How Does Connecting Online Activities to Advertising Inferences Impact Privacy Perceptions?
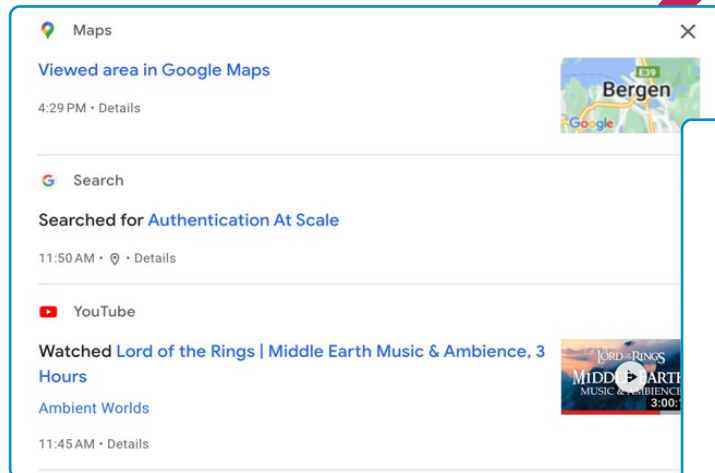
Florian M. Farke, David G. Balash, <u>Maximilian Golla</u>, and Adam J. Aviv
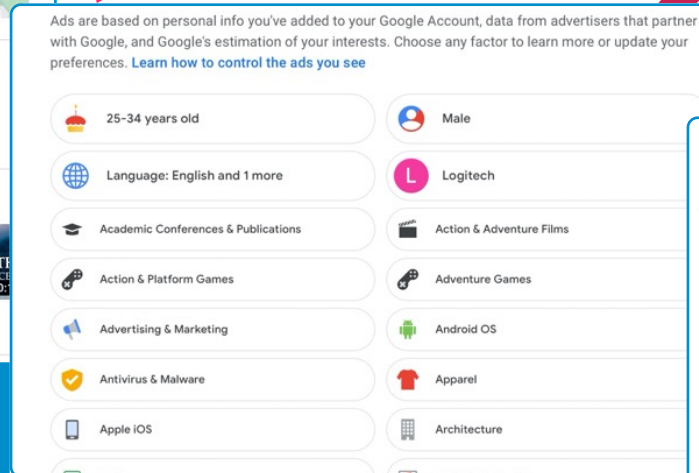
Online Behavioral Advertising

… display highly relevant ads to users based on their browsing behavior.

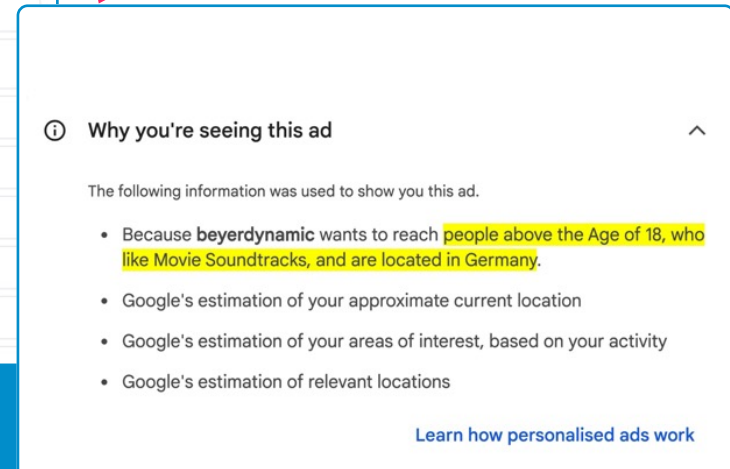# Google's Online Behavioral Advertising 101

1) Activity Collection

Google collects **activities.**

(searches, page visits, map views, and watched videos)

2) Inferencing

Based on these activities, their algorithm derives **interests.**

3) Targeted Ads

Advertiser **target** selected audiences that share certain "interests," or "behaviors."
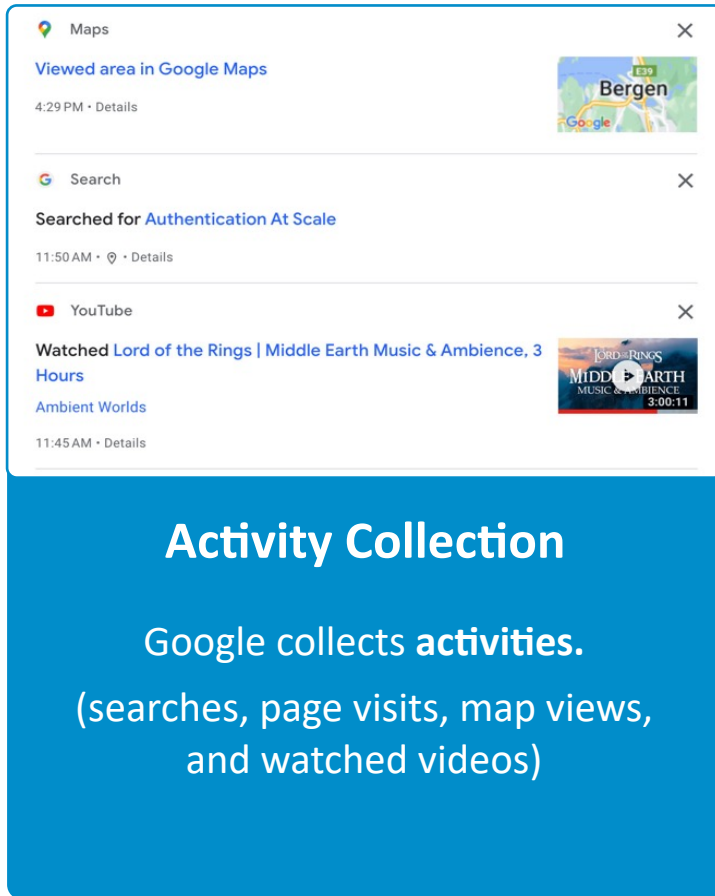
YouTube activities: 17 | 15 | 32 | 48 | 30

