URL: https://stvp.stanford.edu/blog/videos/making-technology-less-manipulative-entire-talk

How good are you at limiting your screen time? Because of the way humans evolved, our brains are no match for the engineers, designers and companies that collectively create the devices and apps that demand our attention all day long, according to technology ethicist Tristan Harris. A former tech entrepreneur himself, Harris is now co-founder of Time Well Spent, a nonprofit movement to create an ecosystem that aligns technology with our humanity.

Transcript

- Thank you guys all for coming.. Yeah it is very weird for me to be here.. I did come here as a student and listen to many famous entrepreneurs come and give lectures.. I remember the first one that I came I had a venture capitalist or it's like a real estate guy come up to me and say well, he gave me his business card and he said here if you ever need office space you know call me and I was like what do you mean and he was like you're going to start a company someday so when you want to just get some office space.. So yeah my role in the world now is very different than I actually thought it would be when I was at Stanford.. I graduated here in 2006.. I was an undergrad in computer science.. I took mostly symbolic systems classes.. I was very interested in cognition, in the mind, in neuroscience, in psychology and I got particularly interested in a lab here called the Persuasive Technology Lab.. Does anybody here know the Persuasive Technology Lab? Some of you, okay, if you didn't know the Persuasive Technology Lab is basically taught by B.J..

Fogg, he's a psychology professor and he pioneered this field of persuasive design.. How do you persuade people how does that work and how could technology persuade us for good? That was the question and this lab was filled with young engineering students, mostly computer science, business, psychology students many of whom went on to join the ranks of you know Facebook and actually my project partner in the class Mike Krieger was here from Stanford and founded Instagram actually using a lot of the techniques that we designed and learned about.. And the thing that caught my attention in that class was this last semester, the very last class of the semester, was about what is ethical persuasion.. What does it mean to be able to ethically and personally persuade a human being? And we actually talked about this thought experiment well what if you could get the perfect profile, let's say every amount of information about how this particular human being's mind works and you could use that information to persuade them towards anything.. So this was sort of a thought experiment that we were running and it turned out that that question became ten years later the thing that I now do basically for a living and am concerned about, and that frankly is one of the most important hidden and invisible topics of our time right now.. If you're aware in the last election including today where the hearings in Congress with the major technology companies, Facebook, Google and Twitter testified before Congress specifically about the ways that their platforms were used to persuade the American public in the election and what Russia's role was in that in using these platforms.. So, and if you didn't know Cambridge Analytica which is a company that basically the Trump campaign used and also for Brexit to basically create a personalized profile of what would persuade people.. So these questions that I was really interested in became really the foundation of possibly one of the most important you know things that's going on in the world right now.. So why does this matter? Why does ethical persuasion matter? Really this is a conversation about values.. You know when you're in this class and you're thinking what does it mean to ethically persuade people you think well who's to say what would be good for someone else..

Well we want to help them, we want to make their world better, we want to make their world more open and connected, we want to help them sell stuff.. So we're starting to define these values right of like we want to persuade people to be more open and connected.. But what I want to talk about today is where our ideas about making the world better meet the rubber meets the road with reality and what it would mean to actually make the world better with persuasion.. So I want to tell you a little story about kind of how I got into this and got concerned.. When I was a kid I was a magician and I basically studied from a very early age that this instrument that you're seeing me through right now your conscious experience can be manipulated.. There's limits to our attention, there's limits to what we can see and not see, there's limits to how we can think about something.. If I split this room in half and said is the number of countries in Africa greater than 50 or less then 50 and I

asked you that and I asked the other group is the number of countries in Africa greater than 150 or less than 150.. Just by anchoring these two groups on different numbers you would get different results.. And what you start to realize as a magician is that the human mind is living inside of a 24/7 magic trick.. You're basically inside of an instrument that you can't see that you're experiencing me right now and that there's certain levers and strings you can pull on to get people to think about things in a certain way..

But I say this because this was a childhood interest and I didn't really do it for that long.. I gave like one or two magic shows, but it gave me the kind of foundations of that thinking.. I want to explain that I later came to Stanford.. I took the Mayfield Fellows classes.. I got very interested in entrepreneurship and I started a small company called Apture which was about learning and helping people learn things on the internet.. And we thought what if we could apply all that stuff I learned at the Persuasive Technology Lab for good, let's use it for good.. And oh I know what good is cause I'm a 23 year old you know smart kid from Stanford.. So we built this thing that basically let the Economist would be one of our customers and you would highlight something on the page and we would persuade you to basically go deeper and learn about something.. We were just making, we called it lighter fluid for sparks of curiosity.. We thought what could go wrong..

We have pure good intentions, we're just trying to help people learn things.. We're persuading them to learn.. And we used all these littler persuasive design techniques.. And as I was building the company, you know we raised venture capital, we had a team of 12 people, as I was sitting there and what got me waking up in the morning was gosh you know I really want to help people learn things.. That was about all I could think about, but at the end of the day everything that my company did was not actually about helping people learn what it really mattered for the business was whether or not we could tell the Economist that we were increasing how much attention or time people were spending on its website.. You guys following, so I had this goal of hey on the web we're going to provide this technology that's going to make it easy to learn about things, but the way we sold that to the publisher of the Economist was by saying we're going to help people spend more time on your website.. So I had 12 people that worked at the company and I recruited them by being really passionate but being just you know classic like we're going to change the world, we're going to make people learn about things, it's going to be great.. And then really believe it, I was very persuasive as I was the CEO of the company.. And I honestly as the founder of the company couldn't admit that there was a gap between my positive intentions as a human being waking up wanting to do this good and the ultimate thing that we were measured by which was to capture human attention.. And you know when you're young and you're raising money and all your friends are starting companies there's this temptation to compare yourself, right? You know they raised this much money from DFJ and I only raised raised it from (mumbles) VC or you know they recruited this brilliant guy from LinkedIn and oh well we recruited these other people..

