ISD

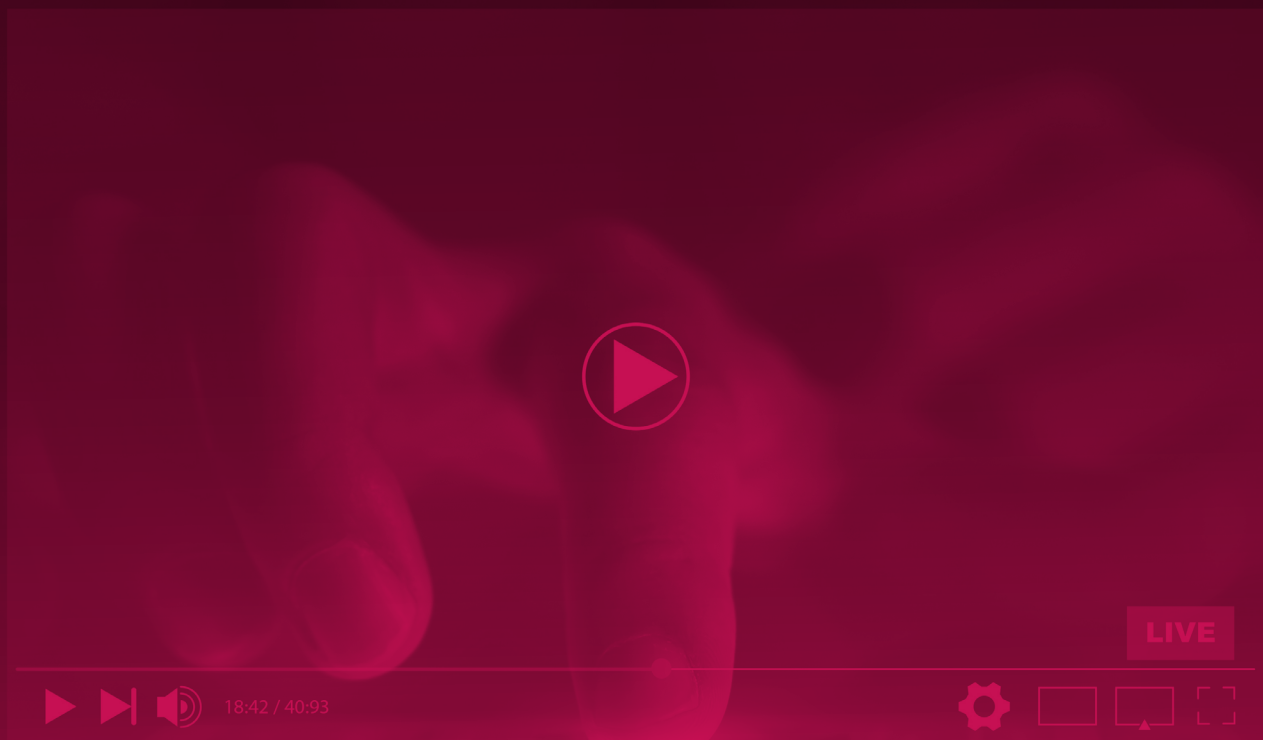# Pulling Back the Curtain

An Exploration of YouTube's Recommendation Algorithm

Aoife Gallagher, Lucy Cooper, Rhea Bhatnagar, and Cooper Gatewood

18:42 / 40:93

LIVE

About the authors

Aoife Gallagher is a Senior Analyst on ISD's Digital Analysis Unit, focusing on the intersection between far-right extremism, disinformation and conspiracy theories and using a mixture of data analysis, open source intelligence and investigative techniques to understand the online ecosystem where these ideas flourish and spread. Previously, Aoife was a journalist with the online news agency, Storyful. She is the author of the book Web of Lies – The Lure and Danger of Conspiracy Theories and co-author of the ISD reports The Genesis of a Conspiracy Theory and Ill Advice: A Case Study in Facebook's Failure to Tackle COVID-19 Disinformation.

Lucy Cooper is a Digital Research Analyst at ISD US. Lucy holds a BA in political science from Barnard College of Columbia University, where she focused on political behavior, voting and elections in the United States. She has previously worked on political campaigns in Michigan and as a research assistant.