Maps activities: 15 | 19 | 45 | 29 | 24

0%   25%   50%   75%

- Absolutely ina...
- Neutral
- Absolutely app...
- Inappropriate
- Slightly appro...
- Slightly ina...
- Appropriat...

Figure 5: The majority of participants found the explan... Google gives as to why they collect activity data appro... (Web: 67%, YouTube: 72%; Maps: 64%).

Understanding of Data Collection. We also reco... mouse movements of the participants during their vis... My Activity page to get an idea of whether and h... interacted with the page. We recorded an average pa... scroll depth of 20553 pixels ($SD = 22285$, $min = 65$... 252735). A single activity height is approximately... which suggests that the average participant scr... ...00 activities during their explorati... ...ty helps to better u...

Prior Work

"Not much I can do to stop it, so why worry?" (P55)

Our Prior Work: Google My Activity

Activity Collection

Google collects **activities.**

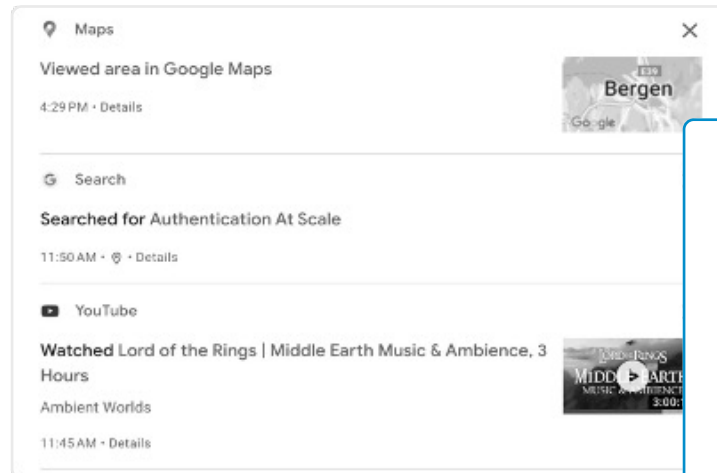(searches, page visits, map views, and watched videos)

Exposure to My Activity, resulted in

- significant **decreases in** perceived **concern** and
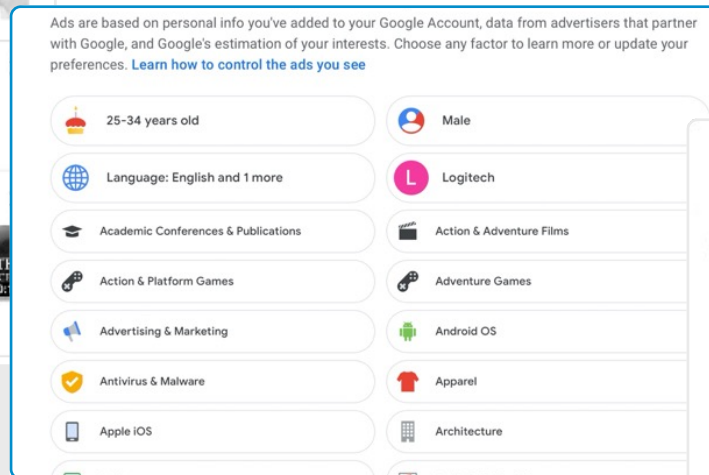- **increase** in **perceived benefit** from data collection
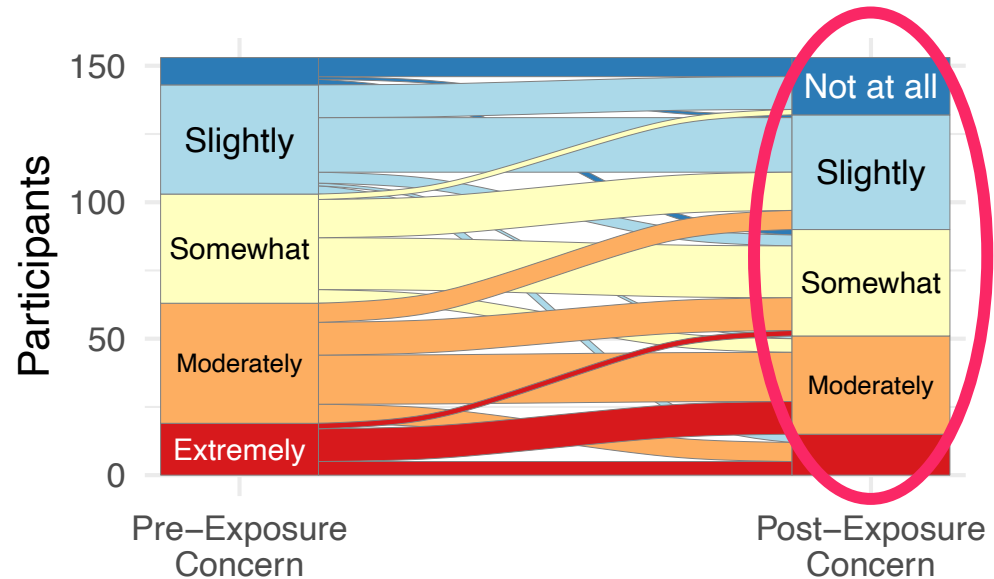
Wilcoxon Sign-Ranked Test W = 2519.5 p < 0.001
with moderate effect r = 0.32

