**Your Mind on Changing Prejudices & Stereotypes, with Dr. Tessa Charlesworth**

[music]

0:00:04.7 Ava Ma De Sousa: Welcome to Minds Matter, a podcast sponsored by the Monash Centre for Consciousness and Contemplative Studies. I'm Ava.

0:00:10.6 Beth Fisher: And I'm Beth. And on Minds Matter, we explore research from neuroscience to psychology whilst talking about our own personal experiences.

0:00:18.7 Ava Ma De Sousa: This week on the podcast, I spoke to Dr. Tessa Charlesworth, who's a Postdoctoral Researcher at Harvard University and the University of Toronto. She's also an incoming Assistant Professor at the Kellogg School of Management. On this episode, we talked about how prejudices and attitudes change over time. Tessa shared research on two specific methods that she's used to investigate this change. First, she uses a task called the Implicit Association Task, which we referred to in the podcast as the IAT. This allows researchers to get a gauge of people's implicit prejudices or the ones that we're not necessarily consciously aware of. And this task is one that is arguably harder to control your responses to than say, being asked how prejudiced you are in a questionnaire.

0:00:57.6 Ava Ma De Sousa: What's really cool about this is that she uses a database called Project Implicit that has 20 million respondents and has been collecting data for almost 20 years, so she can really examine attitude change on a really fine grain level. We also talk about how prejudices are embedded into broader cultural products. The research that Tessa does for this uses a method called Word Embeddings and Natural Language Processing, which is borrowed from computer science and artificial intelligence, it basically creates a mathematical representation of words, which are also called vectors. In non-technical words, this just allows her to examine what words and traits are associated with different racial groups, for example, over time in big cultural products such as the books that we read.

[music]

0:01:41.6 Dr. Tessa Charlesworth: Hi. I'm Tess Charlesworth. I'm currently a Postdoctoral Researcher at Harvard University and the University of Toronto. And I'll soon be starting as a faculty member at Kellogg School of Management. And in terms of my research, I'm interested in how and why we change our attitudes and our stereotypes about social groups. And the keyword there really is change. Why is it that some of our biases and our attitudes and our representations move us towards more acceptance of some people over history, whereas other kinds of biases actually become less tolerant and even hostile over history? What is it about our minds and the actual mental representations, but also equally important, what is it about the kind of social structures and the settings across history that can give rise to more or less change over time?

0:02:30.6 Ava Ma De Sousa: One of the really cool things about your work is that you move away from the traditional psychology type of research and analysis. Typically when we think psychology, we think getting people into the lab, maybe having a confederate to an actor, messing with people and then seeing what happens, or maybe running some online studies where you have a specific question and a specific sample that you're using to do that. But in your work, you really don't do much of that. And one of the first things that I wanted to talk about is work that you do with a constantly open website where people are able to just take tests for fun, so I was just wondering if you could talk about Project Implicit and also discuss the main test that is administered on Project Implicit, which is the Implicit Association Test.

0:03:12.9 Dr. Tessa Charlesworth: Yeah. Project Implicit is a non-profit that was founded by Mahzarin Banaji, my advisor, as well as her student, Brian Nosek and her advisor, Tony Greenwald. And together they had developed this test called the Implicit Association Test, which as some of your listeners will know, is basically a [0:03:30.6] \_\_\_\_ time task. The easiest way I think about the IAT, the idea behind the IAT is when it comes to sorting a deck of cards. It's very easy to sort a deck of cards based on color, that's just automatic, easy, fast, congruent associations that you can do. It's a lot harder if I ask you to say sort clubs and hearts to one side, and diamonds and spades to the other side, that's less congruent and a little bit harder for you to do by just relying on automatic reactions. The IAT is actually a very similar kind of task, we basically have you sort social groups and attributes, good and bad, for instance. And a congruent task might be old and bad to one side and young and good to another side. And an incongruent task, the one we'll do a comparison with on your speed, would be something like good and old and bad and young. And we can use the difference in the reaction time to see how automatically biased you are about those two kinds of associations.

