Siara (00:00)
We are all rats in a BF Skinner experiment acting nonsensically

This problem that we have with smartphones is not happenstance. This problem was intentional and most importantly, this problem is profitable.

They've figured out how to weaponize human psychology to get us to do things subconsciously.

they track every single second that you move through these apps. It is that exact of a science.

consumers should know what tactics are being used against them. They don't want us to talk about them, but it's a free country, so I'm going to.

Hey,

Happy August.

Question for you. Have you ever opened an app, done your thing on the app, closed the app, and then 30 seconds later gone right back to the app?

I do this sometimes, and I'm always like, what did I expect to change? Like, why am I back here? You are actually being trained to do that. We are all rats in a BF Skinner experiment acting nonsensically is where I'm at.

And I don't say that to make you feel bad. I say that to tell you that this urge is engineered.

This problem that we have with smartphones is not happenstance. This problem was intentional and most importantly, this problem is profitable.

Engineers very successfully use neuroscience and product design to keep you on your phone and on these apps. I think I've always noticed this in the background, but it became very clear to me after reading this book I found a few years ago.

I was looking for professional advice on how to have a better relationship with my phone.

I was getting distracted a lot and I was looking for answers. So I found this book, Indistractable, How to Control Your Attention and Choose Your Life. Very good book.

I learned some good stuff. There are some great concepts in here. I was actually super into it. You can see the obsessive notes I was taking. I was like locked in. But three years later, the most interesting thing to me about this book is the precursor to the book.

The exact same author, Nir Eyal published a book five years before this one called Hooked, How to Build Habit-Forming Products. I want to read you a really quick excerpt from the intro of Indistractable that actually talks about Hooked.

It's a cookbook of sorts. The book contains a recipe for human behavior, your behavior. These tech companies know that in order to make money, they need to keep us coming back. Their business model depends on it. I know this because I've spent the past decade researching the hidden psychology that some of the most successful companies in the world use to make their products so captivating. He notes that companies like Facebook, Google, PayPal, Slack, they keep this book in their offices and recommend it to their employees.

Okay, so if there's a cookbook, that means there's a recipe. So when he says habit forming or recipe for human behavior or keep us coming back, I hear them all as euphemisms for how to manufacture an addiction.

Like habit forming doesn't have to be an inherently bad thing, of course. There are apps like Insight Timer, for instance, that are trying to get you to form the habit of meditating, which I do think is a net positive for someone's life. But when that habit stops being about utility and improving someone's life, what now? We can see that these apps have gone far beyond utility with social media, gambling apps, games, and other time wasting, energy sucking, and in some cases life-ruining apps.

They've figured out how to weaponize human psychology to get us to do things subconsciously. And I think in my opinion, that's sort of messing with human will.

Understanding these tactics has helped me tremendously. I can now recognize when I'm being teased into a habit loop that I know is not going to help me in the long run. I'm also a lot more critical of the actions that I take on an app now, especially if it's a new app that I haven't formed any habits for yet. some apps have successfully formed my habits, unfortunately. But now that I can see what's happening, I can fight them a lot more strategically.

And I want everyone else to have that same knowledge because these tactics aren't often talked about in mainstream circles.

They're usually talked about within the tech industry and usually with the aim of getting users to use more of their technology, which is fine. Like business is business, but consumers should know what tactics are being used against them. They don't want us to talk about them, but it's a free country, so I'm going to.

Let's cover a few.

The first tactic I'll share is actually the central model in Hooked. I'll explain it to you in 15 seconds. It's a four-step loop: trigger, action, variable reward, and investment. And when I read about that model, I was reminded of the Fogg behavior model.

If you haven't heard of BJ Fogg, he founded Stanford's Persuasive Technology Lab. He coined captology, meaning the study of computers as persuasive technologies.

He also published a book called Persuasive Technology, Using Computers to Change What We Think and Do in 2003. That was over 20 years ago. But Facebook was founded one year after in 2004, and even more curiously, in 2007, BJ Fogg co-taught a standard computer science class where students created Facebook apps, optimized them using behavioral data, and they were partly graded on their app's combined engagement score.

One class session was literally titled Engineering Your Way to Viral Adoption. His model says that a behavior happens when three things converge at the same moment: motivation, ability, and a prompt. You need some motivation to perform the behavior. The behavior needs to be easy enough to complete, and then something needs to prompt you to do it. So back to Niri Eyal's model. The first step, trigger.