Rhea Bhatnagar is a Digital Research Analyst at ISD US where she focuses on online extremism, hate and disinformation in the US and globally. Prior to joining ISD, Rhea worked for Media Matters for America as an Internet and Misinformation Researcher, where she concentrated in platform accountability and the spread of international online misinformation. Her research has previously been cited in The New York Times, The Washington Post and Columbia Journalism Review. Rhea holds an undergraduate degree in International Affairs with a concentration in Conflict Resolution from the George Washington University's Elliott School of International Affairs.

Cooper Gatewood was a Senior Manager within ISD's Digital Analysis Unit, focusing on quantitative research into the spread of hateful and polarising narratives online, and how they are leveraged by extremist actors. Cooper also developed monitoring and evaluation frameworks to measure the impact of many of ISD's intervention projects. Cooper contributed to ISD's research on disinformation campaigns, particularly those aimed to influence and disrupt election processes. Cooper holds a Masters of International Affairs from Columbia University and a Masters of International Security from Sciences Po and is fluent in Spanish and French, as well as speaking proficient Japanese.

ISD

Powering solutions
to extremism, hate
and disinformation

Amman I Berlin I London I Paris I Washington DC

www.isdglobal.org

Contents

Executive Summary

This is a four-part research project aimed at examining YouTube's recommendation algorithm from different angles to shed light on the technology that powers 70% of all video views on the platform. Each investigation sought to examine the kinds of content recommended to users who showed interest in certain subjects. Additionally, each investigation tested whether a particular variable would impact the content recommended to users. To do this, two accounts were created for each investigation.

Investigation One looked at whether the gender of an account would impact the content recommended to teenage users interested in gaming. For the second investigation, two male accounts were created with an interest in "male lifestyle gurus" — one a 13-year-old and the other a 30-year-old — to test whether age plays a role in the kind of content recommended.

The third and fourth investigations examined users with an interest in "Mommy vloggers" and Spanish-language news respectively.[1] Instead of testing account setting variables, these investigations looked at whether slight variations in the content watched by users would have an impact on the algorithm. The Mommy vlogger accounts varied their interest in news, with one watching a left-leaning news channel and the other a right-leaning one. For the Spanish-language news users, one account showed an interest in fringe conspiratorial content, while the other watched a mainstream lifestyle channel.

While this project did shed some light on the functionality of YouTube's algorithm, it also raised questions and concerns. Analysts found a lack of conclusive evidence to suggest that the age and gender of accounts impacted the kind of content recommended, but they did find evidence that young users were exposed to inappropriate and potentially harmful content as a result, including videos related to self-harm and suicide.

Additionally, this research shows that YouTube is showing its users large number of videos featuring content unrelated to their interests. Although on the surface this is not necessarily problematic, ISD's research shows that harmful content is often among these unrelated videos. This includes health misinformation and videos featuring Andrew Tate, the misogynistic social media influencer being investigated

in multiple jurisdictions for sexual crimes. Relatedly, the presence of content related to Christianity across the recommendations of all four investigations also raises questions about its consistent presence in YouTube's algorithm.

The following report outlines the trends found across all four investigations, as well as details on the methodology and platform policy recommendations. See each separate investigation for more detailed analysis.

Key Findings & Introduction

Key Findings

- YouTube is not properly safeguarding its young users. Accounts set up to imitate teenagers were recommended harmful content, including sexualized gaming videos, misogynistic content, and videos relating to self-harm and suicide.

- YouTube's search function and recommender feed surfaced health misinformation and other problematic content. Despite users not searching for it, users were exposed to anti-vaccine content and videos featuring Andrew Tate, who has been banned from the platform.

- Across all four investigations, religious content was recommended to users despite none of the accounts showing an interest in it. These videos were almost exclusively related to Christianity, raising questions as to why this happens.

- While there were differences seen in the recommendations of accounts with different account settings, there were few conclusive indications that these were related to the age or gender of the accounts. These variations signify that each user journey is unique and that unknown factors decide what kind of content is recommended.

Introduction

As social media has evolved over the last 15–20 years, platforms have begun to rely more on machine learning algorithms to serve content to their users. These systems are increasingly being criticized by policymakers and activists as evidence mounts that platforms are recommending harmful and problematic content to their users.[2]

According to YouTube, the recommendation algorithm on the platform drives 70% of all video views.[3] This means users are, in many cases, being served content YouTube believes they wish to watch, rather than searching for and manually selecting content themselves.