0:04:30.2 Dr. Tessa Charlesworth: They developed this task back in 1998, and immediately thought, "I wonder if we could just put it online for some demonstrations for other educators, other researchers to use," having no idea that it would really break the Internet and blow up in the way that it has. Because a couple of years ago when I checked, we had 20 million people who had taken the IAT through this website. And it's since exploded, this was also pre-2020, it's since exploded throughout the pandemic following the murder of George Floyd. I'm assuming we probably have at least 30 million people at this point who have come on to this website and taken these tasks. And the cool thing about the website is because it was put on shortly after the development of the test, it was put on in the early 2000s, we have really rich time series data, really rich data on these kinds of automatic attitudes that have been collected now over two decades. That's basically where, knowing about this data was where my first projects in grad school really started. It was actually a staff midterm project where we had to use some sort of existing dataset, and all of my existing datasets at the time were 20 kids, I was trained as a developmental psychologist, so it was 20 children and we couldn't do really fancy methods to them or anything like that, it was basically a t-test or a chi-square test.

0:05:45.1 Dr. Tessa Charlesworth: And so knowing that this dataset existed, I thought "Why not play around and see whether there are any interesting trends that have happened over this long-term time span of 20 years?" And the first trends that we started to see, my advisor Mahzarin Banaji didn't believe them. I was showing this drop in implicit racial attitudes over time, and she was like, "That can't be right, we know that implicit attitudes do not change over time. Implicit attitudes, these kinds of tests that we get through the IAT are thought to be super stable, they're called cognitive monsters, their automatic habits, all these things they can't change over time." And so we went back and we tried another task, sexuality. And sexual orientation biases have also dropped, by even greater degree, at this stage now up to the end of 2020, by 78% drop. And so it was, already we were looking from the staff's midterm project of seeing, "Huh, it turns out these implicit biases, when we're thinking about them as this really long historical time scale over this much more aggregated level of analysis across our population, maybe they could change in a very different way from how exactly as you described, how we might typically think of attitude change in an experiment in the lab, one or two time points for a single individual."

0:07:04.5 Dr. Tessa Charlesworth: And then of course, the last question that we asked is, "Okay. Is this the case that implicit attitudes are always gonna change or is this something just unique about maybe race and sexual orientation, which are unique in many ways in our society?" There are attitudes that we talk about a lot, there are attitudes that we've prioritised through protests, through legislation, through media representation. And so we tested a few other attitudes, including implicit biases about age, disability and body weight. And those three tasks or those three tests, to see biases have not changed on our implicit attitudes. That, again, told us a little bit more about the likely sources of where those changes had been coming from, things like those unique features of race and sexual orientation relating to the kinds of protests and media representations that aren't going to extend to every single group in our society. It's not that we have this massive wave of just societies becoming nicer over time in society, it no longer has implicit biases, it's really only for some groups and only in some ways. That was a long-winded answer, but that's Project Implicit and the first papers that have come out with that data looking at change in a much more macro level approach to understanding.

0:08:22.0 Ava Ma De Sousa: Have you compared the implicit attitudes to explicit attitudes? For the things that didn't change in the implicit attitudes like ageism and stuff like that, are there changes in explicit attitudes or is it also reflecting this thing that we don't really talk about it so people don't maybe care about it as much anyway?

0:08:38.3 Dr. Tessa Charlesworth: Explicit attitudes have been changing for all three of those topics of age, disability and body weight. And so, I mentioned something about this, it's not the case that we have this massive wave of society becoming less biased, actually, we do on explicit measures. When it comes to explicit biases, every single topic we've looked at has dropped in bias over time. And that, again, really is pointing to the source of those explicit bias changes. That is, when it comes to explicit bias change, it does seem to be this wave of what we might call a wave of social desirability concerns. A wave in which people are becoming more and more concerned with saying things like, "I don't like old people," or "I don't like people with disabilities" or whatnot. And they're dropping in those biases over time. But that's not the case for implicit. There isn't a same wave that's affecting our implicit biases because those tests aren't as vulnerable to self-presentation and social desirability concerns. We can actually see a really interesting dissociation in the likely sources of these changes, depending on which measure we're looking at.

0:09:47.4 Ava Ma De Sousa: I guess this is kind of hopeful, the fact that we've been really trying to change some of these attitudes, there have been a lot of protests. It still feels like, in body weight and body shaming, there's a lot of talk about that, I guess, at least online. I guess I'm wondering, if you could pinpoint exactly what it is about certain types of movements. Is it just that they're more in the discourse?

