Topic	Pages interacted †	Related Interests (Pakistan) †	Unrelated interests (Pakistan) †	Related Interests (US) †	Unrelated interests (US) †
Page like	Baking	My Baking Addiction, Dessert Recipes	Baking, Desserts, My Baking Addiction		Baking, Desserts, Food, My Baking Addiction	
	City	Islamabad		Multan		
	Clothing Brands	Diners Clothing, Outfitters	OUTFITTERS, Outfitter, Diners		OUTFITTERS, Outfitter, Diners	
	Electronics	Apple, Samsung	Samsung, Apple Inc., Apple, Multinational corporation, Conglomerate (company), Samsung Electronics		Samsung, Apple Inc., Apple	
	Food	Food Directory Pakistan Pakistani Food,	Pakistani Cuisine		Pakistani Cuisine	
	Hotels	Marriott Hotel, Avari Hotel Ramada Hotel	Marriott Hotels and Resorts	Pakistan, Lahore, Punjab Pakistan Multan, Gross domestic product	Marriott Hotels and Resorts	Pakistan, Lahore, Punjab Pakistan Multan, Gross domestic product
	News Channel	Express News	CNN, Current events, BBC, Express News (Pakistan)		CNN, Current events, BBC, Express News (Pakistan)	
	Pizza	Domino's	Pizza, Domino's Pizza, Pizza Hut, Restaurants	Franchising, Yum! Brands	Pizza, Domino's Pizza, Pizza Hut, Italian cuisine	cheese
Like/love react	Car	BMW, Mercedes, Ferrari	Automobiles, Mercedes-Benz, Automotive Industry, BMW, Ferrari			
	Cricket	Sky Sports Cricket, PTV Cricket				
	Culture	BBC Culture, World Culture Forum	Culture,	Member states of the United Nations, Languages of Pakistan, National Language, Otaku, Japan, Japan Otaku Mode, Prefectures of Japan		
	Decor	Decor by Ihsan, Everlasting Decor	Weddings, Decorative arts, Interior design	Design		
	Gym	UFC Gym, Gym feed, Yoga.com	Yoga.com			
	Perfume	Perfume.com, Fragrance Direct	Fragrances, Fragrance Direct, Perfume.com	Odor		
	Personality	Mian Nawaz Sharif		Prime minister of Pakistan	imran khan official	Human, Man
	University	LUMS, FAST, NUST	Lahore University of Management Sciences, LUMS, Academic degree		Academic degree	
	Watches	Rado, Rolex, Blancpain	Blancpain			
Positive Comments	AI Assistants	Google Assistant, Amazon Alexa, Siri, cortana	cortana, Amazon Echo, Artificial intelligence, Amazon.com	Android (operating system)	Amazon Echo, Amazon.com, Siri, cortana, Artificial intelligence	Multinational corporation, Android (operating system), Shopping
	Animals	Cats and Kittens, Dogs Lovers, The Rabbit haven	Dogs, Cats and Kittens, Rabbits, Hamster	Mammal	Dogs, Kitten, Cats, Cats and Kittens, Cats, Rabbits, Hamster	Mammal
	Bikes	Harley Davidson, Ducati, Kawasaki	Ducati, Ducati Multistrada, Ducati Monster, KAWASAKI V-twin engine, Ducati Apollo, Ducati Desmosedici RR, Kawasaki Heavy Industries, Ducati Pantah, Ducati 851, Ducati MH900e, Motorcycles, Ducati Desmosedici, Types of motorcycles, Sport bike	Grand Prix motorcycle, insurance	Ducati, Ducati Multistrada, Ducati Monster, KAWASAKI V-twin engine, Ducati Apollo, Ducati Desmosedici RR, Kawasaki Heavy Industries, Ducati Pantah, Ducati 851, Ducati MH900e, Motorcycles Ducati Desmosedici, Types of motorcycles, Sport bike	Grand Prix motorcycle racing
	Deodorant	Axe, Degree, Dove	Dove Men+Care, Deodorant	Dove (chocolate)	Dove Men+Care, Deodorant, Dove (toiletries)	
	Furniture	IKEA, Liberty Furniture, Stanley Furniture	Furniture, IKEA, Ready-to-assemble furniture	Retail, Liberty (department store)	Furniture, IKEA, Ready-to-assemble furniture	Retail, Liberty (department store)
	Medicine	Panadol, Medicine, Diazepampillsuk	Medicine			
	Novels	Harry Potter, Game of Thrones, Warner Bros., Rupert Grint	Harry Potter, Game of Thrones, Daniel Radcliffe, Fantasy films, Harry Potter (film series), Television programme	London, coming-of-age story, Televisions, Funny or Die		Funny or Die
	Scientists	Albert Einstein, Issac Newton, Marie Curie	Marie Curie, Albert Einstein		Marie Curie, Albert Einstein	
	Singer	AKON, Shakira, Rihanna, Inna	Pop music, Shakira, Akon Popular music, World music	Arts and music, Eurodance	Akon, Shakira, Pop music, Popular music, World music	Eurodance
	Studio	Coke Studio, Nescafe Basement, Pepsi battles of bands	Coke Studio (Pakistan)	9GAG, Soft drinks, Haven (TV series), Maribel Verdu,	Coke Studio (Pakistan)	Soft drinks, Pepsi, 9GAG, Maribel Verdu, Electronic music
	Tech (Software)	Google, Facebook, Microsoft	Microsoft, Facebook, Online, Google, Social network, List of google products, Cloud computing	NASDAQ-100, Dow Jones, Microsoft Developer, Industrial Average, Software developer, Computer hardware	Google, Facebook, Online, Social network, List of Google products, Cloud computing	NASDAQ-100
Scroll only	Airline	PIA, Emirates, American Airlines	Emirates (airline), The Emirates Group			
	Board game	Ludo, Chess				
	Buildings	Burj khalifa, World Trade Center	One World Trade Center, Burj Khalifa			
	Construction	Vinci, Power China, Strabag				
	Guitar	Yamaha, Fender, Gibson Guitars				
	Online transit	Zipcar, Urbvan				
	Ride services	Uber, Careem, Lyft				
	Soap	Dettol, Lifebouy	Lifebuoy			
	Space exploration	NASA, Space X				
	Superhero	Superman, Spiderman				
	Tea	Lipton, Tapal, Vital Tea				
	Tech news (web)	The Next Web, The Verge, Engadget				
	Wall paint	Nippon Paints, Berger Paints				
	Weapons	Weapons World, Future Weapons, Weapon Lovers				