And there's all these dynamics when you're running something after you've built together a team and you're basically out there in the world with your company and your product that how deeply will you be able to examine your own motivations and beliefs and see maybe what I'm doing isn't about helping people learn at all.. If that were true would I be willing to see that that was true.. Now think about it if you've got you know we raised four million dollars, we had a team of 12 people, their families, their livelihoods depended on me.. I had a huge amount of responsibility.. I was just trying to make the business work and here I was starting to question whether the entire premise, the values, the core of what I was doing everyday were we actually doing that.. Some friends of mine who were founders of tech companies as well got together one weekend.. And we got together because we all talked about how there's all this pressure that you face when you're an entrepreneur, this pressure to succeed and to talk about how you're growing, how things are going better that you can never express your doubts.. Right, every time someone says how's this going you say oh well we just raised this much money, we just closed these ten more customers, it's going great.. You can never just say like you know I don't even know what I'm doing.. I'm not sure that I know why I'm doing this anymore..

Like where is it safe to utter that sentence.. So my friends got together this one weekend and we had a group called Doubt Club and Doubt Club was basically acknowledging that it's not safe anywhere to talk with other founders about your doubts about your company, your product and your life.. And we basically went around in a circle, that was hyper confidential, it was very safe.. We basically said look the commitment here is even above your connections and other employees who might or might not work for you or if other investors are looking at your company, this comes first in this small private group.. And we basically shared our doubts about what we were doing and it was in that moment that I realized that there was a huge gap between my positive intention that I really wanted to do good, but there was a gap between this learning goal with our company Apture and the actual delivered thing that we had to do everyday which was capture human attention.. I was able to see that.. And then I had to figure out what am I going to do.. So Tina's told this story that makes it sound like I'm some, oh he sold his company to Google and what a successful entrepreneur, but I'm going to tell you the truth which is I was terrified about what we were doing.. I felt really lost about why I was doing everything I was doing and we had to figure out what to do.. And we ended up you know being in a situation, we could have raised more money, but we were also shopping the company around and seeing what would happen and ultimately we soft landed the company at Google..

And so we had a nice out.. We had two other, two of the other people in Doubt Club ultimately one got acquired, talent acquired for their company, we got talent acquired, and the other actually just shut down his operation, decided to move to Berlin and do what he was passionate about.. But I say this because it's incredibly, incredibly hard to question what you're doing.. Upton Sinclair the writer wrote that you can't ask someone to question the thing that their salary depends on.. It's even harder to get someone to question something that they deeply believe is their purpose.. It's like the thing about

questioning someone's religion or something like that or questioning someone's identity.. And the reason I'm talking about this is because all of us in this room I imagine have good intentions for where we want the world to go and what it means to value certain things that we care about.. But we're only as good as the way that we can self-examine our own values.. So here I was at Google and I was kind of recovering as an entrepreneur.. I felt honestly like I'd failed..

I need to get psychotherapy, I went to Burning Man.. I had to kind of let go of you know the life that I was living and realizing that there was more to life than just starting companies right.. And I saw all these friends that had started companies and then I was sitting at Google and I was actually joined the Gmail team and working on some future looking personal assistants type projects and I was working with Gmail and I was in the room with the people that make this product.. I was fascinated, here's this Gmail product that literally I don't know a billion people actively use and you know that hundreds of millions of knowledge workers that's the window, it might be open on some of your laptop screens right now for all I know, right.. We people live in this product and it has this extraordinary influence on the thoughts that arrive in people's heads, right.. Just like you're sitting there if an email comes in right now it's going to push thoughts into your mind and you don't get to choose whether that happens it just happens.. And so I was sitting in this room thinking about you know I really cared about products that really delivered a positive benefit in the world and I was in the room with the people who would be thinking the most deeply about what it would mean to truly benefit people's lives, what email should do, what are the values behind email.. Like of all rooms to be in in the world to be in that design room is like that's the room.. And I was, I don't want to say disappointed, but the way that conversations were had were about let's make it really fun to use, let's make it engaging.. What if we make it bounce up when you scroll..

What if it expands vertically instead of slides horizontally and there was all these design questions like that and I felt that there was something missing.. I couldn't put my finger on it for a while but I realized later it's like when is email actually adding up to a real net positive difference in your life.. Think about the emails that you send whether it's love letters or apartment searches where you're actually getting some real actual delivered life change, a real benefit, a real value.. And then think of all the other stuff that we're just kind of shuffling messages back and forth and I felt like of all rooms to be asking this question we were not asking the biggest question and after a year at Google I was kind of burnt out and I decided that I was basically going to leave, at least I thought I did, and before I did I kind of took these concerns and I made a presentation.. And the presentation was basically never before in history have 50 designers in California at three tech companies influenced what two billion people will think and do right now.. And we have an enormous responsibility as a company that shapes the screen in terms of what we are causing people to do.. And the reason I was even thinking about that is because I had this background as a magician and in the Persuasive Technology Lab where I learned that minds really are steered by forces that they don't see.. So I have to say that in making this presentation I actually thought I was going to get fired.. It was not alarmist or angry or upset, but it was very critical and it was very existential.. Like what does it mean for actual email to be a benefit to people's lives..