0:10:13.7 Dr. Tessa Charlesworth: Yeah, it's a great question, and one, I'm excited to continue studying at Kellogg because there are really amazing social movement scholars there. I'm sure they'll... Collaborating with sociologists and people who specifically study social movements, I think, will be a great path forward to identify what are those features that characterised the same-sex marriage movement and Black Lives Matter movement that weren't present in the body positivity movement, for instance, or in the disability rights movement as well? Yes, I think when it comes to social movements, that clearly can't be the whole story because you're exactly right, there are parallel social movements that have happened for these other identities, although to a lesser degree, there hasn't been the same kind of movements that we see with pride protests or Black Lives Matter movements, for instance. There is another factor that does distinguish these groups though, that could be another key ingredients that we might need, and that is that those three groups, age, disability and body weight, are all stigmas are all social groups that are physical, and they're read on the body in a really particular way that we maybe don't think about as much with sexuality and race, which are much more socially constructing, much more in the social milieu.

0:11:25.0 Dr. Tessa Charlesworth: And there are really interesting theories from the Sigma literature, which comes more from public health, about how those different... That distinction between more body-related stigmas and more socio-demographic-related stigmas might serve different functions in our society. Socio-demographic stigmas like race, for instance, often serves a function of keeping people down and getting our resources and resource domination. Whereas a body-related stigma like age and disability are thought to reflect functions that are more about pathogen avoidance or keeping people away that might have something that's harmful to our health or harmful to our ideas of well-being. And so, those different functions of stigma might make them differentially malleable when it comes to change over time. It might be really hard to change a stigma that we think is fundamentally tied into health and well-being, versus a stigma like sexuality and race that we might think is more tied to resources, or dominance or power or something like that. That's an open empirical question, so I don't have the answer, but it is another access that divides those groups above and beyond just protests.

0:12:45.3 Ava Ma De Sousa: I'm gonna ask you to speculate on this...

0:12:46.5 Dr. Tessa Charlesworth: Yeah.

0:12:48.2 Ava Ma De Sousa: A little more. I'm just thinking about the differences between groups of sexuality, race versus something like disability or weight. I feel like with disability and weight, those are things that could happen to anyone, whereas race or sexuality, although they're socially constructed in a sense, they're not really changeable. I was just wondering, do you think that plays into it as well, that maybe there's more of a fear attached to it for people and that leads to more stigma that is more movable?

0:13:16.7 Dr. Tessa Charlesworth: Yes, I think absolutely. It's another big factor. And there are methods that we're trying to do now to test these hypotheses, and we'll hopefully talk about them a little bit later. But right now, the reason why we can't perfectly quantify all the factors that might contribute to some groups changing and others not, so protests, whether it's body-related, whether it's changeable or controllable, whether there's a risk that you might assume that identity and so you wanna push it away and avoid it. All of those features are things that we can speculate and do back-of-the-envelope calculations on across the six groups that we've looked at so far. But what we would need to actually quantitatively test that, is a much bigger sample size of groups that vary along those dimensions to basically give us the actual variance that we would need and change to then explain based on different features. Obviously, we can't do that with the kind of archival attitude data, because we just can't go back and tell researchers from the 2000s, "It would be great if you collected some data on wheelchair-bound stigma," or "It'd be great if you collected some data on smoking stigma."

0:14:24.7 Dr. Tessa Charlesworth: Unfortunately, we don't have that, but what we do have is text and language. One of the new methods that we've been developing is "How do we extract these same kinds of stereotypes and representations and biases through language?" Because that just opens the door to being able to now study as many groups as we want across, really, as long as it has been as we have language for it. And we've just started to do that and to actually map out the space of how these different groups are changing. And then what we can do now is look at some of those correlates, how are groups rated on the perceived change ability or the perceived risk of their identity? How are they rated in terms of perceived threat or perceived peril, and then see whether that predicts the degree to which they change our time?

0:15:07.9 Ava Ma De Sousa: I don't know if you can do this with the data that you have, but because you're tracking these changes over time, and there are certain big changes that happen societally or that just happen in the news, for example, that can both seemingly push forward social movements, but also bring them back. So I'm thinking, obviously, the election of Barack Obama, there's also, during the time of your data collection, the Supreme Court ruling that same-sex marriage should be legalised. On the flip side, there's been presidents in the past that have said things that are maybe a little bit counter to the feminist movement, for example, or very disparaging things about migrants and things like that. So I was just wondering, do you see that reflected in the data? And is there differences between how positive things that might decrease bias or legislation that might decrease bias influence the data versus negative things that might pull things back?