* Except for liking a page, all other activities were performed without liking the respective page. † Blank entries means no interests were inferred.

D LIST OF INTERACTIONS FOR ACCOUNT 5 AND 6

Table 15. Account 5 and 6 interaction table

Activities*	Topic	Pages interacted †	Related Interests (Pakistan) †	Unrelated interests (Pakistan) †	Related Interests (US) †	Unrelated interests (US) †
Page like	Baking	My Baking Addiction, Dessert Recipes	Baking, Desserts, My Baking Addiction		Baking, Desserts, Food, My Baking Addiction	
	City	Islamabad		Urdu, Multan		
	Clothing Brands	Diners Clothing, Outfitters	OUTFITTERS, Outfitter, Diners		OUTFITTERS, Outfitter, Diners	
	Electronics	Apple, Samsung	Samsung, Apple Inc., Apple, Multinational corporation, Conglomerate (company), Samsung Electronics		Samsung, Apple Inc., Apple	
	Food	Food Directory Pakistan, Pakistani Food,	Pakistani Cuisine		Pakistani Cuisine	
	Hotels	Marriott Hotel, Avari Hotel Ramada Hotel	Marriott Hotels and Resorts	Pakistan, Lahore, Punjab Pakistan Multan, Gross domestic product	Marriott Hotels and Resorts	Pakistan, Lahore, Punjab Pakistan Multan, Gross domestic product
	News Channel	Express News	CNN, Current events, BBC, Express News (Pakistan)		CNN, Current events, BBC, Express News (Pakistan)	
	Pizza	Domino's	Pizza, Domino's Pizza, Pizza Hut, Restaurants	Franchising, Yum! Brands	Pizza, Domino's Pizza, Pizza Hut, Italian cuisine	cheese
Like/love react	Car	BMW, Mercedes, Ferrari	BMW, Ferrari			
	Cricket	Sky Sports Cricket, PTV Cricket	PTV Sports			
	Culture	BBC Culture, World Culture Forum	Culture,	Member states of the United Nations, Languages of Pakistan, National Language, Otaku, Japan, Tokyo Otaku Mode, Prefectures of Japan		Otaku, Japan, Tokyo Otaku Mode, Prefectures of Japan
	Decor	Decor by Ihsan, Everlasting Decor	Weddings, Decorative arts, Interior design	Design		
	Gym	UFC Gym, Gym feed, Yoga.com	Yoga.com			
	Perfume	Perfume.com, Fragrance Direct	Fragrances, Fragrance Direct, Perfume.com	Odor		
	Personality	Mian Nawaz Sharif		Prime minister of Pakistan	imran khan official	Human, Man
	University	LUMS, NUST, FAST	Academic degree		Academic degree	
	Watches	Rado, Rolex, Blancpain	Blancpain			
Positive Comments	AI Assistants	Google Assistant, Amazon Alexa, Siri, cortana	cortana, Amazon Echo, Artificial intelligence, Amazon.com	Android (operating system)	Amazon Echo, Amazon.com, Siri, cortana, Artificial intelligence	Multinational corporation, Android (operating system), Shopping
	Animals	Cats and Kittens, Dogs Lovers, The Rabbit haven	Dogs, Cats and Kittens, Rabbits, Hamster	Mammal	Dogs, Kitten, Cats, Cats and Kittens, Cats, Rabbits, Hamster	Mammal