This Work: Google Ad Settings

1) Activity Collection

Google collects **activities.**

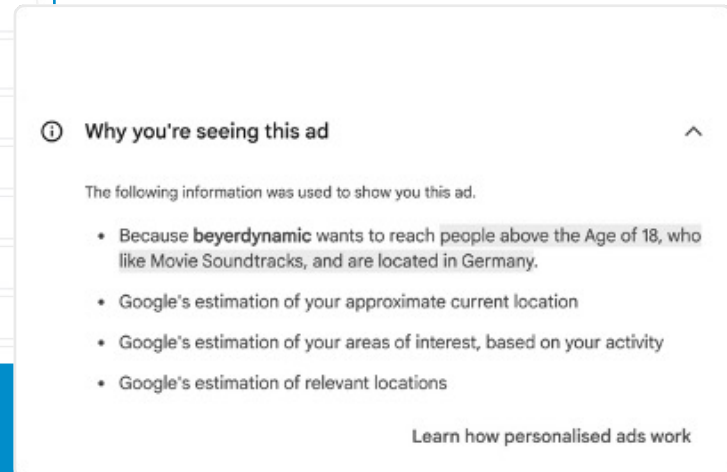(searches, page visits, map views, and watched videos)

2) Inferencing

Based on these activities, their algorithm derives **interests.**

3) Targeted Ads

Advertiser **target** selected audiences that share certain "interests," or "behaviors."

This Work

Ad personalization

Google makes your ads more useful on Google services (such as Search or YouTube), and on websites & apps that partner with Google to show ads. **Learn why you're seeing an ad**

Ad personalization is ON

Advanced

How your ads are personalized

Ads are based on personal info you've added to your Google Account, data from advertisers that partner with Google, and Google's estimation of your interests. Choose any factor to learn more or update your preferences. **Learn how to control the ads you see**

🎂 25-34 years old	👤 Male
🌐 Language: English and 1 more	Ⓛ Logitech
🎓 Academic Conferences & Publications	🎬 Action & Adventure Films
🎮 Action & Platform Games	🎮 Adventure Games
📣 Advertising & Marketing	🤖 Android OS

5 Interest Types, 26 Topics, >2400 Interests

Aggregated
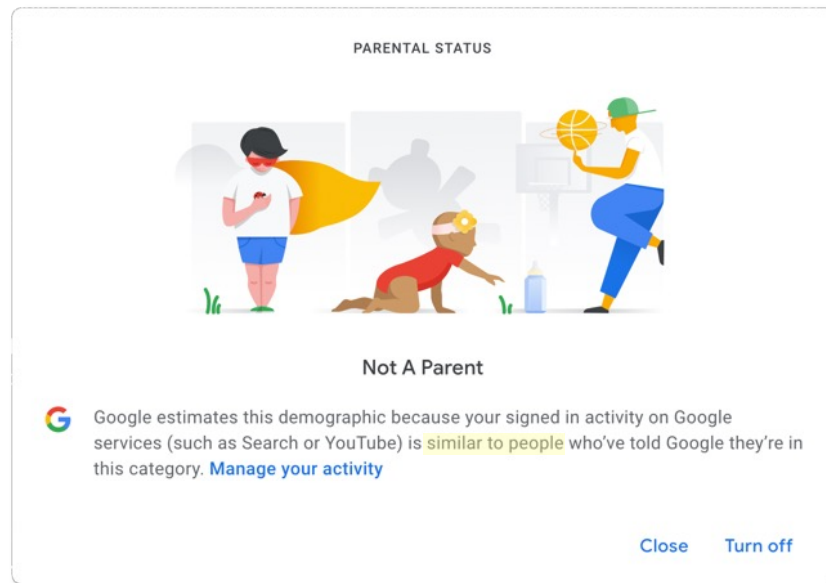
"...similar to people..."