So a notification pulls you in. The second step is action. You open the app. Third step, variable reward. You get some sort of unpredictable or predictable validation that tells you, yes, it was a good idea to open this app, you should stay. In social media's case, this might be a message, a funny video,

perhaps a little rage bait that makes you angry enough to keep scrolling. And then the final step, investment. First and foremost, you are investing your time, of course. But you also might follow a new creator, post something about yourself, you may save a video, or their absolute favorite thing, you click on an ad. All the while you're giving the algorithm more and more information about you with every move that you make on the app.

that investment makes the app more personalized, but it also makes it harder to leave because now your memories, relationships, preferences, followers, followings, this entire digital profile of yourself is sitting inside the app. So on the platform side, your account is basically an arrangement of numbers that just tells the algorithm exactly how to keep you there. If you haven't watched The Social Dilemma on Netflix, which one of our guests actually helped produce, they explain this phenomenon perfectly.

The next tactic is one we've covered in depth on this show, Infinite Scroll.

The infinite feed removes the end of the experience. I know there are some audience members that are in between 18 and 22 listening, so I don't know if you'll remember this, but for everyone else, do you remember the apps like Instagram back in the day where there was actually a stopping point on a feed? Like you could scroll posts strictly for who you were following, and eventually it would say, okay, you're all caught up. No reason to stick around, really.

so nostalgic.

But could you even imagine that today? How could apps optimize revenue if they don't try to keep you there longer than you really need to be?

If you haven't seen it yet, in episode 10, we actually spoke with the inventor of Infinite Scroll. His name is Aza Raskin. He explained to us the original purpose of the feature. It was always intended to genuinely improve people's experience on the internet. But now it has been mutated to steal our time and attention in a manner so effective it became one of the defining features of a worldwide attention problem. You can listen to that conversation for more of his thoughts on the manner. It's very fascinating.

Tactic number three is a concept called uplift modeling. In this context, uplift modeling is a statistical method used to estimate whether a specific intervention, like sending a notification, actually changes someone's behavior. A normal model might ask, is this person likely to open this app? While an uplift model asks, will this notification cause this person to open the app when they otherwise would not have?

So maybe on my lunch break, I had no plans to reach for the Pinterest app. But because they sent me a notification about trendy dresses for city girls in spring, whatever thing they decided I'm interested in, I do open the app.

a sneakier version of this is ghost notifications. That's when you get notified from an app, but then when you go to open the app, there's actually no notification there, or it's a notification that's been there for a long time. It repeats itself. You've already seen it.

That's annoying. I really hate these and I think I've seen it happen on Facebook more than anything.

I must say I do not have confirmation that this is intentional, but I don't have confirmation that it's not. More frequently are notifications that are essentially meaningless. They're just meant to get you in the app in general. I call them notification bait. It's when you see the little red badge with the number of notifications that you have and it's telling you that you have many. Go open the app. Then you open the app and it's like, you have a follow suggestion, or Sally Jones posted a thread you might like.

You don't need these notifications. A real notification on a social media app, for instance, comes from a peer that you consented to following. Maybe someone performed an action and they intended for you to see it. They followed you. They sent you a message. They liked to post. It is a social exchange. Anything other than that.

Is a waste of time in my opinion.

And seemingly, no matter how many times you click show me less of this or turn off notifications like this, they still manage to send them to you. Not only do I think this is a terrible user experience, but I think it's manipulative. When a user wasn't going to open an app, but they see

a notification and think, it could be a message from so-and-so, they're chasing a social connection. This is an extremely human urge. They're maybe hoping to hear from someone that they care about.

So by not getting that social dopamine hit that they're looking for, it's an easy trick and trap because now they're sitting in the app that they've designed to be so addictive and they're scrolling and scrolling even though they didn't intend to, and they're looking for that meaningful dopamine hit that they never get. Yuck. I hate it.

another good example is LinkedIn. it'll be like you have a new notification and it's like this person looked at your profile or you have five networking suggestions and it's like hey this is not urgent like this is not something that you needed to notify me for. Please do not contact me again for anything like this. But alas they do.

When I became really aware of how ridiculous and spammy the notification situation was, I remember going through each app and changing my notification preferences.

It didn't help that much. It really barely helped. I suspect that some companies actually just completely override your preferences. Or they like slowly roll them back because they want to keep testing how to get you in. at the end of the day, they will do whatever they want because it's their app and you are at the mercy of what they want, not what you want.

And did you ever notice like some notifications create a false sense of mystery? They'll say, like, someone liked your post, someone viewed your profile, see what you missed. The information is deliberately incomplete so that you go to the app to satisfy your curiosity. It's just

A really good indicator that this type of tactic is effective is something called phantom vibration syndrome. If you haven't heard of that, it's when someone feels their phone vibrate even though it didn't. This is a real thing. You might have experienced it yourself. I swear I've actually had this happen with notification sounds a couple of times where I'm thinking I'm hearing a text tone or some other notification when nothing actually happened. Or like if you're with someone else and a notification comes from their phone instead of yours and you think it's yours, so you go to check your phone.

phone

but then there's nothing there and it's almost like disappointing when it's not your phone. I suspect that these are all side effects of uplift modeling.

my advice to you when it comes to notifications is to go to the higher ups. What is your actual smartphone settings? Cut off all the notifications from the apps that are gonna waste your time. They cannot override Android or iOS settings. So try that. For me, very few apps on my phone have notification privileges because.

