

JESSICA DOUIEB: Today I am thrilled, absolutely thrilled, to be joined by Tristan Harris. He is now Co-Founder and President of the Center for Humane Technology, whose mission is to reverse human downgrading, and realign technology with humanity.

Tristan, thank you so much for joining us today.

TRISTAN HARRIS: Of course, Jessica. It's really good to be here. I'm excited to get into all of this with everyone.

JESSICA DOUIEB: Awesome. Well, first, two things I want to say. First, thank you. Thank you for bringing to light this extremely important topic of our time in *The Social Dilemma*. For those of you who have not seen the film yet, we won't give away too, too many spoilers. But *The Social Dilemma* really disrupts the disruptors, right? So, unveiling these sorts of hidden algorithms, the words algorithms, these machinations that are happening behind everyone's favorite social media platforms and search platforms, et cetera. So, you know, essentially the film says human willpower is not strong enough to be able to compete, right, with the most sophisticated AI that we have on the planet.

So, Tristan, share with us, right, like how did this start for you? How did your journey as a design ethicist at Google led you down this path or brought you to these realizations?

TRISTAN HARRIS: Yeah. Well, it may be relevant for this audience. You know, actually before this, people don't think I'm just some kind of Pollyanna activist or reformer type. I actually came from the tech industry. So, I went to Stanford and studied computer science. My friends in college were the founders of Instagram and were many of the early employees of these companies. So, I'm very close with and intimately familiar with and have attended the weddings of some of the people whose business models we're now criticizing.

And I had a technology company myself called Apture, which was a tiny, tiny, tiny negligible talent acquisition by Google about eight years ago or something like that. So, I know what it's like to raise venture capital, to play the startup game, to try to, you know, build your proof cases.

And that's what really woke me up to what's really going on and how no matter what narrative each startup founder I and my colleagues knew about or talked about, whether it was the Instagram guys or others, it really boiled down to a race for

human attention. Because no matter what service you're building or whether you're running a political campaign or you're running a conference or you're running an investment bank, what do you need first to get a customer? You need to get their attention. Can you vote for someone or vote for a proposition that you've not actually heard of or have heard any buzz about? No. Can you vote for a politician you haven't heard about or don't recognize? No.

So, it really is about attention. And I saw that that race for attention was getting more and more constrained because as everybody, you know, is feeling-- and my first entre into this was actually distraction and feeling incredibly distracted constantly. The inability to focus. The inability to get anything done in the way that email could just hijack a day. I think people spend hours in Gmail. They just feel like after all those emails were sent back, like, was that a day well spent? Was that time well spent? Or is this just, you know, kind of empty calories of work?

And I actually worked, after I got acquired by Google, I worked in the Gmail team. And so, I was right there in the belly of the beast, thinking with the people, the designers, about what the next vision of Gmail would be like.

And as they talk about in the film, as I talk about in the film, I found it fascinating that all the psychological issues with e-mail, whether it's distraction or addiction or addictive checking, pulling the slot machine to see if you got more, reading an email and then rereading it again tomorrow, and again tomorrow, and not responding, and that loop of anxiety that it creates. I was wondering, would we try to tackle any of those psychological issues: distraction and addiction? And I found it fascinating that we weren't really tackling that.

And I thought, if there was one group in the world that would care about the wellbeing of this sort of digital habitat, because email is the perfect example of a place where people really live a huge percentage of their work and personally life, hours a day in an internet café you'll see people have it open, and they're really wasn't that conversation.

And so then, zooming back a little bit to give people context, I also have a background in human persuade-ability and the human psyche and what manipulates the human mind, because-- don't worry, I'm not manipulating you right now.

JESSICA DOUIEB: Okay, good.

TRISTAN HARRIS: The lab at Stanford I studied in a class called The Stanford Persuasive Technology Design class. And there is an associated lab run by a professor named BJ Fogg. And it really ingrained in me that human beings are persuadable. And we were studying persuasion for good. Because you help people have the habits that they wanted to in their lives, whether it's working out more at the gym or flossing. And then the founders of Instagram and I were in that class together. We actually worked on a project together that was about alleviating depression using persuasive technology.

But many of the alumni of that class went on to work at the big tech companies. And you use that persuasive design to create more and more manipulative products because they were caught and trapped in this race for attention. And that's really what the film talks about, is the way that all of the harms we're seeing boil down to not one technology, overwhelms our strengths and IQ and takes our jobs, which is the singularity that everybody talks about. That's the, like, what's going to happen with AI taking humanity's jobs? There's this much earlier point where technology undermines human weaknesses. And really that diagnosis, which is in the film, is at the root of all of the harms that we're seeing.