And I was surprised because I sent it to about ten people and when I came into work the next morning I got some emails back and I clicked on their comments and I went to the presentation, Google presentations and it showed the number of simultaneous viewers was 150.. And when I clicked on it later that day there was 450 simultaneous viewers.. And basically I saw that it had spread virally around the entire company.. Went all the way up to Larry Page.. He was in three different meetings that day where people brought it to his attention.. And suddenly I was in this moment where I felt like you know the whole company was woken up to this question of are we actually influencing the world in a positive way.. And that led me to the next three years and actually a executive at Google saw this and basically generously offered to host me in a little corner in his lab in New York where I could study ethical design, which I basically self-titled this role of design ethicist and this new field of what is ethical design or ethical persuasion.. What does it mean to ethically influence what two billion people are going to be thinking and doing.. And so that's basically what I did for the next three years and I went really deep into understanding first of all you know if you take the human evolutionary code, you're living inside of a meat suit mind body, right, that was tuned millions of years ago.. So our predilection for sugar, salt, fat was tuned millions of years ago when they were scarce, but here we are now and they're abundant..

So we've got all these tunings and you're living inside of it and it can't really change.. You don't really have much optionality.. So you're living inside of this thing and if you made a map of every single string you could pull on this mind body system to persuade it, if you could make a map of how could you addict the human body or a human mind, how could you pull on its sense of belonging, how could you make it feel like it's missing out, how could you get it to do certain behaviors, how could you get it to think about certain things, how could you get it to make certain choices.. So if you had a map of every single way that a human being could be manipulated that was the first part of the study, the task.. Second part of the study was what would it, yeah what would it mean to ethically persuade, so what's constituted in ethically pushing this human animal around in the world and then the last question is values.. What are we pushing it around for and who's to say, how do we know we have value we can stand on, that we can actually persuade people in an ethical way.. So I basically studied this topic and I want to say that I didn't know what I was doing.. I was basically trying to figure out this answer to this question of whether we want to or not Google's going to bump its elbow and Apple's going to bump its elbow and Facebook's going to bump its elbow and a billion people are just going to go in these different directions, right? Because when you wake up in the morning it's important just to set the context you know two billion people that wake up in the morning and the first thing they do is they like check their phone.. And we check it 150 times a day you know in the bathroom, coffee line, going to sleep.. We spend a lot of time on these devices and even when we're not looking at the device the thoughts that are in your mind right

now are still you know partly set by the time you did spend looking at the device..

So we have this kind of 24/7 immersion in this environment.. And I didn't know what I was doing in studying this question.. I just found it to be fascinating and important and interesting and here I am literally whatever it is three days, three years later.. It's November 1, 2017 and the US Congress is questioning Facebook, Google and Twitter about exactly the stuff that I've been interested in for the last three years.. And I honestly find myself right now at the center of one of the most important and visible problems I think in history.. Which is that it's not just that there's this system that's kind of bumping its elbows into people's psychology, but we gave the system a set of goals that are causing tremendous harm.. So what do I mean by that? The reason I told you this story about why I have these positive intentions and with my company Apture we wanted to help people learn and how there was this Doubt Club process to try to figure out what that might have been, you know what really would have been underneath there and you know that I had to actually examine the core beliefs that I had is because I think that just like everybody in the technology industry all these companies that were up there today in Congress have very positive intentions for the world.. Everybody at Google and Facebook and Twitter that I know really, really cares about delivering the best possible you know world that we can create.. And yet I would say that something like the statement our mission is to make the world more open and connected is just the same as me saying I want to help people learn about stuff.. Because what is that actually about? What it's actually about is capturing human attention..

Mark Zuckerberg actually faces the exact same dilemma that I faced except he raised a lot more money and his company's in the public stock market and one of the most profitable companies in history and controls what two billion people will think everyday.. But how could someone like Mark Zuckerberg question whether or not everything that they do everyday is actually about making the world more open and connected.. If you think it's hard as a founder when you're comparing yourself to your friends and whether or not you hired the good people or you raised money from some impressive people.. Talk about what it would be like to run one of the most powerful corporations in history influencing what two billion people in every language think and believe, the terms of people's social relationships and to actually question whether or not the thing that you think you're doing, because you know you're good, you know that you're trying to do good, but how that might be different than the actual result.. And the first thing you have to do is pay attention to incentives.. And I think this is really important as an entrepreneur because all of us, I'm assuming many of you are going to go off in the world and start companies, is pay attention to who's paying who.. What is the actual thing that you're beholden to? With my company it was the Economist, they were our customer and we had to do one thing for them which was keep people on their website for longer.. Let's look at Facebook, their business model is advertising.. No matter what good they want to do in the world their stock price is dependent on how much attention they capture.. YouTube's, Google's stock price, part of YouTube, is dependent on keeping people's attention..