I hate that. There's also no badges on my home screen. I cannot stand the badges. Because I was like checking them. I did feel the need. Okay, so the next tactic is called exploration versus exploitation. This sounds harsher than it is, but in some ways it is harsh. But just let me explain. So again, in the context of social media only, an example of exploitation would be when an algorithm serves you content that it already knows will work on you.

For example, my algorithm knows that if you put an aesthetic cooking video in front of me, Or if you put clips of Rory McIlroy smacking a driver, there's a good chance that I'm going to stay and watch it. If you always stop for, say, like dog videos or cat videos,

They know that this is a pretty safe bet. It can keep exploiting that interest to get more of your time. Very simple.

Exploration is when the algorithm tries pushing something that it is less sure about because it wants to collect new information. It's testing you. Maybe you have never shown any interest in celebrity gossip, conspiracy theories, or weight loss tips, but your algorithm may slip in one of those videos to see what you do. TikTok has openly said that its recommendation system deliberately introduces videos outside of a person's established preferences, to improve discovery and learn what works across a broader range of audiences.

Another good way to think about it, and you've probably heard this metaphor before, are slot machines.

There's actually a tactic called the multi-armed bandit algorithm. It's named after slot machines, which were called one-armed bandits. Imagine each type of content is a different slot machine and the payout is your attention. The algorithm will mostly put the machines in front of you that have worked on you in the past, but they will occasionally test a different one to see if it performs any better. Your reaction will tell it which machine to show you next. And guys, this goes really deep.

Like I can tell you for certain algorithms, I believe that they're actually using algorithms like this based on your mood, not just generalized everyday preferences.

So let's say you're in a particularly irritated mood. They might push a specific type of rage bait video in front of you, see that you're into it, and then continue to push that type of rageful content for that scrolling session.

for me, I always try to note my mood when I'm scrolling. and I notice something kind of funny. When I'm in a good mood, my algorithm will show me more dance videos. I think it's because it knows, okay, like she's in a good mood right now. The video's gonna capture her attention, let's just keep running it back and keeping her here as long as we can. And it works because it knows that I used to dance. I love watching people dance, and so it detects that I'm in a good mood.

shows me dancing people which I think is kind of funny. But yeah, next time you're scrolling, see if you can notice some of the patterns about maybe the time of day or your mood and just be a bit more curious about what the algorithm is pushing you in that moment because you might notice some patterns that surprise you.

I think it's also really important to note that you do not have to like or comment on a video to tell the algorithm that you're interested in it. You might watch till the end, you may replay it, you may open the comments, visit the creator's page, send it to a friend, you might even just slow down for a few more seconds longer than you normally would.

You need to understand that they track every single second that you move through these apps. I can tell you as a creator, when I post a video, they give you analytics for how each piece of content performs. And they will tell you what the average watch time was. They'll also tell you if people clicked away, and they'll give you an average number of seconds that people stay. It is that exact of a science. So they know when something captures your attention, whether or not you engage with a piece of content.

My guess is that when they're trying to determine what content to push in front of you as an exploratory measure, they're tying you to other groups that have similar demographics and are more likely to like this type of content. So all of your demographic data is relevant. Of course, any other data that they collect about you, maybe what you've purchased, what you've looked at on other websites in Meta's case, all of it is relevant and all of it is tracked. It's a lot.

The next tactic is called variable schedules of reinforcement. You may have heard this before. This is where my BF Skinner rat comparison came from how it works:

a predictable reward is easy to walk away from because if you knew that every tenth video would be funny and the other nine would be garbage, you might decide that the funny video was not worth all of that effort.

A variable reward works differently because you don't know when it's coming or what form it will take.

So the next scroll could be great or it could be a dud. The problem is that you are likely to keep scrolling until you get what you want.

Unpredictability is the point.

what I found is that engineers use something called slate recommendation systems. A slate is a group of posts, videos, or options the app places in front of you all at once. Engineers will actually build systems that evaluate groups of recommendations together, and the algorithm will learn which combinations keep you engaged over time.

so the important thing to understand here is that the algorithm may be doing more than just choosing a single video you are more likely to watch. It may be learning the order that will keep you there. I thought that was interesting.

Okay, the last tactic, and this one's super easy to understand, I promise. It is the simple method of gamification. The most obvious example of gamification on these platforms are streaks. Think about like Snapchat streaks back in the day. You didn't want to lose your streak, so you made sure to send at least one person a Snapchat every day.