	Bikes	Harley Davidson, Ducati, Kawasaki	Ducati, Ducati Multistrada, Ducati Monster, KAWASAKI V-twin engine, Ducati Apollo, Ducati Desmosedici RR, Kawasaki Heavy Industries, Ducati Pantah, Ducati 851, Ducati MH900e, Motorcycles, Ducati Desmosedici, Types of motorcycles, Sport bike	Grand Prix motorcycle, insurance	Ducati, Ducati Multistrada, Ducati Monster, KAWASAKI V-twin engine, Ducati Apollo, Ducati Desmosedici RR, Kawasaki Heavy Industries, Ducati Pantah, Ducati 851, Ducati MH900e, Motorcycles Ducati Desmosedici, Types of motorcycles, Sport bike	Grand Prix motorcycle racing
	Deodorant	Axe, Degree, Dove	Dove Men+Care, Deodorant	Dove (chocolate)	Dove Men+Care, Deodorant, Dove (toiletries)	
	Furniture	IKEA, Liberty Furniture, Stanley Furniture	Furniture, IKEA, Ready-to-assemble furniture	Retail, Liberty (department store)	Furniture, IKEA, Ready-to-assemble furniture	Retail, Liberty (department store)
	Medicine	Panadol, Medicine, Diazepampillsuk	Medicine			
	Novels	Harry Potter, Game of Thrones, Warner Bros., Rupert Grint	Harry Potter, Game of Thrones, Daniel Radcliffe, Fantasy films, Harry Potter (film series), Television programme	London, coming-of-age story, Televisions, Funny or Die		Funny or Die
	Scientists	Albert Einstein, Issac Newton, Marie Curie	Marie Curie, Albert Einstein		Marie Curie, Albert Einstein	
	Singer	AKON, Shakira, Rihanna, Inna	Pop music, Shakira, Akon Popular music, World music	Arts and music, Eurodance	Akon, Shakira, Pop music, Popular music, World music	Eurodance
	Studio	Coke Studio, Nescafe Basement Pepsi battles of bands	Coke Studio (Pakistan)	9GAG, Soft drinks, Haven (TV series), Maribel Verdu,	Coke Studio (Pakistan)	Soft drinks, Pepsi, 9GAG, Maribel Verdu, Electronic music
	Tech (Software)	Google, Facebook, Microsoft	Microsoft, Facebook, Online, Google, Social network, List of google products, Cloud computing	NASDAQ-100, Dow Jones, Microsoft Developer, Industrial Average, Software developer, Computer hardware	Google, Facebook, Online, Social network, List of Google products, Cloud computing	NASDAQ-100
Scroll only	Airline	PIA, Emirates, American Airlines				
	Board game	Ludo, Chess				
	Buildings	Burj khalifa, World Trade Center	World Trade Center, Burj Khalifa			
	Construction	Vinci, Power China, Strabag				
	Guitar	Yamaha, Fender, Gibson Guitars				
	Online transit	Zipcar, Urbvan	Zipcar			
	Ride services	Uber, Careem, Lyft				
	Soap	Dettol, Lifebouy				
	Space exploration	NASA, Space X				
	Superhero	Superman, Spiderman				
	Tea	Lipton, Tapal, Vital Tea				
	Tech news (web)	The Next Web, The Verge, Engadget				
	Wall paint	Nippon Paints, Berger Paints				
	Weapons	Weapons World, Future Weapons, Weapon Lovers				

* Except for liking a page, all other activities were performed without liking the respective page. † Blank entries means no interests were inferred.