Twitter's stock price is dependent on keeping people's attention.. Everything else they say is an intention that's outside of that, it's a dream.. It's something that they'd like to have happen, but the end of the day, the thing that they're beholden to is capturing human attention.. And I want to talk about why the situation, and I'm going to take you down a little journey which might leave some of you more alarmed than you intended to be, but I promise that I'll turn it around at the end.. I want to scare you for a moment about where we are because I think we're in a much more dangerous situation than people tend to recognize.. The goal of capturing human attention becomes this arm's race for who's better at being a better magician and pulling on the strings of the human mind.. Right, so you have these products competing and everyone's trying to figure out how can I get more attention and then my friend Mike Krieger on Instagram says hey, just like with Twitter, let's add the number of followers that you have to our product.. If we add the number of followers we have then everybody has to login everyday to see how many they have and they want to get more so they have to come back to the product everyday.. So this one tiny design choice, right, these products just evolve, it's like an organism that evolved this new hand and that hand is really adaptive for the adaptive environment which is what's good at capturing attention.. We just invented this new persuasive thing..

But what could the consequences of that be? And how would you as an engineer, designer be thinking about that? You might say we're helping the world because now people know like who's following them and now people know that they can connect with certain people that are interested in their interests and you can see the list of who follows you and there's all sorts of benefits, positive things that could come from that.. And when you're a human being living inside of the eyes and mind and beliefs of someone who makes Instagram or Twitter you're thinking about it in terms of those positive things because that's what you're trying to do, but how would you know that might cause a whole bunch of other externalities because right now the number of people who define their self-worth based on the number of followers they have.. David Brooks wrote a book called The Road to Character and he talks about the world values survey and how people value different things over time and one of the big things that's changed is that people went from valuing fame as the number 18 on the list to valuing fame as the number one on the list.. So, and that's in the last, sorry that's in the last I think ten years or something like that.. And I would argue that the reason that like literally billions of people now value fame higher on the list is actually because of little tiny seemingly innocuous things like putting the number of followers that you have in the core of our software interfaces, right? So we have this natural situation where everyone's competing for attention and we are evolving these systems to get better and better at extracting human attention and we, then it gets really competitive and we have to add something else.. So what do we add, we add AI.. So now instead of just offering a product to you I actually have to predict with big amounts of data and machine intelligence what's going to keep you on the screen.. Instead of just offering some stuff you can click on I'm going to actually predict from millions of things I could show you, if I'm YouTube, what's the video I can

put in front of you.. If I'm Tinder I'm going to pick from millions of people I can show you what's the perfect reason to not be with the person that you're with.. If I'm Facebook I could say all the million things I could show you in the news today what's the perfect thing that's going to get you to like, click or share, okay..

So you have an AI that's basically been given this goal.. Now I want you to put this in context when we think about AI we have to remember you know when you point the AI system at chess, right, first it kind of wiggles around and it makes some kind of funny looking moves and then it starts making some smart moves as it gets better and then it makes some surprisingly smart moves and then it beats Garry Kasparov.. And when it beats Garry Kasparov it doesn't unbeat Garry Kasparov.. It's now better than all human beings at chess, right? So you take that same AI and you point it at the Go game and that took 30 years.. It made these funny looking moves and now with AlphaGo, with Google's AlphaGo, it beat all of Go players and when it beats all of Go players it doesn't unbeat all of Go players.. So we built these AIs and then we actually invisibly gave it a new target.. We pointed it at this and we said whatever gets this human being, play chess against this human being's mind and play 20 steps ahead of where their mind can even possibly see and show stuff to them that is either the perfect next video on YouTube, the perfect outrageous news story on Facebook, which controls what two billion people will think everyday, the perfect reason for you to cheat on your spouse with Tinder, the perfect political message so we can vary 60,000 political messages and we can actually combine word choices and different contortions of politician's faces and colors of buttons to perfectly animate a response from your brain stem.. We're playing chess against ourselves and we have every click that we give this system it gets stronger, right? Every click or you share you're feeding it attention which feeds it more dollars which feeds it more resources which feeds it more computing power which means it's better at playing chess on your mind.. The reason this should be alarming is that these systems are not neutral.. We have tendency to say you know we've always had computers, video games, radio, TV, we always worry about them so why should we be so concerned this time? And there's a few different reasons that this is so different and the biggest one is this AI enhancement that I just mentioned..

Couple others are that no other medium could pull on your social psychology.. So no other medium could show you an infinite set of reasons why other people are living better lives than you are.. An infinite set of reasons for why you should feel like you're missing out, other people are having fun without you.. An infinite set of reasons why you owe people responses.. You didn't open your TV and it said you know you owe 100 people responses you better start getting back to them, right? So we're doing such enormous, I don't want to say harm or damage, but we're doing, there is so much that is now put on the human evolutionary animal, right, more than we've ever put and especially when you add in AI.. So now the question becomes what is the goal of that AI.. What does it actually want? So if this is not a neutral product, if this is not just sitting here, but it actually wants something from me well if you're Mark Zuckerberg you think the thing you programed the AI to do is to make the world more open and connected, which you translate into let's engage people.. Let's show people whatever engages them.. And that's going to be very persuasive to someone who's living inside of that mind.. But we have this problem because the actual thing that this is all based on is attention..