YouTube states that a variety of factors influence what its users will see in the recommended feed, including search and watch history, the videos a user likes, the channels they subscribe to, and the playlists they create. Other information, such as the country you live in and the time of day, can also influence recommendations.[4]

Past research into YouTube's algorithm has found evidence of the platform "facilitating problematic content pathways", pushing users into "mild ideological echo chambers", and serving up content containing misinformation, violence, and hate speech.[5] For example, research conducted by ISD in 2022 on YouTube Shorts — the platform's TikTok-style vertical video feed — found that the feature was promoting anti-feminist and misogynistic content to boys and young men.[6]

These investigations and others like it are often caveated by the statement that without meaningful access to data to effectively study YouTube's recommendation systems, it can be difficult to draw absolute conclusions from the research.

Project Overview

This ISD investigation aimed to pull the curtain back on the YouTube recommendation algorithm and attempted to answer some questions about how it serves content. Four investigations were conducted that examined the algorithmic recommendations of users with different interests — gaming, male lifestyle gurus, Mommy vloggers, and Spanish-language news.

Two accounts were created for each investigation to examine different aspects of the recommendation system. Investigations One and Two looked at variations in account settings — namely age and gender. Investigations Three and Four looked at slight variations in the kind of content watched. See Tables 1 and 2 for more details.

Table 1: The interests and variables tested in Investigations One and Two.

	Account setting variation investigations	
	Investigation 1	Investigation 2
Persona interest	Gaming	Male lifestyle gurus
Variable tested	Gender (one male, one female)	Age (one 13-year-old, one 30-year-old)

Table 2: The interests and variables tested in Investigations Three and Four.

	Content variation investigations	
	Investigation 3	Investigation 4
Persona interest	Mommy vloggers	Spanish-language news
Secondary interest (variable tested)	News (right-leaning or left-leaning)	Vloggers (fringe or mainstream)

Building the personas

According to research, users spend an average of 19 minutes per day on YouTube.[7] This project aimed to examine what the recommendations on the platform present to users after five days of watching YouTube (with watch time over five days averaging 19 minutes per day.

For the accounts to act as similarly to regular users as possible, they engaged in different actions, including searching for channels, searching for specific keywords and phrases, subscribing to channels, and watching videos. A period of scoping research was carried out to identify popular channels, keywords, and trends associated with each persona.

See individual reports for details on how personas were created for each investigation, the content the accounts watched, and the actions taken.

Analysis

After the fifth day of persona building, a browser automation tool was used to record the recommendations displayed on the homepage of each account every six hours for one month. The recommendations for each account across the four personas were then coded (see individual investigations for different coding guidelines used) and analyzed according to the following criteria and principles:

- The number of videos recommended to each account;

- The distribution of video recommendations to each account (i.e. how many videos were recommended multiple times to each account);

- All videos recommended more than once were coded according to the coding guidelines outlined in each investigation;

- Unique videos recommended to each account — that is, videos that were recommended to one account but not the other — were coded using the same guidelines;

- The ads recommended to each account were categorized and analyzed across each account.

Findings

YouTube is not adequately safeguarding young users from harmful content

Both Investigations One (gamers) and Two (male lifestyle gurus) involved analyzing the content recommended to teenage users of YouTube, and both investigations surfaced findings that indicate gaps in YouTube's efforts to safeguard young users.

During the set-up of YouTube accounts (detailed below), it was noted that parent/guardian approval was not required for creating accounts for both 13-year-old and 14-year-old-users. Google states that the minimum age for creating a Google account in the US is 13. In many other countries, particularly in Europe, the minimum age is between 14 and 16.[8] ISD could not find any indication that there are safeguards in place to stop young users simply lying about their age to bypass these restrictions.

ISD also found multiple instances of harmful content being recommended to young users through YouTube's algorithm. In Investigation One's analysis of teenage users with an interest in gaming, both accounts were recommended gaming videos containing sexually explicit content and others with themes of self-harm and suicide; only one of these videos contained a content warning. Videos related to guns and ballistics and glorifying the use of weapons were also recommended to these accounts. YouTube did not place content warnings or age restrictions on any of these videos.