E LIST OF INTERACTIONS FOR ACCOUNT 7 AND 8

Table 16. Account 7 and 8 interaction table

Activities*	Topic	Pages interacted †	Related Interests (Pakistan) †	Unrelated interests (Pakistan) †	Related Interests (US) †	Unrelated interests (US) †
Page like	Baking	My Baking Addiction, Dessert Recipes,	Baking, Desserts, My Baking Addiction		Baking, Desserts, Food, My Baking Addiction	
	City	Islamabad		Multan		
	Clothing Brands	Diners Clothing, Outfitters	OUTFITTERS, Outfitter, Diners		OUTFITTERS, Outfitter, Diners	
	Electronics	Apple, Samsung	Samsung, Apple Inc., Apple, Multinational corporation, Conglomerate (company), Samsung Electronics		Samsung, Apple Inc., Apple	
	Food	Food Directory Pakistan, Pakistani Food,	Pakistani Cuisine		Pakistani Cuisine	
	Hotels	Marriott Hotel, Avari Hotel Ramada Hotel	Marriott Hotels and Resorts	Pakistan, Lahore, Punjab Pakistan Multan, Gross domestic product	Marriott Hotels and Resorts	Pakistan, Lahore, Punjab Pakistan Multan, Gross domestic product
	News Channel	Express News	CNN, Current events, BBC, Express News (Pakistan)		CNN, Current events, BBC, Express News (Pakistan)	
	Pizza	Domino's	Pizza, Domino's Pizza, Pizza Hut, Restaurants	Franchising, Yum! Brands	Pizza, Domino's Pizza, Pizza Hut, Italian cuisine	cheese
Like/love react	Car	BMW, Mercedes, Ferrari	Automobiles, Automotive Industry, BMW, Ferrari, Mercedes-Benz			
	Cricket	Sky Sports Cricket, PTV Cricket				
	Culture	BBC Culture, World Culture Forum	Culture,	Member states of the United Nations, Languages of Pakistan, National Language, Otaku, Japan, Tokyo Otaku Mode, Prefectures of Japan		Otaku, Japan, Tokyo Otaku Mode, Prefectures of Japan
	Decor	Decor by Ihsan, Everlasting Decor	Weddings, Decorative arts, Interior design	Design		
	Gym	UFC Gym, Gym feed, Yoga.com	Yoga.com			
	Perfume	Perfume.com, Fragrance Direct	Fragrances, Perfume.com	Odor		
	Personality	Mian Nawaz Sharif		Prime minister of Pakistan	imran khan official	Human, Man
	University	LUMS, NUST, FAST	Lahore University of Management Sciences, LUMS		Academic degree	
	Watches	Rado, Rolex, Blancpain	Rolex, Blancpain			
Positive Comments	AI Assistants	Google Assistant, Amazon Alexa, Siri, cortana	cortana, Amazon Echo, Artificial intelligence, Amazon.com	Android (operating system)	Amazon Echo, Amazon.com, Siri, cortana, Artificial intelligence	Multinational corporation, Android (operating system), Shopping
	Animals	Cats and Kittens, Dogs Lovers, The Rabbit haven	Dogs, Cats and Kittens, Rabbits, Hamster	Mammal	Dogs, Kitten, Cats, Cats and Kittens, Cats, Rabbits, Hamster	Mammal