For me, I remember seeing this a lot in high school and college. I remember people having

Two or three, even five year streaks, which is wild. But what did maintaining that streak actually do for them? Nothing, really. It was completely trivial. There really is no value to that. What did it do for Snapchat? Everything. It gave you a reason to open the app every day and it kept the app alive. A streak takes an ordinary action and turns it into something you feel like you own. And for some people, it becomes part of their identity. So once you've maintained a streak.

Streak for you know 500 days, skipping one day feels like destroying something that you built. A really good example is Duolingo. People do not play around with their Duolingo streaks because it signals to them a commitment, a commitment to learning a new language. How many people are fluent in said language? Who is to say?

Streaks are just one example of gamification. There's also like think of any time an app awards you points or badges. Think about leaderboards, daily challenges. These are all using game mechanics to keep you coming back. Now, with gamification, just like any of these tactics I've mentioned, I don't think any of them are inherently a bad thing. I think an app that's well-meaning can use gamification to get someone to form a good habit.

Like I mentioned before, that Insight Timer has a gamification system on their app where they do reward you when you get a streak of meditating every day. In that case, I'm forming a habit that I actually intend to form. But if you find yourself using an app that is trying to get you to build a streak that doesn't actually align with your values or what you want to do, I would question if that streak matters at all, and I would completely ignore it. Just take away its power.

Now, I'm sure there are many other engineering tactics that we could find that companies use. We didn't even get into how AI amplifies all of these effects, but I think that these core tactics can really help you recognize when you're being pulled into a behavior that you didn't consciously choose. Human will is a beautiful thing. Please exercise it and understand when companies are trying to take it from you.

I also wanna reiterate that like my beef with some of these tactics does not lie with the engineers.

A lot of these brilliant engineers can be great people. My episode with Aza Raskin is a really perfect example of someone who created something, it got turned into something else, and he doesn't like what he sees. So my frustration lies with the companies that create these perverse incentives and leave us all to deal with these urges that we didn't plan on having. And I do want to just say that.

okay, two new juicy stories for you.

First story is about the EU AI Act. As of August 2nd, 2026, just a few days ago, Article 50 of the EU AI Act requires greater disclosure around certain consumer-facing AI uses. I like to hear this. The rules cover four main things. One,

People must be told when they are directly interacting with an AI system unless it is already obvious. I love that. And side note, I actually hope that they expand this eventually and say that any humanoids in the future also need to disclose that they are in fact humanoids and not humans, whether over the internet or in person. I know that makes me sound a little bit crazy, but I truly have the utmost faith in engineers' ability to make these humanoids indistinguishable from humans

I think that we would be very happy if we made that a rule. Because I feel like we're like one step away from ex machina. two, providers of generative AI must add machine readable markers to AI generated or manipulated content. This is huge. I can interpret this in a lot of different ways, but we're not 100% sure yet. So we'll just wait and see.

Three, deep fakes must be visibly disclosed. Great.

The fourth one is that people must be informed when they are exposed to emotion recognition, biometric categorization systems, or AI-generated public interest text that had no human review or editorial oversight. This is good stuff. This doesn't like solve all of our AI woes, of course, but I'm gonna keep my eyes on this one and see how this will be applied because I'm super curious how they will enforce it.

It's only the EU, obviously, but that's kind of the way these things go. With regulation, the EU is often more careful than the US. but I think that this is overall a step forward and better than nothing.

The second story is about Claude. So about a week ago, users reported that Claude chats were showing up in Google search results. That's not good. These were conversations that users had turned into shareable links. So not just any Claude chat, but you had to share it via a public link. Claude chats are private by default, but once someone creates a public link, anyone with that link can view the conversation snapshot

Search engines then started indexing some of those public links like ordinary websites, which is kind of just how the internet works, like it's not that surprising. But some of the searchable chats

reportedly contained medical information, internal company documents, resumes, API keys, login credentials, and personal contact information.

Again, to reiterate, the problem was that people may have understood shareable link to mean send this to one person or a group of people, but it actually creates a public web page that could spread much further. So Google has removed the chats from its search results, thankfully. But for now, if you do use Claude just for extra measure, I would review your shared chats and just unshare anything that you wouldn't want a stranger to find to be extra safe.

I think the lesson there is be careful what you put in LLM, which you should be doing anyway.

before I leave, if you've listened to my interviews of other people, you know that I always ask them, How are you logging out? and how are you logging in. I do implore you all to think about that for yourself. So as you go into the weekend, think about like how do you kind of want to log out this weekend? Do you wanna

Deactivate Instagram for a few days, or

Maybe you wanna go touch grass and leave your phone in a drawer all of Sunday and go hang out with friends and family. Doesn't that sound perfect?

That's what I'm gonna do.

Get off your phone. Bye.