Activity-Based
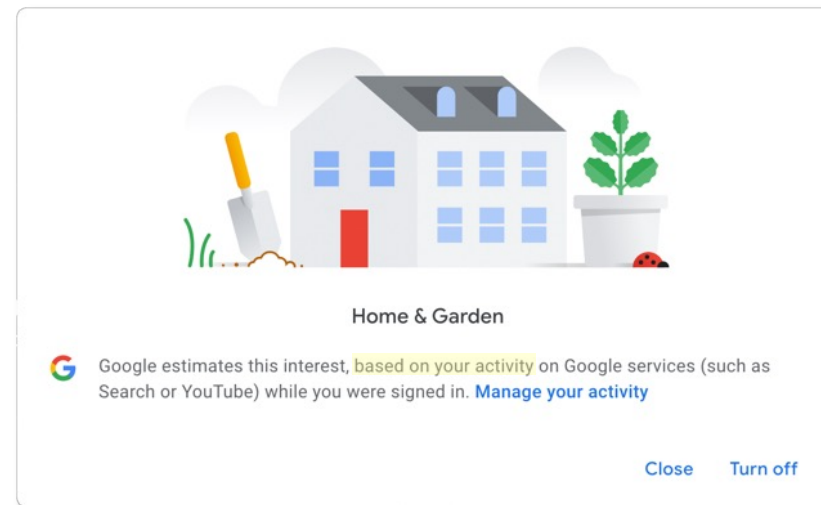
"...based on your activity..."

- ⭐ Celebrities & Entertainment News
- 🛋 Clocks
- ☕ Coffee & Tea
- 👊 Combat Sports
- ✴ Comics & Animation
- 🎮 Computer & Video Games
- 🖥 Computer Drives & Storage
- 💻 Computer Hardware
- 💻 Computer Peripherals
- 🪛 Construction & Power Tools
- 🍲 Cooking & Recipes
- 💳 Credit Cards
- 🚲 Cycling
- 👥 Demographics
- 📕 Dictionaries & Encyclopedias
- 🖥 Distributed & Cloud Computing

Our Prior Work: Disconnected Privacy Dashboards

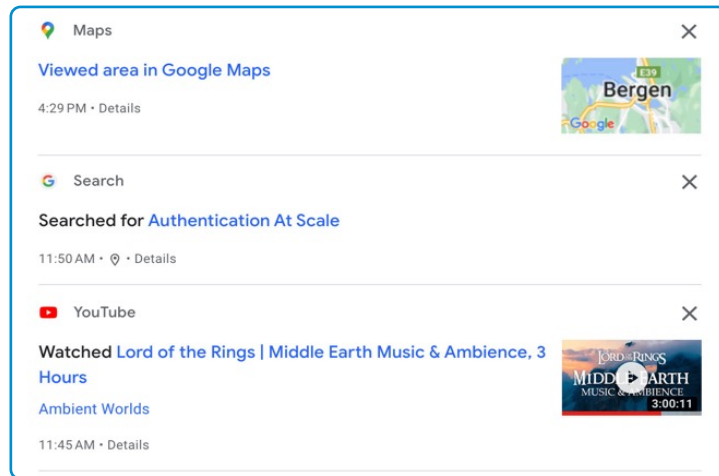**At the time, we thought that...**

...the disconnect of **what data is collected** and **how that data is used** could be responsible for the dramatic shift in concern and benefit.

Collected Activities

Google My Activity
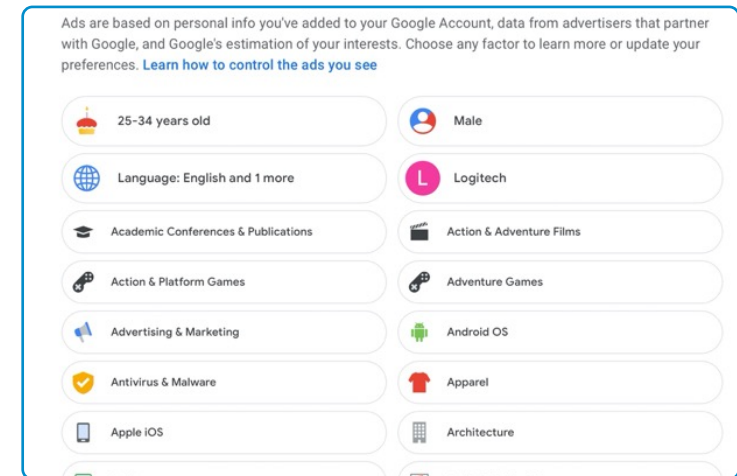
Inferred Interests

Google Ad Settings
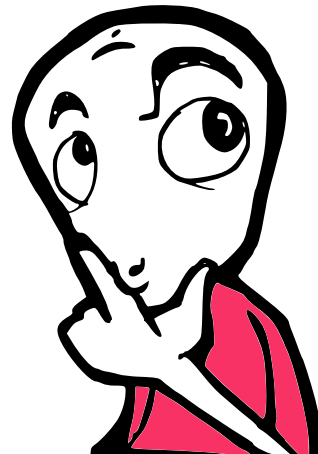
In this Work: Let Participants Connect the Dots!