In Investigation Two's examination of content recommended to users interested in male lifestyle gurus, ISD found few indications that the content recommended to a 13-year-old user differed from that recommended to a 30-year-old based on the age of the account. This meant that age-inappropriate content, including videos that sexualized women and promoted misogynistic ideas were shown to a teenage user. Other videos recommended included anti-trans content and videos and channels related to wider "culture war" issues.

Additionally, videos of Andrew Tate were also recommended to both the child and adult accounts despite neither account showing an interest in him and Tate being banned from the platform. Again, YouTube did not place any age restrictions or content warnings on these videos.

In both investigations, accounts were recommended clips from TV shows that are not exceptionally explicit, but are targeted at adult audiences, including Family Guy and South Park.

YouTube states that it "may place an age-restriction" on content that "doesn't violate our Community Guidelines, but … may be incompatible with YouTube's Terms of Service or not appropriate for viewers under 18." The examples listed by the platform of content that would fall into this category include "a video containing adults participating in dangerous activities that minors could easily imitate, such as handling explosives or challenges that cause bodily injury" and videos that contain nudity or sexually suggestive content including those "where the subject is in a pose that is intended to sexually arouse the viewer" or "where the subject is in clothing that is considered unacceptable in public contexts, such as lingerie". Other examples include videos "with heavy profanity in the title, thumbnail or associated metadata".[9] YouTube also states that the examples listed are not exclusive.

YouTube search results surface channels with a history of sharing health misinformation

In Investigation Three, where users displayed an interest in Mommy vloggers, a search on YouTube for the term "natural immunity in children" surfaced a video from Dr. John Campbell at the top of the results.

Dr. John Campbell has a doctorate in nursing but is not a physician and has a history of amplifying health misinformation and anti-vaccine rhetoric.[10] His videos have been fact checked at least a dozen times according to Google's Fact Check Explorer.[11]

Because the two accounts watched the Dr. John Campbell video that appeared at the top of the search results, both were then recommended more of his content, some of which included misinformation about COVID-19 and vaccines.

The high prominence given to Dr. John Campbell in YouTube's search results and the number of his videos subsequently recommended to the accounts highlight how easy it can be for users to be led into misinformation rabbit holes on the platform. This also shows clear gaps in YouTube's efforts to stem the flow of such content.

Videos with religious themes are recommended to accounts that have shown no previous interest in such content

An interesting finding across all four investigations was the prominence of videos with religious themes, primarily Christianity-related content, that was recommended to the accounts. None of the accounts analyzed in these investigations watched religious content during the persona-building stage or displayed an interest in religion in any other manner.

Despite this, religious-themed videos appeared in the recommendations of all the accounts, albeit in different amounts. These videos included religious sermons, full-length livestreams of religious services, and meditation videos with religious overtones.

Table 3 shows the number and percentage of religious-themed videos that were found during the analysis of unique videos in each investigation (videos recommended to one account but not the other). While the proportion of such videos was relatively low in Investigations One and Two, the third and fourth investigations revealed a higher percentage. This is possibly reflecting that users interested in Mommy vloggers and Spanish-language news tend to be more interested in religion; however, further research is needed to understand this.

The ubiquity of these videos across investigations, as well as the fact that almost all the videos were related to Christianity, raises questions as to why YouTube recommends such content to users and whether this is a feature of the platform's recommendation system.

What we learned about the algorithm

Across all four investigations, it was found that the content most frequently recommended to accounts was related in some way to the user's interests in either gaming, male lifestyle gurus, Mommy vloggers or Spanish-language news. This is unsurprising given that the accounts were built to have an interest in these areas.

There were some noteworthy differences in both the number of videos recommended to each account and the number of ads seen across the investigations even though data was collected for all accounts for exactly one month (see Figures 1 and 2).

These differences were greater in Investigations Three and Four, where the content watched by the accounts varied. This suggests that the number of videos and ads recommended to users may differ based on slight variations in the content watched and that the type of content watched also impacts the number of ads seen by

Table 3: Results of the unique video analysis in each investigation, showing the number and percentage of religious-themed videos recommended to each account.