	Bikes	Harley Davidson, Ducati, Kawasaki	Ducati, Ducati Multistrada, Ducati Monster, KAWASAKI V-twin engine, Ducati Apollo, Ducati Desmosedici RR, Kawasaki Heavy Industries, Ducati Pantah, Ducati 851, Ducati MH900e, Motorcycles, Ducati Desmosedici, Types of motorcycles, Sport bike	Grand Prix motorcycle, insurance	Ducati, Ducati Multistrada, Ducati Monster, KAWASAKI V-twin engine, Ducati Apollo, Ducati Desmosedici RR, Kawasaki Heavy Industries, Ducati Pantah, Ducati 851, Ducati MH900e, Motorcycles Ducati Desmosedici, Types of motorcycles, Sport bike	Grand Prix motorcycle racing
	Deodorant	Axe, Degree, Dove	Dove Men+Care, Deodorant	Dove (chocolate)	Dove Men+Care, Deodorant, Dove (toiletries)	
	Furniture	IKEA, Liberty Furniture, Stanley Furniture	Furniture, IKEA, Ready-to-assemble furniture	Retail, Liberty (department store)	Furniture, IKEA, Ready-to-assemble furniture	Retail, Liberty (department store)
	Medicine	Panadol, Medicine, Diazepampillsuk	Medicine			
	Novels	Harry Potter, Game of Thrones, Warner Bros., Rupert Grint	Harry Potter, Game of Thrones, Daniel Radcliffe, Fantasy films, Harry Potter (film series), Television programme	London, coming-of-age story, Televisions, Funny or Die		Funny or Die
	Scientists	Albert Einstein, Issac Newton, Marie Curie	Marie Curie, Albert Einstein		Marie Curie, Albert Einstein	
	Singer	AKON, Shakira, Rihanna, Inna	Pop music, Shakira, Akon Popular music, World music	Arts and music, Eurodance	Akon, Shakira, Pop music, Popular music, World music	Eurodance
	Studio	Coke Studio, Nescafe Basement Pepsi battles of bands	Coke Studio (Pakistan)	9GAG, Soft drinks, Haven (TV series), Maribel Verdu,	Coke Studio (Pakistan)	Soft drinks, Pepsi, 9GAG, Maribel Verdu, Electronic music
	Tech (Software)	Google, Facebook, Microsoft	Microsoft, Facebook, Online, Google, Social network, List of google products, Cloud computing	NASDAQ-100, Dow Jones, Microsoft Developer, Industrial Average, Software developer, Computer hardware	Google, Facebook, Online, Social network, List of Google products, Cloud computing	NASDAQ-100
Scroll only	Airline	PIA, Emirates, American Airlines				
	Board game	Ludo, Chess	Chess			
	Buildings	Burj khalifa, World Trade Center	World Trade Center, Burj Khalifa			
	Construction	Vinci, Power China, Strabag				
	Guitar	Yamaha, Fender, Gibson Guitars	Fender Musical Instruments Corporation, Fender Guitar, amplifier, Bass guitar Yamaha Corporation			
	Online transit	Zipcar, Urbvan	Zipcar			
	Ride services	Uber, Careem, Lyft				
	Soap	Dettol, Lifebouy				
	Space exploration	NASA, Space X				
	Superhero	Superman, Spiderman	Spider-Man, DC Comics, American comic book, Superman, Superhero, Kryptonian	Scholastic Corporation		
	Tea	Lipton, Tapal, Vital Tea				
	Tech news (web)	The Next Web, The Verge, Engadget				
	Wall paint	Nippon Paints, Berger Paints				
	Weapons	Weapons World, Future Weapons, Weapon Lovers				