You know, our brains have short-term memory of seven plus or minus two things that we can remember. That's why phone numbers are seven digits long. When technology blows past those short-term memory limits, we feel that as a problem called distraction or information overload. Like, I was just doing something. I was checking my e-mail. But then I forgot why. And now I'm looking at something else. That's overloading our short-term memory. Or hacking our limbic system by the fact that we have early emotions for, say, outrage and outrage negative emotions last longer, stick around longer, and are more viral than positive emotions. And then that's how you get polarization, because the most outrageous stuff gets to the top of the newsfeed and then we end up polarizing our societies.

But really, the lens that we use comes from a lens of how is the human mind hackable? And it's never been more obvious when you have, as you said, a super computer that's pointed at your prefrontal cortex, your executive control part of your brain, the front, which is really monitoring, you know, what am I trying to do? What are my goals? And what are my-- how do I stick to what I was intending to do, my willpower? That is the

very thing that's being compromised by a supercomputer that's predicting the perfect thing to overwhelm or undermine those systems.

JESSICA DOUIEB: Talk through a little bit about the impact of social media on mental health and in society today.

TRISTAN HARRIS: Fundamentally, I think the issue is, you know, are the thousand engineers behind the screen, when they make those super computers and predictive models to figure out, you know, what to show you next, are they asking what's best for you or for society or for democracy? Or are they just asking what would keep you on the screen? And their stock price is directly connected to what will keep you on screen.

So, you know, in terms of mental health, I mean, there's so many different things. Right? So let's take a simple example. You know, Jessica, if you posted a photo on Instagram and you got 100 comments for that photo and 99 of those comments are positive, just everything's great, but one of those comments is negative, where does your attention go after? This is a universal human experience. Because we're evolutionarily evolved to care about when there are negative people saying negative things about us because our reputation is on the line with our tribe.

But if you're a teenager, this is especially damning because now it's never been easier to not just see the one negative comment, but to see the tree of then that person was retweeted by someone else, and then that person said it, confirmed and piled on. And now you can go down the hate spiral. And then when you close the laptop or you close your phone and your mind is sitting there, you're doing something else, do you think your mind just automatically lets go of that? Or does it stay with you for the rest of the day?

Another thing is the fame lottery. The idea that we have a number of followers and we have a number of, yeah, basically the number of followers that we have. That is sort of a status indicator for how much influence we have in the world. And when that was invented by some friends at Twitter, actually did that first and then Instagram followed, that was invented because it was a good way to get you to come back everyday and to see how many followers do I have? Am I more famous than I was yesterday? It's sort of like, you know, "he loves me, he loves me not." Like "mirror, mirror on the wall, am I more famous than them all?" And that was all done for this race for attention.

And then that sets up this loop where it's this treadmill of there's always infinite reasons to see more likes, more comments, and for everyone. But the question is, is it leading to a better society? And also, what is it displacing? Because now our society is hyper focused on compressed into the present. It's almost impossible, our attention spans have shortened dramatically over the last ten years, but I think the real message of the film is so long as we the users are not the customer, but we are the product, we're worth more when we can be turned into predictable automatons.

And much like, you know, cows are worth more when we domesticate them for producing greater meat and greater milk, so we don't have regular wild cows, we have the kind of cows now that are best for us. My concern is that we're taking humanity and there's, like, the wild, alive, and fulfilled, and free, and informed humans that are kind of-- we can at least aspire to some notion of who we could be at our best as humans, but we're actually worth more when we're addicted, distracted, polarized, narcissist, attention-seeking, and disinformed than if we're this alive, free, informed human being. Or an alive child who's playing with their friends off the screen. Because that child is worth more if they're a 15 year old caring about the number of Instagram followers they have and making \$10,000 a month and looking at their comments and addictively posting more photos than if they're actually just free, you know, as a human. So, that's really what we have to change at the end of the day.

JESSICA DOUIEB: When you look into the future a little bit, maybe it's only three to five years, right, can't predict too far out, but what do you see in a positive way?

TRISTAN HARRIS: Well, I think it's the degree to which we're willing to make radical change is the degree to which we understand how much is at stake. Because, for example, I think it's important for everyone to get no matter what you care about, you can care about human trafficking, you can care about climate change, you can care about racial inequality, you can care about any of these existential topics that we have to deal with, our ability to work on that problem, like take climate change which I'm very passionate about, depends upon our ability to see the same reality and understand the timelines that we have. Because if everyone understood the same things, we could coordinate on how we want to fix it. Right?