It's the same contradiction that I felt.. We now have this system that's, you know we always talk about, I don't know how much you talk about it here at Stanford, but there's all of this discussion about runaway AI.. What if in the future we were to build a runaway AI, like a paperclip maximizer and we give it this goal to make paperclips and it turns the whole world inside out just to create paperclips.. And what if that would happen and how would we make sure that wouldn't happen? So there's all these people working on AI safety and the amazing thing is that this basically already happened.. Because we already built a runaway AI that's steering what two billion people's thoughts are and we hid it from society by calling it something else.. We played a magic trick on the human mind because if you call it a newsfeed or you call it YouTube recommended videos or you call it Tinder recommendations people won't even notice, that's how easy it is to fool the human mind.. So let's see where do I want to go.. What we're really dealing with is we have systems that have exponential impact, right? Facebook just disclosed in the hearings today and yesterday that there are more than five million advertisers on Facebook, okay.. And that means that there are also potentially hundreds or thousands or millions of campaigns per advertiser.. So you've got this combinatoric explosion and so if you have China or North Korea or Russia or someone else who's trying to spread manipulative advertising, not even for a politician, but just for divisive issues or conspiracy theories or lies..

The problem is we've created an exponential system of persuasion without exponential guidance or exponential ethics to control it, to govern it.. And so where we find ourselves is basically the situation that many of the people who worried about runaway AI have been talking about forever and that's the situation where we're at.. And I don't mean to depress you, but that's actually where we are like right now like today and the Congress is just now waking up to basically the reality of what this system has the power to do and the cat's out of the bag.. Now you could say well hold on a sec when we talk about these future scenarios where if you have runaway AI there's always this discussion of like we'll let's put it inside of an airgapped computer and it won't like get out, right? Or we can always pull the plug, let's just shut the thing down.. We could always pull the plug.. It's not going to be a problem because we can always just turn the computer off if we have this runaway AI in the future.. But the problem is if you think about what that would be today, that would basically be like turning the lights off at Facebook.. Now it's not as if from an existentialism perspective as a human being you can't like get out of your chair and walk out of the room in space and time and tell the board of directors you know we're going to shut this thing off.. You could physically do that, but there would be consequences and we can't turn off this system that we've created.. So the only way to solve a system that's runaway pursuing its own goals is not to manage your relationship to it..

I'm saying this because I want to contextualize, a lot of people think of, if you've read any of the work I've done that it's

about we've got to better manage our relationship to our phones.. Let's be more mindful with our phones.. That's like saying when chess is kicking your butt let's manage our relationship to the chess that's just totally overpowering where we are.. So the only way to solve this problem is actually to put the AI on the same side of the table as us.. We have to be really honest about what the lines of power are.. So right now the thing is pointed at us, it's extractive.. It basically says I need to extract as much attention out of you as possible and the classic line is it's a race to the bottom of the brain stem.. So it's not enough that I offer you a product, I have to reach deeper down into the brain stem and create an unconscious habit so now you actually pull for the phone more often.. That's not enough I have to reach even deeper down to the brain stem and own your social psychology.. So the way that Snapchat basically owns people's social relationships..

That's not enough I have to reach deeper down into the brain stem and get to your self-worth and control your sense of how often people like your self-worth.. So I told you I'd go dark, I apologize.. So we have this extractive economy and if you'll notice it doesn't look that different from how the broader economy is structured when it's extractive.. It's the same as the environment.. We have a runaway system that makes more money the more you extract without putting something back in balance.. So the fundamental situation here is we need to find a more ergonomic relationship with the boundaries of our architecture.. In the case of how it's extracting us.. And if you don't care about us let's talk about our kids because this thing is just eating kids alive for breakfast.. So we can do that, but I think that what this is about is just like the point where we were where we realized you know after we extracted coal and we created this incredible economic prosperity after generating energy with coal.. We also polluted the external environment and no one really saw that at the time..

Good intentioned people created this bad stuff and we just realized it later.. And we're at that point where we're just realizing that this extractive attention economy where we're capturing people's attention is basically polluting the external environment, it's also polluting the inner environment and now we need to invent solar and green basically solutions.. Like things that are regenerative or ergonomic or replenishing to the human being.. So you can make all the same metaphors of the environment, right? And that's going to mean that things don't grow the same way.. One of the reasons why advertising is such a great business model is because it grows you know incredibly fast and is super profitable.. I mean who wants Facebook to be regulated when the stock price is through the roof right now.. But also who wants the changes that we're going to need to make for climate change to be abated given that we're going to have to make huge sacrifices.. And so the conversation actually comes back to values.. It comes back to am I willing to value something else besides money.. Am I wiling to value something else besides status? Now this is an open question when it comes to corporations who basically they don't have control or choice..

We created these robots called corporations that are just maximizing profit and they have to.. There's no like way they can't do that.. So we have to acknowledge when that misalignment exists.. And it's the same for our own lives.. You know just because I can start a company and my friends, you know I have friends personally who've sold a company for a billion dollars.. Now that's very seductive.. We can be pulled, it's a magic trick.. We can be pulled by basically other people's success or incredible amounts of money, right? But the conversation always comes back to what do we care about, what is this for, what is the problem that this technology is the solution to.. Why am I doing this? And if we ask those questions that's really the essence of ethics.. You know I have this fancy title of being the former Google design ethicist..