Investigation	Account	#	%
1	Male	8	2.6
	Female	3	1
2	Child	5	2
	Adult	1	0.3
3	Right	14	5.5
	Left	14	3.9
4	Fringe	14	3.6
	Mainstream	34	7.2

Figure 1: The number of videos recommended to each account across all four investigations.

Number of videos recommended

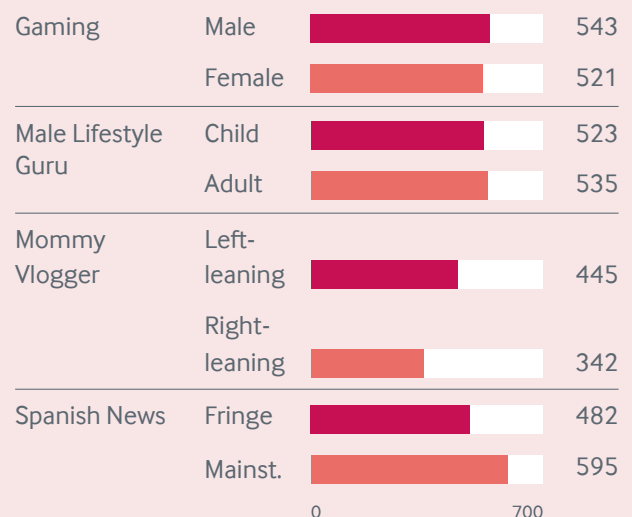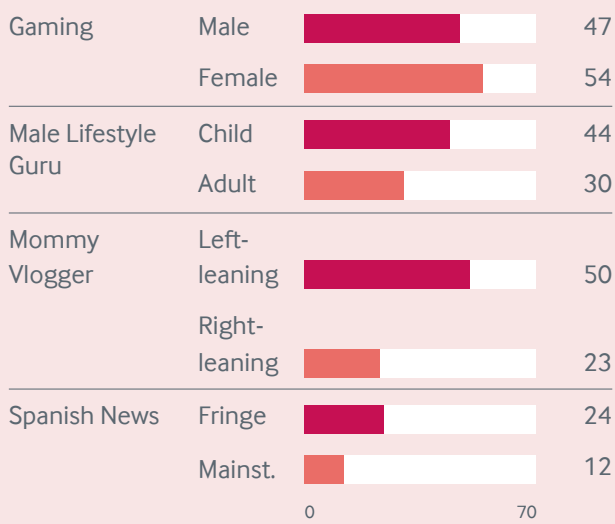|		
Gaming	Male	543
	Female	521
Male Lifestyle Guru	Child	523
	Adult	535
Mommy Vlogger	Left-leaning	445
	Right-leaning	342
Spanish News	Fringe	482
	Mainst.	595

0   700

Figure 2: The number of ads seen by each account across all four investigations.

Number of ads seen

Gaming	Male		47
	Female		54
Male Lifestyle Guru	Child		44
	Adult		30
Mommy Vlogger	Left-leaning		50
	Right-leaning		23
Spanish News	Fringe		24
	Mainst.		12

0   70

Table 4: Results of the unique video analysis in each investigation, showing the number and percentage of religious-themed videos recommended to each account.

Investigation	Account	% unique	% unique and unrelated
1	Male	56	52
	Female	55	55
2	Child	55	46
	Adult	57	44
3	Right	74	55
	Left	80	64
4	Fringe	74	58
	Mainstream	79	45

users. Overall, these findings show that user journeys are unique and based on factors that are as of yet unknown. It is therefore difficult to come to definitive conclusions without greater access to algorithmic data.

What is perhaps more interesting is the analysis of unique videos in each investigation and the proportion of these unique videos that were unrelated to the accounts' interests (see table 4).

For Investigations One and Two, an average of 56% of the videos recommended to each account were unique – in that they were not recommended to the other account. For Investigations Three and Four, this average was 77%. This difference may be explained by the variations in the content watched by the accounts in Investigations Three and Four, signifying that slight differences in the kind of content watched can have an outsized impact on the videos recommended.

The unique videos that contained content not related to the accounts' interests also tell an interesting story. Across all four investigations, the average percentage of these videos was 52%. This is likely an effort by the algorithm to detect other interests of the user, but such a high percentage could also be problematic and has the potential to expose users to harmful content without them searching for it directly. For example, it was within this set of unique and unrelated videos that the accounts in Investigations One and Two were recommended adult entertainment content and anti-trans content respectively.

