🕐 This article was published more than **1 year ago**

(Sean Loose/For The Washington Post)

WE THE USERS

I gave Instagram photos of my baby. Instagram returned fear.

Algorithms prey on you. What if you could reset them?

By [Geoffrey A. Fowler](#)

May 12, 2022 at 6:00 a.m. EDT

W hen my son was born last year, friends from all over wanted to share in my joy. So I decided to post a photo of him every day on Instagram.

Within weeks, Instagram began showing images of babies with severe and uncommon health conditions, preying on my new-parent vulnerability to the suffering of children. My baby album was becoming a nightmare machine.

This was not a bug, I have learned. This is how the software driving Instagram, Facebook, TikTok, YouTube and lots of other apps has been designed to work. Their algorithms optimize for eliciting a reaction from us, ignoring the fact that often the shortest path to a click is fear, anger or sadness.

For all its wonder and convenience, technology too often fails us. Lately, I've been exploring ideas about how we can make it better. High on my list of demands: We the users need transparency about how algorithms work — and the ability to press reset when they're not serving us.

I learned this firsthand by going on a hunt to unravel how my baby's Instagram account got taken over by fear.

More than a billion people spend time on Instagram in part because they enjoy it. I made my son a private Instagram account, and posted nothing but photos and videos of him smiling and snuggling. I followed the accounts of a handful of other babies from friends also longing to connect when covid-19 kept us apart.

But there was a darker dynamic at work, too. On the app's home screen and other tabs, Instagram mixes photos from my baby friends with suggested posts from strangers. At first, these algorithmically generated recommendations were neutral, such as recipes. After a few weeks, something caught my attention: Instagram was consistently recommending posts of babies with cleft palates, a birth defect.

Soon after came suggested posts of children with severe blisters on their lips. Then came children attached to tubes in hospital beds. In my main feed and the app's Explore and Reels tabs, Instagram was building a crescendo of shock: There were babies missing limbs, babies with bulging veins, babies with too-small heads, babies with too-big heads, even hirsute babies. Lots of the images were shared not by parents, but by spammers posting nonsense captions and unrelated images.

On Instagram's Shopping tab, things were also getting dark: T-shirts with crude dad jokes gave way to anti-vaccination propaganda, then even sexually explicit toys.

When I open Instagram today, more than 1 in 10 of the images I see just aren't appropriate for my baby photo album.

I shared dozens of examples of these posts with Instagram, which is owned by Facebook's parent, Meta. The company took down some of the shopping ads for violating its policy against adult products. But as for the suggested posts involving babies, spokeswoman Stephanie Otway says the company doesn't think there's anything un-recommendable about them. "Parents use Instagram to get advice, share their experiences, and seek support from other parents, including when their children have special needs," she says.

Of course parents can and should share photos and videos of their children, including when they have blisters or are in the hospital, to build community. But of all the millions of images across the app, these are the ones Instagram chose to show my son's account — and I have no way of knowing why.

> We the users need transparency about how algorithms work — and

What I question is how Instagram decided to show me these specific images, and at this volume, when I have no connection to these families.

the ability to press reset when they're not serving us.

Other new parents on Instagram tell me they also feel they're being recommended posts that prey on our specific insecurities, from breastfeeding to vaccination. "I found Instagram to be particularly devastating to my already fragile mental state in the postpartum period," says Nicole Gill, the co-founder of Accountable Tech, a progressive tech advocacy group. "Getting suggested posts on 'how to lose baby weight in 6 weeks,' for example, almost immediately after having my daughter was not pleasant."

Instagram would only describe in vague terms how its systems work, and wouldn't explain why it recommended this specific category of baby content.

So I called up an expert who would explain: Frances Haugen, the most prominent Facebook whistleblower.