I will tell you I didn't study that much philosophy at Sanford.. I think the deepest ethical question you can ask is simply asking questions and asking why.. And so with that I will open it up for some questions.. So thank you.. (applauding) Okay.. - [Audience Member] Can I take the first question? - Sure.. - [Audience Member] So Tristan this is really fascinating and really important.. Are there examples of companies that are doing this well? - Yeah, are there examples of companies that are doing this well.. You know I think that, yeah so we can get into the theory of change.. There is a whole bunch of companies who's business models are aligned with us..

Anyone you're paying, right, the money is going directly, you're paying the person that's serving you.. So you know we actually on Time Well Spent's website we have a bunch of different products that we basically show, you pay for the product and they basically help you.. Calendly helps you schedule meetings efficiently, basically name any product that you pay, it delivers value and you pay like on a monthly basis.. So anything like that.. I don't want to give, I mean for some reason these examples are sort of fleeting my mind, but there's all sorts of products that are like that.. I think the main thing that we have to do is recognize that the advertising supporting products are the ones whose incentive is to extract from you and it's always hard to acknowledge that because we get also benefits from advertising.. We get the fact that we enjoy getting an ad for Nike shoes that we really wanted to buy.. But the problem isn't the ad, it's the misalignment of incentives, so yeah.. And what if we paid for Facebook for example.. But question in the back..

- [Audience Member] Hi, so I see your talk as a magic trick in itself and is the first part where you present something ordinary, this humble CEO of a startup in Silicon Valley and then you have a turn which is and here's what's behind the curtain and the Doubt Club and your journey of you know discovering the extraordinary on top of the ordinary, but then the third and most difficult part of any magic trick is to prestige, is how you get it back, how you come, how you bring this extraordinaryness into the ordinary and so your magic trick was about your journey and your insights and in fact how this journey allows you to create a kind of magic through design or through whatever it is you're doing.. But my question is is it the purpose of this to create people that believe in magic or is it to create more magicians because what I don't understand is that we're spending all this time trying to create the perfect system, the perfect Gmail, the perfect platform which is in itself a

magic trick so that people can continue using it so that people believe in its magic ethicacy.. But my question is why don't we spend more time creating magicians - Yeah.. - [Audience Member] and decentralizing design or allowing people to design their own experience.. - Yeah.. - [Audience Member] Different contexts.. - Yeah, I appreciate your question.. I think you're taking the magic metaphor further than I was intending.. (laughing) Yeah, it's not about the magic of ethics or values or these kinds of things.. I think the point of the magic frame is that it exposes the fact that we're inside of a system that can be pulled or drawn..

It's more delicate or vulnerable or manipulable than any of us would like to admit.. And right now all of our design institutions and our metrics don't account for that model of human nature.. They're governed by simpler ideas like we're giving people what they want or if you clicked it that's what you wanted, which is very persuasive when you tell someone that, but then you can imagine if you said well are there any cases when someone clicks something and it's not what they want, that they were outraged, they were upset, they were, you know that kind of thing.. Another question? - [Audience Member] So you said that these features of companies who compete against each other, but for instance if you were to tell one of them to just remove, like for instance how do we persuade a company to remove one of these time consuming features without them jeopardizing their competitiveness in markets.. How would you - Sure.. - [Audience Member] tell Snapchat to remove Snapchat streaks? - Yeah.. - [Audience Member] You know how would you do that? - So the question is really in the attention economy where everyone, doesn't even matter if you're building Snapchat or you're building a meditation app, even meditation apps need your attention and to figure out how to put a habit inside of your body so that you use it everyday.. How do you tell any company not to get people's attention or to subtract some of these manipulative features from their products is your question.. So the answer is you cannot obviously get any person whose business model is to maximize how much attention they get from you to not do that.. You can't tell YouTube to not try to show people the next videos, you can't get Snapchat to now subtract the streaks feature so the way to do it is you have to go up a level and you go up a level by going to the device and the platform..

So in other words we can't actually ask Facebook to do something against its business model but you can ask Apple, Microsoft or Amazon or Samsung who are companies whose business models set up the choice architecture, which is the interface between you and this army of things that wants your attention.. And so right now that choice architecture because we have this model of human nature that says people are clicking or choosing freely is just wild west right.. Like basically when you get a phone it's like Apple and Google give totally open door access of all those apps to just reach into you and do all the stuff that they want to do.. So there's a bunch of ways they can clean up and mediate that relationship.. So it's like if we're jacking people into the matrix Do we want to jack their impulses or do we want to jack basically the top hud of the reflective minds and there's ways of designing the choice architecture on a phone like a home screen or notifications that are basically asking the most reflective part of ourselves what we want.. And to make that concrete, that's the theory, the concrete version are things like conversational interfaces.. So Siri, the air pods, watches, things that have supplemental interactions, that are peripheral interactions in which you basically are doing something with a device but you're not giving people a bottomless bowl of things to scroll through and explore in an infinite world which is what they're all based on now.. So that's one theory of change is the platforms.. The other one is governments.. We can talk about that more later yeah..