* Except for liking a page, all other activities were performed without liking the respective page. † Blank entries means no interests were inferred.

F LIST OF INTERACTIONS FOR ACCOUNT 9 AND 10

Table 17. Account 9 and 10 (Mixed accounts) interaction table

Activities*	Topic	Pages interacted †	Related Interests (Pakistan) †	Unrelated interests (Pakistan) †	Related Interests (US) †	Unrelated interests (US) †
Page Like	Baking	My Baking Addiction, Dessert Recipes	My Baking Addiction, Baking, Desserts		My Baking Addiction, Baking, Desserts, Food	
	Electronics	Apple, Samsung	Samsung, Apple Inc., Apple Multinational corporation, Conglomerate (company)		Samsung, Apple Inc., Apple	
	Food	Food Directory Pakistan, Pakistani Food	Pakistani Cuisine		Pakistani Cuisine	
	Hotels	Marriott Hotel, Avari Hotel Ramada Hotel	Marriott Hotels and Resorts	Pakistan, Lahore, Punjab Pakistan, Multan, Gross domestic product, Karachi	Marriott Hotels and Resorts	Pakistan, Lahore, Punjab Pakistan Multan, Karachi Gross domestic product
	Pizza	Domino's	Pizza, Domino's Pizza, Pizza Hut	Domino's Pizza, Pizza Hut	Pizza, Italian cuisine,	cheese
	University	LUMS, NUST, FAST	Academic degree, NUST, Lahore University of Management Sciences, National University of Computer and Emerging Sciences, LUMS	Foundation (non-profit)	Academic degree, NUST, Lahore University of Management Sciences, National University of Computer and Emerging Sciences, LUMS	
Angry React	Car	BMW, Mercedes, Ferrari	Mercedes-Benz, BMW, Ferrari, Automotive Industry, Automobiles			
	Cricket	Sky Sports Cricket, PTV Cricket				
	Perfume	Perfume.com, Fragrance Direct	Fragrances, Fragrance Direct, Perfume.com	Odor		
	Shoes	Nike, Borjan	Shoes			
	Watches	Rado, Rolex, Blancpain	Blancpain			
Positive Comment	Animals	Cats and Kittens, Dogs Lovers	Dogs, Kitten, Cats, Rabbits, Hamster, Mammal Cats And Kittens		Kitten, Cats, Rabbits, Hamster, Mammal Cats and Kittens	
	Bikes	Harley Davidson, Ducati, Kawasaki	Ducati, Ducati Multistrada, Ducati Monster, KAWASAKI V-twin engine, Ducati Apollo, Ducati Desmosedici RR, Kawasaki Heavy Industries, Ducati Pantah, Ducati 851, Ducati MH900e, Motorcycles, Ducati Desmosedici, Types of motorcycles, Sport bike	Grand Prix motorcycle, insurance	Ducati, Ducati Multistrada, , Ducati Monster, KAWASAKI V-twin engine, Ducati Apollo, Ducati Desmosedici RR, Kawasaki Heavy Industries, Ducati Pantah, Ducati 851, Ducati MH900e, Motorcycles Ducati Desmosedici, Types of motorcycles, Sport bike	Grand Prix motorcycle racing
	Deodorant	Axe, Degree, Dove	Dove Men+Care, Deodorant	Dove (chocolate), Brand Deodorant	Dove (toiletries), Dove Men+Care,	
	Scientists	Albert Einstein, Issac Newton, Marie Curie	Marie Curie, Albert Einstein		Marie Curie, Albert Einstein	
	Singer	AKON, Shakira, Rihanna, Inna	Pop music, Shakira, World music, Akon Popular music	Eurodance	Akon, Shakira, Pop music, Popular music, World music	Eurodance
Negative Comment	AI Assistants, Shopping	Google Assistant, Amazon Alexa, Siri, Cortana	cortana, Amazon Echo, Siri, Artificial intelligence, Amazon.com	Android (operating system)	Amazon Echo, Amazon.com, Siri, cortana, Artificial intelligence	Multinational corporation, Android (operating system), Shopping
	Furniture	IKEA, Liberty Furniture, Stanley Furniture	Furniture, IKEA, Ready-to-assemble furniture	Liberty (department store), Retail	Furniture, IKEA, Ready-to-assemble furniture	Liberty (department store), Retail
	Novels	Game of Thrones, Harry Potter	Harry Potter, Game of Thrones, Harry Potter (film series), Daniel Radcliffe, Fantasy films, Television programme, Warner Bros., Rupert Grint	London, coming-of-age story, Televisions, Funny or Die	Harry Potter, Game of Thrones, Harry Potter (film series), Daniel Radcliffe, Fantasy films, Television programme, Warner Bros., Rupert Grint	London, coming-of-age story Funny or Die, Televisions
	Studio	Coke Studio, Nescafe Basement Pepsi battles of bands	Coke Studio (Pakistan)	9GAG, Soft drinks, Pepsi, electronic music, Maribel Verdu, Haven (TV series)	Coke Studio (Pakistan)	Haven (TV series), Pepsi, Maribel Verdu, Electronic music
	Tech (Software)	Google, Facebook, Microsoft	Microsoft, Facebook, Online, Google, Social network, List of google products, Cloud computing	NASDAQ-100	Google, List of Google products, Facebook, Online, Social network Cloud computing	NASDAQ-100 Multinational corporation

* Except for liking a page, all other activities were performed without liking the respective page. † Blank entries means no interests were inferred.

Received January 2021; revised July 2021; accepted November 2021