... **connecting activities to** advertising **interests**, stabilizes perceived benefit of and concern for online data collection.
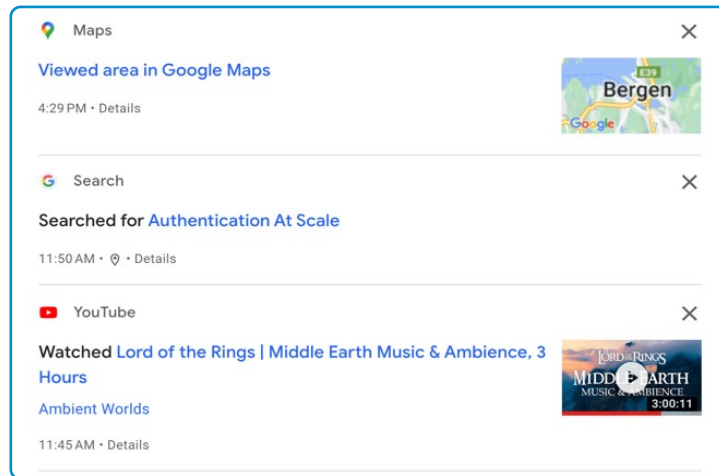
Collected Activities
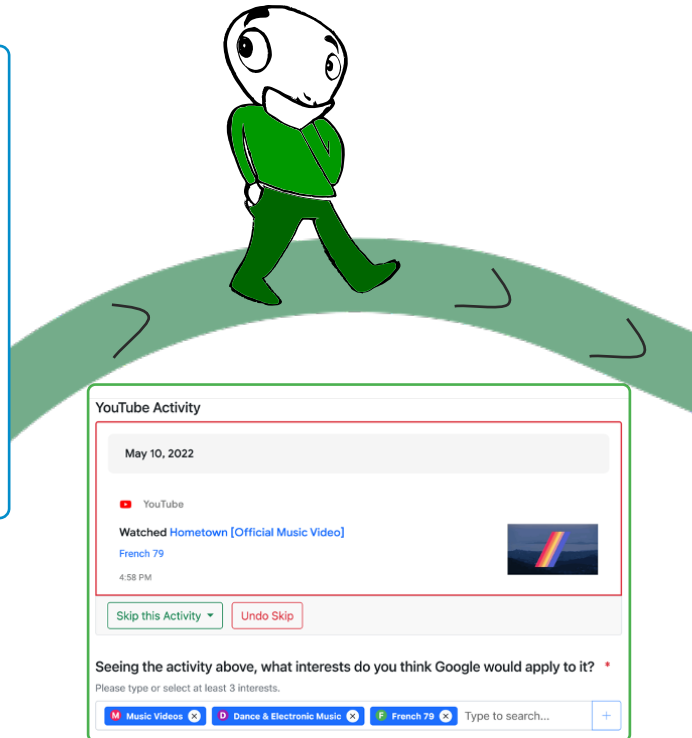
Google My Activity

Labeling Task

Inferred Interests

Google Ad Settings

User Study

"... I find them creepy because it means Google is watching me every step online." (P172)

Study Design

Intervention I

YouTube Activity

May 10, 2022

How to personalize the ads you see on YouTube a...

Manage Ad Settings

Watch on YouTube

[Music Video]

Seeing the activity above, what interests do you think Google would apply to it? *
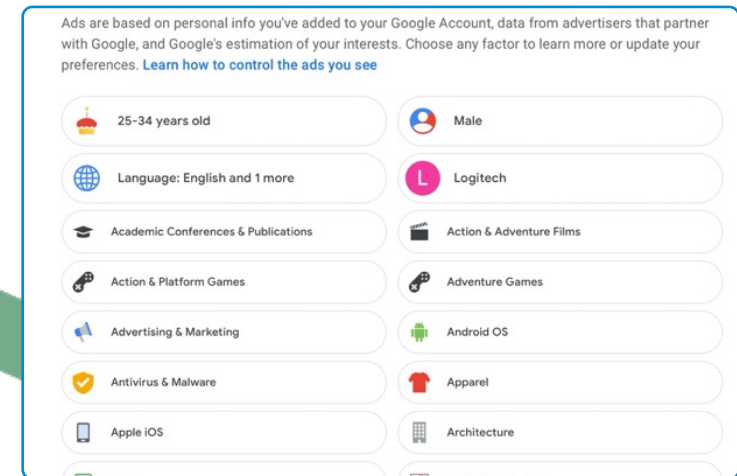Please type or select at least 3 interests.

Music Videos ✕   Dance & Electronic Music ✕   French 79 ✕   Type to search...   +

Intervention II

Ads are based on personal info you've added to your Google Account, data from advertisers that partner with Google, and Google's estimation of your interests. Choose any factor to learn more or update your preferences. Learn how to control the ads you see

25-34 years old	Male
Language: English and 1 more	Logitech
Academic Conferences & Publications	Action & Adventure Films
Action & Platform Games	Adventure Games
Advertising & Marketing	Android OS
Antivirus & Malware	Apparel
Apple iOS	Architecture

Pre-Task Perception → **Labeling Task** → **Post-Task** Perception → **Interests Exposure** → **Post-Exposure** Perception

12

Labeling Task

YouTube Activity

May 10, 2022

▶ YouTube

Watched Hometown [Official Music Video]

French 79

4:58 PM

Skip this Activity ▾   Undo Skip

→ Google My Activity

Seeing the activity above, what interests do you think Google would apply to it? *

Please type or select at least 3 interests.