But right now we live in a world we don't understand the same

things. We have a bunch of people, a large percentage of the population who doesn't believe in climate change. And as we talk about in the film, if you type in "climate change is" into Google, whether you saw climate change is real or climate change is a hoax depend on not what was true, but where you were searching from. So, if you're searching from a high fracking country, you might see "climate change is not real." By the way, one fun thing for your listeners is, we'll talk about Russia manipulating the election, one of my favorite examples is Russia actually went into pro environmentalists groups who were anti fracking and they actually dialed them up. They were sort of putting their bot armies to amplify all the anti fracking groups. Why? Because if the United States is against fracking, it's more reliant on foreign oil and it makes Russia more viable, et cetera.

So, there are a lot of games that are being played here. But I think that what we have to do to get regulation and these other things to happen, I'm actually hoping that an incoming Biden administration, if we can get past the current polarized moment, would actually treat this as an issue on par with economic recovery. It's almost like democracy recovery. And specifically a kind of digital recovery of our society. And taking it seriously. And when I say this, I want people to understand, I don't mean the government regulates what people can and can't say, right, and free speech. It means asking the question of what kind of social media environment and business models would upgrade our capacity to make sense of the world and make choices together and solve problems? And what business models are going to fundamentally lead us into conflict?

JESSICA DOUIEB: I know you touched on the new administration. So, given the economics that are fueling the division through social media, juxtaposed with the politics of our country and of the world, what realistic solution do you pose?

TRISTAN HARRIS: Well, I'm actually just thinking about this in real time. In the last 24 - 48 hours I've become really convinced that we need to make some kind of TV or Netflix series of just-- that's going to take too long, frankly. We need immediate demonstrations of people being able to, from their micro realities, have a conversation and actually open up a shared reality. We need demonstrations of that. We need evidence to that. We need to help break down how it is that we got to be so polarized and everyone is operating from the politics of grievance, because they think they've seen infinite evidence of

the other side being unfair, mean, you know, angry, naïve. And everyone has seen infinite evidence of that because there actually is infinite evidence of the other side being naïve, mean, unsympathetic, et cetera.

And if we live from that space of grievance, we are never going to solve our problems. Because there's only escalation from that. As we say in the film, where this goes in that case is more towards conflict and civil war. And we have to show that that's not actually going to be a better result for anyone. Even though it might seem like it's better.

And I think the nice thing is that people are actually unaware of the degree to which social media has really splintered us into these micro realities. One thing I'm very excited about, to be honest, and optimistic about, is that everything we've been talking about the last hour has actually been received positively by both sides of the political aisle. Right? We have incredible fans that are on the left and on the right. That this is the thing that's making us all go crazy. And I think that is so exciting. That the one thing that can make us, you know, nonpartisan is the way that technology drives at partisanship. You know? The one thing that can reclaim common ground is the revealing of how common ground got destroyed.

And so, I think now what the film does, I think, is create common ground about the breakdown of common ground.

JESSICA DOUIEB: Well, Tristan, I feel like I could continue with you for quite some time. But we want to be mindful of everyone's time and your time, most importantly. I want to thank you so much. And it's nice to end on that positive note. Thank you so much.

TRISTAN HARRIS: Thank you so much. Really enjoyed doing this with you. And I hope people can check out more resources if they want to through humanetech website.

JESSICA DOUIEB: Will do. Absolutely. Thank you.

TRISTAN HARRIS: Thank you.

This transcript should not be copied, distributed, published or reproduced, in whole or in part, or disclosed by any recipient to any other person. The information contained in this transcript does not constitute a recommendation from any Goldman Sachs entity to the recipient. Neither Goldman Sachs nor any of its affiliates makes any representation or warranty, express or implied, as to the accuracy or completeness of the statements or any information contained in this transcript and any liability therefore (including in respect of direct, indirect or consequential loss or damage) is expressly disclaimed. The views expressed in this transcript are not necessarily those of Goldman Sachs, and Goldman Sachs is not providing any financial, economic, legal, accounting or tax advice or recommendations in this transcript. In addition, the receipt of this transcript by any recipient is not to be taken as constituting the giving of investment advice by Goldman Sachs to that recipient, nor to constitute such person a client of any Goldman Sachs entity.

This transcript is provided in conjunction with the associated video/audio content for convenience. The content of this transcript may differ from the associated video/audio, please consult the original content as the definitive source. Goldman Sachs is not responsible for any errors in the transcript.