Last fall, Haugen, a former Facebook product manager, exposed internal discussions about how the company's algorithms work, and its own research into the toxic outcomes. Among the most shocking revelations was the impact on teenagers: 32 percent of teen girls have told Facebook that when they felt bad about their bodies, Instagram made them feel worse.

Algorithms aren't just preying on teenagers, Haugen told me. Chances are, your feeds have also dragged you into rabbit holes you didn't ask for, but also can't avert your eyes from. Maybe you've experienced it in your Netflix queue, your Google search results or the recommended videos on YouTube.

Unraveling what happened to my son's Instagram account can explain how it happens — and offer some good ideas for how to stop it.

How they drag you down a rabbit hole

When we sat down together, I showed Haugen the recommendations in my son's Instagram account. "I'm so sorry that you keep getting exposed to these kinds of disturbing images," she says. "We're kind of on a runaway loop led by the algorithm right now."

To explain what's happening, she says, we have to start with what motivates Instagram and Facebook. Their business is based on showing you ads, so they want as much of your attention as possible.

Once upon a time, Instagram's main feed could actually come to an end, saying "you're all caught up" after you'd seen everything shared by your friends. But over time, the company decided your friends alone aren't enough to keep you opening its apps. So in 2020, Instagram started adding in algorithmically selected content you didn't request to keep you around longer.

So how does it decide what to show you? The algorithms used by Instagram and Facebook look for "signals." Some are obvious: Liking a post, following an account, or leaving a comment on a post are all signals.

In my case, I didn't do any of that with Instagram's suggested posts. But Haugen explained you don't have to "like" a darn thing for Instagram to pick up signals, because it's monitoring every single thing you do in the app.

"The reality of being a new dad is that you are more vulnerable to the suffering of children," Haugen says. "And I am sure when you run into one of the shocking photos, you're not intending to spend time on that photo, but you pause. And the algorithm takes note of that longer duration."

It's called "dwell time." Otway, the Meta spokeswoman, confirmed even the speed of your scroll is a signal that feeds Instagram's algorithm. So are a few other things Haugen said I likely did out of shock when I first saw these posts, such as tapping into an image to take a closer look. In a blog post last year, Instagram chief Adam Mosseri said the app is on the hunt for thousands of signals.

Instagram's judgments are, for the most part, invisible to us. If you're a power user, you can get a few more clues by requesting to download all your Instagram data. Buried in the files is "your topics," a list of everything the algorithm thinks you're interested in, which is used to create recommendations.

When I did that, I saw Instagram had assigned my son's account some 327 interests. Those included "disability" and "fear."

That's right, fear. I gave Instagram photos of my baby, and Instagram returned fear.

Said Otway, the Meta spokeswoman: "Our recommendations allow people in this community to find one another, but they can always let us know in the app if they're not interested in something recommended to them."

She's half right. You can't edit that list of "your topics" — but you can give feedback on an individual recommended post, if you know where to look.

Reporting this column, I learned Instagram offers this one lever of control over its algorithm: When you see a suggested post (or an ad), in the upper right corner there are three dots. Tap on them, and up pop a number of options, including a button at the bottom labeled "Not interested."

Humans out of the loop

It's not that Instagram and Facebook want to lead us to dark places, Haugen told me. But amplifying extreme content is one of the consequences of training algorithms to focus on what it calls "engagement," or content that leads people to interact.

According to the documents Haugen leaked, changes to Facebook's algorithms in 2018 and 2019 — to encourage what it called "meaningful social interactions" between users — had the consequence of promoting posts that sparked arguments and division.

Extreme content can also become a gateway to misinformation about vaccines, scams, or even sharing illicit images and information.

On my son's account, I witnessed another unintended consequence: what Haugen calls "engagement hackers." They're a kind of spammer who has learned how to hijack Instagram's logic, which encourages them to post shocking images to elicit reactions from viewers and thus build their credibility with the algorithm.