M Music Videos ⊗   D Dance & Electronic Music ⊗   F French 79 ⊗   Type to search...   +

→ Google Ad Settings
(Labeled by participants)

13

Labeling Task

YouTube Activity

May 10, 2022

▶ YouTube

Watched Hometown [Official Music Video]

French 79

4:58 PM

Skip this Activity ▾   Undo Skip

Seeing the activity above, what interests do you think Google would apply to it? *

Please type or select at least 3 interests.

Type to search...   +

Activity:
Extracted from My Activity
3x {*Search, Maps, YouTube*}
3x {*recent, 3-, 18-months old*}

Labeling Task

YouTube Activity

May 10, 2022

▶ YouTube

Watched Hometown [Official Music Video]
French 79
4:58 PM

Skip this Activity ▾   Undo Skip

Seeing the activity above, what interests do you think Google would apply to it? *

Please type or select at least 3 interests.

Type to search...   +

Activity:
Extracted from My Activity
3x {*Search, Maps, YouTube*}
3x {*recent, 3-, 18-months old*}

Task:
"What interests do you think Google would apply?"

Labeling Task

YouTube Activity

May 10, 2022

▶ YouTube

Watched Hometown [Official Music Video]
French 79

4:58 PM

Skip this Activity ▾   Undo Skip

Seeing the activity above, what interests do you think Google would apply to it? *

Please type or select at least 3 interests.

[M] Music Videos ⊗   [D] Dance & Electronic Music ⊗   French 79   ⌫   +

F French Cuisine

F French Guiana

F French Overseas (DOM-TOM)

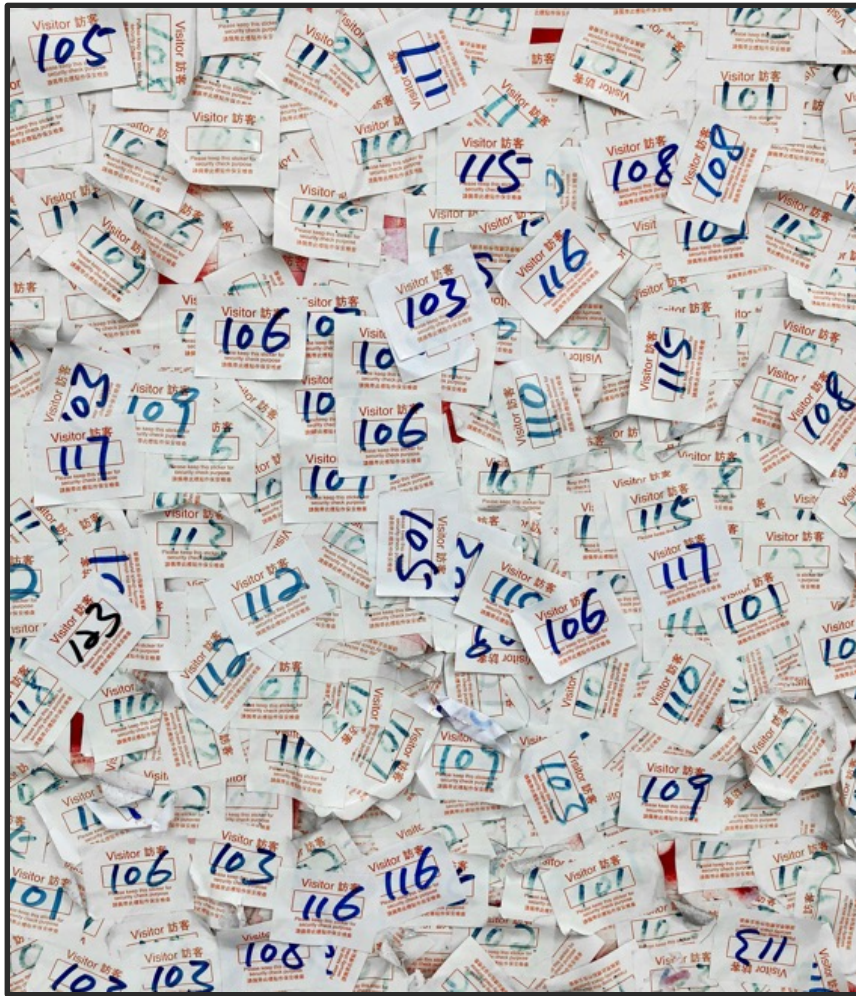3 results

RUB GW

Results

"... as I use them for everything, they know more about me than I know about myself." (P101)

People's Understanding of Inferencing

2615 labels:

27% were **on** Ad Settings page

73% were **not**

70% existing label

30% custom label

Participant

D Dance & Electronic Music ✕ == 🎵 Dance & Electronic Music

M Music Videos ✕   ✕ N/A

F French 79 ✕   ✕ N/A

Google

Too Generic
M Maps ✕
I Images ✕
Y YouTube ✕
K Knowledge ✕

Overly Specific
M Minecraft ✕
D Disney+ ✕
T Taylor Swift Red ✕
S Seattle Seahawks ✕