Platform Policy Recommendations

- YouTube should increase its moderation of gaming videos and provide updated flags to moderators about harmful themes appearing in gaming videos. Given that YouTube already lists several categories of content we found recommended to teen accounts under categories that could potentially violate its child safety policy, the platform should commit more resources to moderating videos that have been identified as popular with teen and child accounts, proactively identify themes within those videos that may be harmful to minors, and tag those videos with content warnings or enact an age restriction to view.

- YouTube should consistently enforce its medical misinformation policy, which states that videos found violating the policy will be removed and that accounts found to be violating the policy frequently may be removed. YouTube should also consider updating its medical misinformation policy to contain more concrete terms around account removal and deletion. While the current policy states that accounts with repeated violations of the policy "may be" terminated, creating a definitive upper bound of violations could make enforcement of the policy easier and more consistent, as well as prove a stronger disincentive to accounts that spread medical misinformation. Additionally, channels with a history of sharing misinformation should not receive prominence in search results or be boosted within the algorithm. This could be added to the list of actions taken for accounts that receive strikes.

- Without greater transparency and data access for researchers into YouTube's algorithm, the ability of external researchers to study, identify, and articulate problematic or violative content being fed to users will always be limited. To get a more comprehensive understanding of YouTube's recommendations, including type, frequency, and the reason for recommendation, either YouTube will need to proactively grant that access to researchers or, more likely, legislation will have to be passed to mandate greater transparency and data access.
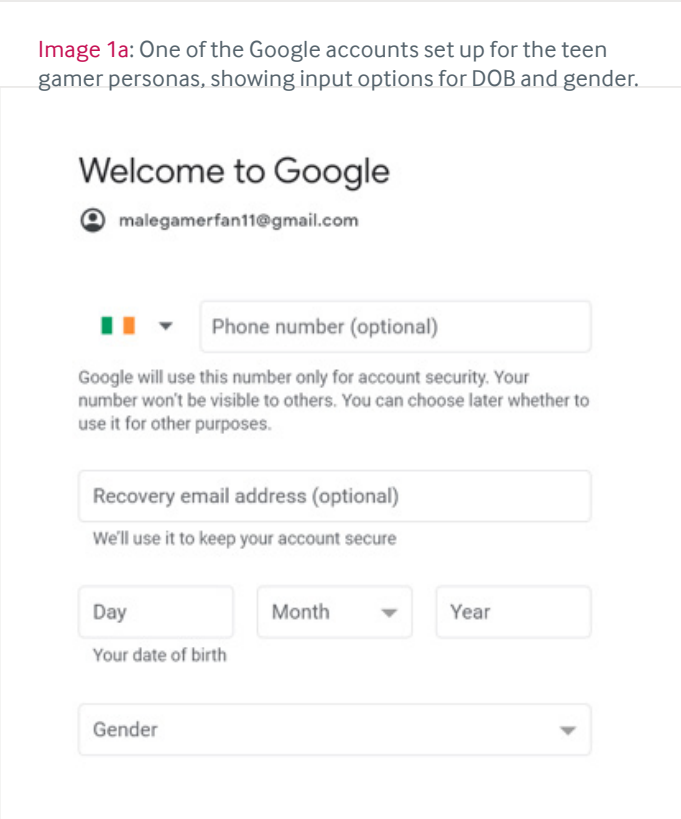
Appendix

Creation of YouTube accounts

YouTube accounts are connected to Google accounts, therefore clean Google accounts were created for the purposes of this project. The details in Table 5 below list the locations (via VPN), ages and genders that were chosen for each of the accounts that were created. The VPN location was kept consistent during the persona-building stage (detailed below). Image 1a shows the input options for gender and date of birth when creating a Google account.

It is noteworthy that Google accounts (and therefore also YouTube accounts) created for teenagers did not require the approval of a parent or guardian before being set up.

Table 5: Details of age, gender, and account location of each YouTube account.