- [Audience Member] So that sounds like you're talking about further up the chain.. I'm just wondering on the other side if you move down the chain, I guess I'm wondering if you're seeing culturally an acceptance that this is a real problem, you know is it widespread or is this just a small group that feels this way today and is that important driving change.. - Yeah, so is it important for driving change, so is there a large group of people who agree or see the problem and is that important for driving change.. I'll say a couple things on that one is I felt so nervous making that first presentation at Google, I thought man I can't be, I honestly thought if this problem existed someone else would have thought of this and said it already and I felt really nervous putting the idea out there because I was basically saying humans don't choose freely, here's all these ways that we don't and I felt uncomfortable saying that.. I felt very vulnerable and it actually has been this process over four years of being repeatedly validated where people understand that this really does happen, this really does work this way, that I've felt more comfortable stepping into it and I'll say since then the acceptance from more and more ex-alumni of the big companies have come out, at least privately to me, in agreement.. And so basically I was on 60 Minutes earlier this year with Anderson Cooper talking about this problem we called it brain hacking, and it was mostly the addiction part of this problem and since that interview Mark Zuckerberg's personal mentor, Roger McNamee who actually convinced Mark not to sell the company to, was it Yahoo or Microsoft for a billion dollars, and now feels incredibly conflicted about his role in creating Facebook.. He's actually partnered up with me and we've been doing all this great work together and there's been more and more people who've been coming out of the woodwork because they understand this is actually what the situation we're in is.. So I would actually say if you don't see the situation this way, and for someone who doesn't and they're in the industry and they actually work on these products what's happening is more of the Upton Sinclair can't get people to question the thing that they're doing thing and we always want to look for the positive, so.. Yeah.. - [Audience Member] Many have a perspective on this..

Is it true the executives of a lot of the companies send their kids to schools where there's a ban on all of this stuff? - Yes the question was is it true that, are we okay? Is it true that executives send their schools to places where they don't have phones and there's limits.. The answer is yes and I think that's always a really telling signal is when the CEOs of companies don't feed their own kids with whatever it is that they're making.. It's an extension of the golden rule, It's like don't just do unto others what they would do to you, do unto others what you would do to your own children.. You know I know Steve Jobs limited his phone, his kids phone and tablet access.. I think Sheryl Sandberg doesn't let her kids use social media as I

understand it.. I know there's people at Apple that are very high up that send their kids to Montessori schools and I think at Google it's very similar.. And I'll just say in parallel when the CEO of Lunchables food line, there's a billion dollar food product, would not let his own children eat Lunchables and you know that there's a problem when that's true, so.. Yeah.. - [Audience Member] So a lot of your talk was my company has this responsibility of ethical behavior but like what can we as consumers do? Like what do you suggest we do if we don't want to like have our minds to be pulled all the time.. - So what should we as consumers do if we don't want our minds to be pulled at..

The first thing to recognize, and this took me a long time to kind of get across, is I interviewed all sorts of people who were like the best experts in the field on persuasion and behavioral economics and magicians and you know I had lunch with, years ago when I started working on this with Danny Kahneman who was actually one of the founders of the field of behavioral economics.. He's the Nobel prize winner in behavioral economics which is how we can kind of fool people.. And in his book Thinking Fast and Slow he says look even though I know how all these techniques work and how these cognitive biases work it still works on me.. Now the important part to realize is okay if you told someone I'm going to give you access to know what your DNA looks like then you don't think well now I can change my hair color from red to something else, right.. So I think the challenge is that we really are inside of this experience.. We're really inside of cognitive biases.. If I did that Africa number of countries thing and I anchor you on one number versus another number whether you want to or not there's all of these invisible influences that are guiding us and so the question becomes what, how do you ethically use those forces and then again for good and for what values.. In terms of being concrete what do you do, you can turn off all notifications on your phone except for when a person wants your attention.. So one of the big things in the industry is just about everything that's a notification that comes at you is actually generated by a machine and the machine's goal is usually like what will get your attention right now.. And so it's automatedly sending you stuff..

Have any of you ever tried the following, you don't use Facebook for like a week and then you watch and suddenly they send you like 3,000 emails.. It's like a drug dealer who wants you to come back.. (laughing) So, I'm not trying to be, I hope when I say this I don't sound like I'm just trying to be against something.. It's more that there's this model that's just not aligned with all of us, right.. I don't think anyone wants a world where this is going.. That's hopefully what I communicated.. Like this is so dangerous where we're headed.. And I want to say one more thing about danger that I didn't mention.. I think the deepest existential risk from this is that the thing that's best at capturing a human being's attention, a single person's attention, is going to be show them an individual reality that confirms their world views.. That's going to be so different from the thing that would confirm our shared sense of reality..

So you can think of Facebook innocually.. Like you give it this innocuous naive goal of like let's help people, let's engage people the most and then that thing is taking entire societies and putting them through a paper shredder where out the other end you get basically filter bubbles and echo chambers where people don't have shared facts and agree on the same reality.. That's like I think the deepest part of this problem.. So if you weren't convinced by the kid's threat or the national security side I think that is the sort of step one for good society.. You.. - [Audience Member] So from a design ethicist perspective I was wondering if you could give an example of one of these persuasive techniques and how do you begin to apply ethical framework to it given the scale at which it's purported.. - Great.. - [Audience Member] (mumbles) - Okay so the question was if you take a persuasive technique how would you apply an ethical framework for deploying it.. Okay so let's take, I love this example, so lets take streaks in Snapchat.. I'm sorry if you've heard me talk about this, but it's really important because if you didn't know Snapchat's the number, how many of the students here use Snapchat as your primary? Okay, a good number, okay, and it's obviously invented by a Stanford alum who has good intentions, but they use this feature which is a technique, a persuasive design technique called streaks..