Banned
C Christianity ✕
P Pornography ✕
C COVID-19 ✕
B Brewing ✕

Context-Dep.
Maps
Directions to 6950 Ringkøbing, Dänemark
?
W Wedding Gift ✕

People's Understanding of Inferencing

2615 labels: Participant   Google

27% were **on** Ad Settings page | D Dance & Electronic Music ✕ | == | ♫ Dance & Electronic Music |

73% were **not**
    70% existing label | M Music Videos ✕ | | ✕ N/A |

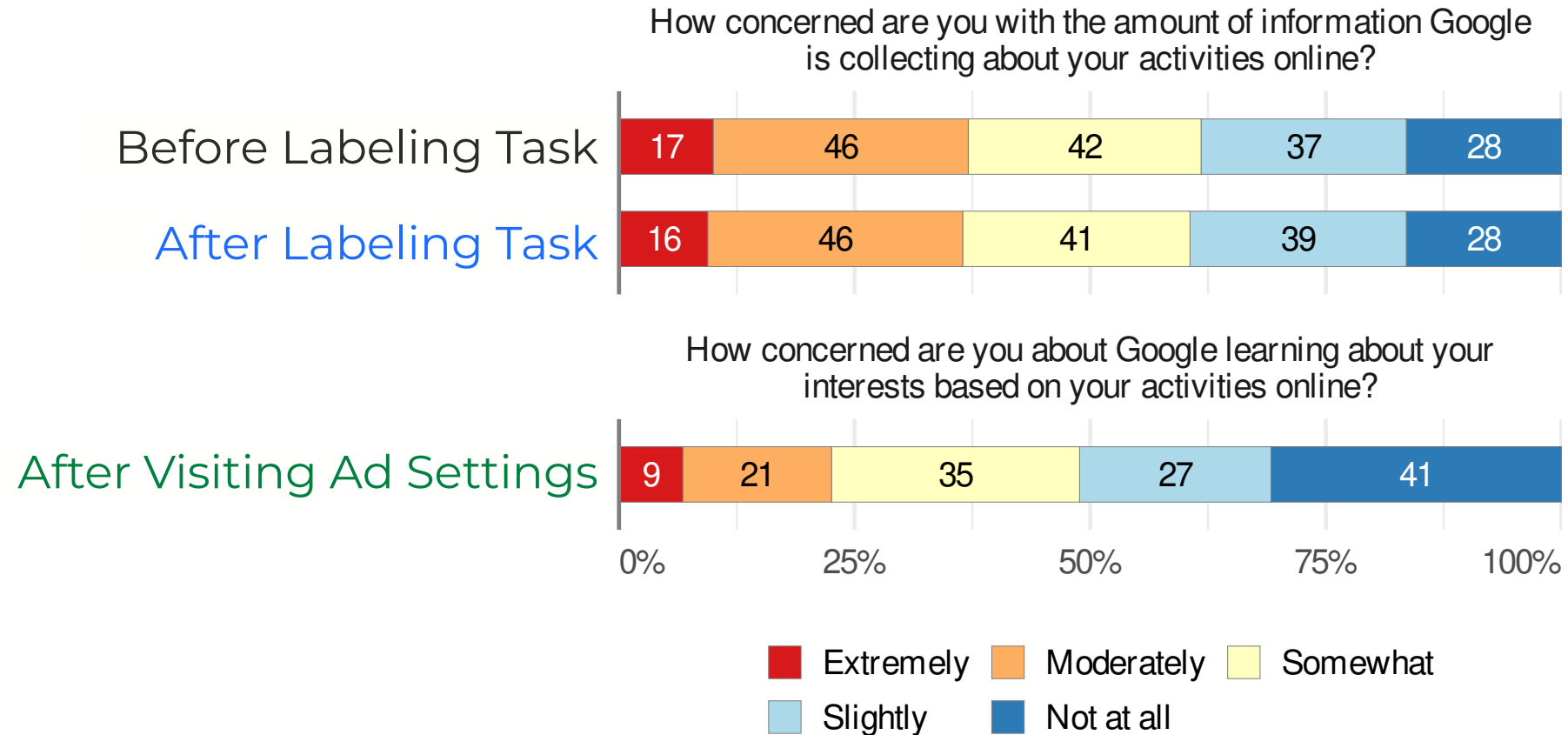   30% custom label | F French 79 ✕ | | ✕ N/A |

We find that...
... participants apply interest labels **differently** than Google.
Labels were often either **too generic** ("Images") or **overly specific** ("Sims 4").

19

Did the Labeling Task Work?

How concerned are you with the amount of information Google is collecting about your activities online?

Before Labeling Task	17	46	42	37	28
After Labeling Task	16	46	41	39	28

😇

How concerned are you about Google learning about your interests based on your activities online?

After Visiting Ad Settings	9	21	35	27	41

🤯

0%   25%   50%   75%   100%

■ Extremely ■ Moderately □ Somewhat
■ Slightly ■ Not at all

Did the Labeling Task Work?

How concerned are you with the amount of information Google is collecting about your activities online?

Before Labeling Task	17	46	42	37	28
After Labeling Task	16	46	41	39	28

😇

How concerned are you about Google learning about your interests based on your activities online?