	Chosen location	Chosen D.O.B	Chosen Gender
Teen gamer	CA	06/01/2008 (14 y/o)	Male
	CA	06/01/2008 (14 y/o)	Female
Male lifestyle guru	IL	06/01/2009 (13 y/o)	Male
	IL	06/01/1992 (30 y/o)	Male
Mommy vlogger	DE	06/01/1982 (40 y/o)	Female
	DE	06/01/1982 (40 y/o)	Female
Spanish-language news	FL	06/14/1990 (35 y/o)	Male
	FL	06/14/1990 (35 y/o)	Male

Image 1a: One of the Google accounts set up for the teen gamer personas, showing input options for DOB and gender.

Welcome to Google

malegamerfan11@gmail.com

Phone number (optional)

Google will use this number only for account security. Your number won't be visible to others. You can choose later whether to use it for other purposes.

Recovery email address (optional)

We'll use it to keep your account secure

Day   Month   Year

Your date of birth

Gender

Endnotes

1   Mommy vloggers refers to YouTubers who share their personal experiences of motherhood through their videos.

2   David, E. (2023). Congress if trying to stop discriminatory algorithms again. The Verge, 13 July.
      Available at: https://www.theverge.com/2023/7/13/23792992/congress-discriminatory-algorithms-ban-law-transparency;
      Lynch, D. (2023). Peers demand curbs on harmful algorithms introduced to online safety Bill. The Independent, 10 July.
      Available at: https://www.independent.co.uk/news/uk/politics/bill-andrew-tate-government-david-lynch-romania-b2372770.html;
      O'Brien, C. (2023). Social Media algorithms causing 'profound damage' to young people — Stephen Donnelly. The Irish Times, 15 May.
      Available at: https://www.irishtimes.com/ireland/education/2023/05/15/social-media-algorithms-causing-profound-dam-age-to-young-people-stephen-donnelly/.

3   Rodriguez, A. (2018). YouTube's recommendations drive 70% of what we watch. Quartz, 13 November.
      Available at: https://qz.com/1178125/youtubes-recommendations-drive-70-of-what-we-watch.

4   Goodrow, C. (2021). On YouTube's recommendation system. YouTube Official Blog, 15 September.
      Available at: https://blog.youtube/inside-youtube/on-youtubes-recommendation-system/.

5   Yesilada, M. and Lewandowsky, S. (2022). Systematic review: YouTube recommendations and problematic content. Internet Policy Review,
      11(1). Available at: https://policyreview.info/articles/analysis/systematic-review-youtube-recommendations-and-problematic-content;
      Brown, A. et al (2022). Echo Chambers, Rabbit Holes, and Algorithmic Bias: How YouTube Recommends Content to Real Users.
      Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4114905; Mozilla (2021). Mozilla Investigation: YouTube Algorithm
      Recommends Videos that Violate the Platform's Very Own Policies. 7 July. Available at: https://foundation.mozilla.org/en/blog/mozilla-in-vestigation-youtube-algorithm-recommends-videos-that-violate-the-platforms-very-own-policies/.

6   Thomas, E. and Balint, K. (2022). Algorithms as a Weapon Against Women: How YouTube Lures Boys and Young Men into the 'Manosphere'.
      ISD, 27 April. Available at: https://www.isdglobal.org/isd-publications/algorithms-as-a-weapon-against-women-how-youtube-lures-boys-and-young-men-into-the-manosphere/.

7   Seitz, L. (2024). Average Daily Time Spent on Social Media (Latest 2024 Data). BroadbandSearch, 4 January.
      Available at: https://www.broadbandsearch.net/blog/average-daily-time-on-social-media.

8   Age requirements on Google Accounts. Google. Available at: https://support.google.com/accounts/answer/1350409#zippy=.

9   Age-restricted content. Google.
      Available at: https://support.google.com/youtube/answer/2802167?visit_id=638447224176354419-3407242767&rd=1.

10   Person: John Campbell. FactCheck.org. Available at: https://www.factcheck.org/person/john-campbell/; Benedictus, L. (2022).
      YouTuber misinterprets Covid-19 vaccine evidence on children from Singapore. Full Fact, 12 August.
      Available at: https://fullfact.org/health/john-campbell-youtube-singapore-children/.

11   Fact Check Tools: "John Campbell". Google.
      Available at: https://toolbox.google.com/factcheck/explorer/search/john%20campbell;hl=en.

ISD

Powering solutions
to extremism, hate
and disinformation

Amman l Berlin l London l Paris l Washington DC

www.isdglobal.org