If you are older in the audience it shows the number of days in a row that you've sent a message back and forth with every contact.. So why is that, is that ethical or not? I want to actually hear, what do you think? If I put the number of streaks in a row, the number of days in a row that two kids have sent a message back and forth what makes that, is that wrong, it that good, is that bad, why? - [Audience Member] I think of like from a utilitarian perspective it adds, it increases your quality of life and helps you feel more connected with your friends.. - Okay so if it increases your quality of life and helps you feel more connected with your friends then that would be one reason why we would say it's good.. Let me tell you actually what the people at Snapchat when they made this feature what I've heard through the grapevine is the way that they justified this was that unlike these other social apps that connect you to all of these friends all at once, you have like all of your relationships all at once, they help you focus on the relationships that really matter because they show you the depth of your friendships.. (laughing) Do you see what I'm talking about? Again you have a belief and a narrative and you're walking, you're stepping inside of this belief where everything you see through these rose colored glasses, where this is deepening friendships.. According to the human being that believes those thoughts, they're inside of a magic trick, and they believe that this is actually about deepening friendships.. And there's definitely evidence for that.. I'm sure that there's someone somewhere two kids who feel like man my friendship is way deeper because I can see the number of days in a row I've sent a message back and forth.. There's someone who's experienced that.. Just like there's magic moments on all these other applications, but there's still something that's not quite right about it and I'll just jump to the chase..

With Snap streaks there's first of all an asymmetric situation so the children don't know that there's like a thousand or a hundred engineers on the other side of the screen and that they're deliberately choosing this technique against them, so that's one thing.. So there's an asymmetry between the knowledge that the persuaders have and the knowledge the persuadee

has.. Okay, the second one is that the, just had it, the goals of the persuader are different than the goals of the persuadee.. So when the people that make Snapchat do this their actual goal is how can I hook you to use the product everyday, cause that works really well.. Like the streaks feature is super addictive.. That's their goal and the goal of the person who's using it, usually the challenge of ethical persuasion is people don't know their own goals so they actually just sit there and then the persuader's goal infects the persuadee.. So now the persuadee is like the matrix, there's a drilled hole in the back and this goal went in and like now I need to keep up with these streaks and now they actually want that.. That's something that they intuitively, independently want.. That's successful advertising.. The persuader's goal has become the persuadee's goal and now they define their friendship based on whether or not they're able to keep up their streaks and if they don't have their streaks they're no longer best friends..

And by the way that's actually true.. So there are children walking around who think that their you know the terms of their friendship are the streaks that they have.. I want to just do one more thing here.. It's a great example, I'm so happy you asked me that.. Is the, you can do the same thing, a streak feature on a meditation app.. Same feature, now we're putting on a meditation app, the number of days in a row you've meditated.. So why is it feeling, like intuitively it feels better there, right? So why is that and I'll just say quickly it's because theoretically the person who's trying to meditate actually cares.. Their goal is aligned with the persuader's goal.. They actually want to sort of mark off the number of days in a row that they've done the thing that they're yearning for.. Whereas in the first case that's not true..

There's a whole field here we should write like a 600 page book about, but that's a great question.. Yeah.. - [Audience Member] So you were recently at (mumbles) and they were talking about YouTube and everything, how YouTube's changed.. I'm spending so much more time, but I see all of the cute videos.. My experience on YouTube is more frustrating now.. - More frustrating for more time.. - [Audience Member] Because I used to look at it like the most (mumbles) such an interesting thing so news and viral videos and other things, but now it's optimized to take most of my time.. The response that I got was yes, but you're seeing more advertisements, spending more time on site, so essentially we're doing our job.. - Right.. - [Audience Member] But I now dislike YouTube when I used to like..

- Right.. - [Audience Member] So is there a way for us to capture our actual attitudes towards the service being provided versus just the time we spend on site and to think there's a rule for pointing to the flaws of measuring interest on ads that are clicked through.. - Yes.. - [Audience Member] I'm very hesitant to buy anything that's-- - Sure.. - [Audience Member] But I'm not sure that that's an accomplishment.. - Well the challenge of advertising is they make money whether or not you buy it or not because they still make money from just the impressions for seeding the ideas into your mind, that's still a success case for advertisers.. I know that we're wrapping up.. - [Woman] No, no, finish please.. - But the challenge here is there are the metrics.. My first Ted talk went through some of them..

The challenge is that any business model, any business or company, technology company, whose business is to capture attention, they just have to do that.. Like there's just no other choice.. YouTube just has to do that.. But if you look at Vimeo as a counterexample they don't auto play the next video, they don't show you the related videos and try to get you into an infinite world and it's because their business model's a little bit different.. And so you can imagine a version where you pay where these choice architectures are not trying to bottomless bowl you into infinite watch time.. So, yeah.. (applauding)..