Several of the accounts behind the images Instagram recommended to my son's appear not to be parents of the children featured in the images. One image I've seen repeatedly, of a baby with what appear to be severe lip blisters, was shared by accounts called kids_past (with 117,000 followers) and another called cutes.babiesz (with 32,000 followers). The captions on the photos don't make sense with image, and don't appear to be related to the other children featured on the account. Both also suggest in their biographies that they're available for paid promotions. Neither account replied to messages asking where it had gotten the blister image.

Instagram doesn't completely throw caution to the wind. It has community standards for content, including guidelines on what kinds of themes can be included in posts that its algorithms recommend. It says content that's either "clickbait" or "engagement bait" is not allowed. In April the company announced a new effort to down-rank content that is not "original."

Haugen says Facebook doesn't have leadership that can ask hard questions about its impact — and accept hard answers. "When you acknowledge power, you also then acknowledge responsibility. And Facebook doesn't want to allocate any more time to features that don't cause it to grow."

How to make algorithms accountable

So how can we the users take back power over algorithms? From researchers and lawmakers alike, there's a growing collection of good ideas.

Instagram declined to let me speak with Mosseri for this column. Let's hope he's open to feedback.

Here's a start: Let us just turn off algorithms. In March, Instagram announced it would bring back a version of its main feed that sorts posts in reverse chronological order. That's good. But to completely shut off Instagram's recommended posts from accounts you don't follow — and make at least your main feed a friends-only experience — you have to select the Favorites-only view, and put all your friends in that category.

An even better idea: Give us an algorithmic reset button. I understand many people really enjoy social media recommendations, especially on TikTok. So give us the power to clear what the algorithm thinks about us without deleting the whole account and losing our friends, just like you can clear your history and cookies in a Web browser.

To give users more control, apps could also stop using unconscious actions — like dwell time while you're doomscrolling — as signals to feed recommendations. Instead, they should focus on the signals where we explicitly say we're interested, such as pressing like or following an account.

Apps also need to provide us better ways to give negative feedback on their algorithmic choices. Right now it's too hard to tell Instagram or Facebook you don't want something. It could move a "no thank you" button out from behind the menu screen, and to right next to the Like button.

Instagram is starting down this path. It tells me it is at the early stages of exploring a control that would allow people to select keywords to filter from their recommendations. To use their example, if you asked that the word "bread" be removed from your recommendations, Instagram wouldn't show posts containing the word "bread."

I'm also intrigued by a bolder idea: Let us choose between competing algorithms to order the information on our feeds. Algorithms can be programmed to show or bury content. Some people might want to see Donald Trump, while others might want feeds that are completely politics-free. It could work kind of like the app store on your phone. Different algorithm developers could compete to organize your Instagram, Facebook or Twitter feed, and you settle on the one you like the best. Or maybe you switch from time to time, depending on your mood.

These are all product fixes, but bigger solutions have to address another problem: We actually know very little about how these algorithms work. Right now, researchers and governments — not to mention we the users — can't see inside their black box for ourselves.

"We the users deserve transparency," says Haugen. "We deserve to have the same level of nutritional labeling for our informational products as we have for our nutritional products. We deserve to see what goes into the algorithms. We deserve to see what the consequences of those things are. And right now, we're forced to just trust Facebook."

It's more than an academic issue. "Platforms get to experiment on their users all the time without letting them know experiments are going on," says Laura Edelson, a researcher at New York University whose Facebook account was cut off by the company for studying political advertisements and misinformation. "Consumers deserve notice."

In the U.S., at least five bills have been introduced in Congress that focus on responsibility for algorithms. One of them is the bipartisan Platform Accountability and Transparency Act (PATA), which would force companies to open up their algorithms by turning over information about how they work — and their consequences — to researchers and the public.

"We agree people should have control over what they see on our apps and we'll continue working on new ways to make them more transparent, while also supporting regulation that sets clear standards for our industry in this area," said Otway, the spokeswoman for Instagram.

Now it's time to hold them to it.