After Visiting Ad Settings	9	21	35	27	41

🤯

0%   25%   50%   75%   100%

🟥 Extremely   🟧 Moderately   ⬜ Somewhat
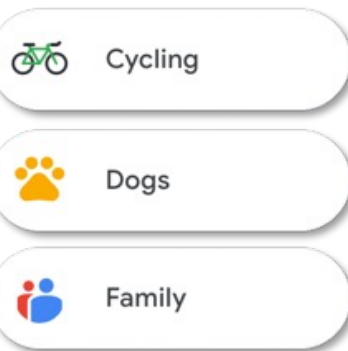🟦 Slightly   🟦 Not at all

In contrast to the labeling task…

… **viewing the labels** that Google applied **significantly decreased concern,** as labels were mostly perceived as **too generic** and **low-risk**.
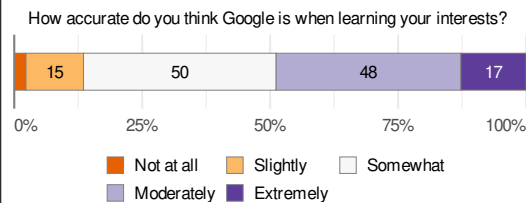
RUB   GW

The Generic Paradox

… by **focusing** on the **generic quality** of interests and not on the **volume** and **accuracy**, people **allow Google** to **continue collecting data** about their online activities.
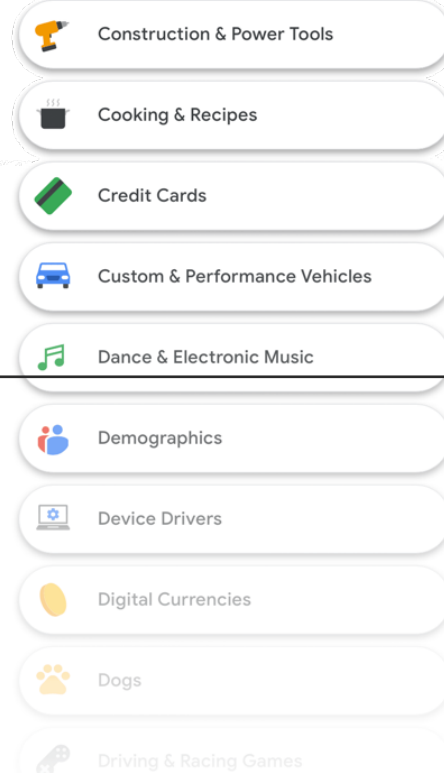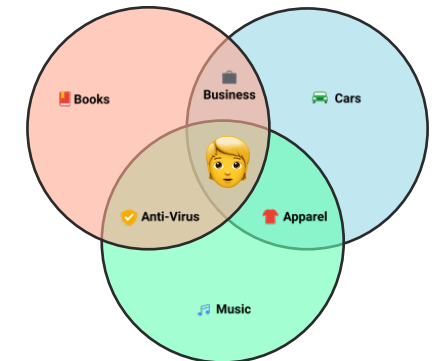
Generic

- Cycling
- Dogs
- Family

Accurate

How accurate do you think Google is when learning your interests?

| 15 | 50 | 48 | 17 |

0%   25%   50%   75%   100%

- Not at all
- Slightly
- Somewhat
- Moderately
- Extremely

> 180 Interests

- Construction & Power Tools
- Cooking & Recipes
- Credit Cards
- Custom & Performance Vehicles
- Dance & Electronic Music
- Demographics
- Device Drivers
- Digital Currencies
- Dogs
- Driving & Racing Games

Abuse Potential

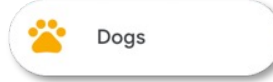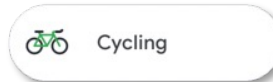Books, Business, Cars, Anti-Virus, Apparel, Music

Takeaways

Labeling Task

Participants apply interests **differently** than Google; Labels were **too generic** or **overly specific**.
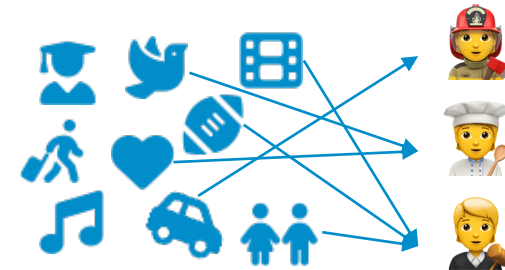

Ad Settings

Interests by Google are somewhat **accurate** and **perceived** as **too generic** and **low-risk**. Viewing them **decreased concerns**.
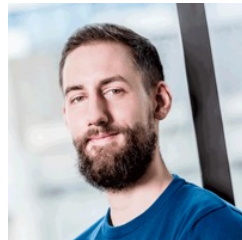

Future Work

Develop dashboards that better explain **how** the **data will be used**, and demonstrate **how** seemingly **low-risk labels** can be combined and **can lead to abuse**.

Florian M. Farke

David G. Balash

Maximilian Golla

Adam J. Aviv

RUB RUHR UNIVERSITÄT BOCHUM

UNIVERSITY OF RICHMOND

CISPA HELMHOLTZ CENTER FOR INFORMATION SECURITY

GW THE GEORGE WASHINGTON UNIVERSITY WASHINGTON, DC