0:16:01.1 Dr. Tessa Charlesworth: Yeah, those are extremely rich and interesting questions that, again, can't wait to continue exploring. We do have our group and some other researchers have, done one-off tests for individual attitudes to test some of those ideas. Eugene Ofosu and Eric Hehman have a great paper on same-sex marriage legislation, and they find that implicit sexuality biases really do show an inflection point following that legislation, at the federal level, but also earlier at the state level. If you were in a state, say Massachusetts, that legalised in 2005, you actually get a little bit of an inflection point there already, so you start decreasing in sexuality bias already in 2005, and then it just gets another inflection point in 2016. What's interesting about that is we're seeing the positive snowballing influence, if you will, of both more local and more federal legislation. There is some evidence that you can have external impacts that create positive change in attitudes. However, there is also data from our own group showing that you can have external impacts that create negative changes in attitudes. The kinds of rhetoric that started with Trump's early presidency in the Republican primaries about disability as well as about body weight and about race, we see actually sharp increases for about a year immediately following that kind of change in rhetoric in the public consciousness.

0:17:33.2 Dr. Tessa Charlesworth: And we specifically see it in those groups that Trump targeted. Trump rarely targeted LGB individuals, lesbian, gay and bisexual individuals. And so we don't really see any increase. In fact, anything we see a slightly sharper drop in sexuality bias, perhaps because of other phenomenally transfer of prejudice and those kinds of things. But for the groups that Trump did target, disability, when he targeted the New York Times reporter and mocked him body weight, when he shamed public figures about being overweight, when he called Miss America a fat pig, those kinds of things. Those shaming events do seem to align with these particular increases and particularly amongst conservative states, in particular among conservative respondents, which tell us that it really is something about this kind of emboldening of rhetoric or this emboldening of prejudice amongst those particular groups that might be most exposed to those structural level or rhetoric level changes, that do create this negative impact on our biases. Now, I will say, you did have such a... The question really is, how do those two impacts a positive change or a negative change.

0:18:48.0 Dr. Tessa Charlesworth: How do they differ from one another? And for me, a really interesting question is, how long does a positive change last and how long does a negative change last? In psychology, there are tons of theories about the unique power of negativity, or it's so much easier to make someone think negative thoughts than it is to make them think positive thoughts, for instance. So it becomes an open question at the more societal or macro level of our attitudes, whether that is also the case, whether it's that much easier to quickly create negative change and push people away from progress than it is to motivate them to become more accepting of a group. And so I think that's a super interesting and ripe area for future work.

0:19:30.9 Ava Ma De Sousa: I just wanted to ask, because what you said about the implicit changes after something like Donald Trump saying something horrible, I guess that's in a sense surprising because the rhetoric that we usually hear around what happened during the Trump presidency is like, "Oh, this was all here, implicitly, and now people are able to express it explicitly," but with your data, it's like... It's every level. Do you have data on whether explicitly there was the same pattern?

0:20:02.9 Dr. Tessa Charlesworth: No. Interestingly, so we do have data, but there were no explicit changes during Trump's presidency, so the only place we're seeing these kinds of bumps in bias are on implicit.

0:20:17.6 Ava Ma De Sousa: Wow.

0:20:18.2 Dr. Tessa Charlesworth: And so the story that we might tell there from a more psychological perspective is that, on a population level, we're probably not seeing that much of this emboldening effect of people on average... Yes, it is true that there is an emboldening effect, especially for more extreme, right-wing people, they didn't exist and didn't have the kind of public endorsement that they do now. However, when it comes to the average American or the average US resident, it seems to be that actually, our explicit biases don't change that much. We're equally reticent to say negative things about another group during Trump's presidency. However, we are continuously... We are more exposed during Trump's presidency to these kinds of reminders about the negative association between, say, being overweight and fatness, because we're listening to the rhetoric, because we're more exposed to these right-wing commentators, because we're seeing these kinds of negative media representations that just weren't allowed from the extreme anymore. And so because the average American is now exposed to these new associations just hanging out in our daily lives, that's gonna shift our implicit attitudes, even if it doesn't shift our explicit attitudes.

0:21:31.6 Ava Ma De Sousa: So there's been some criticism of the IAT in general, and particularly one that I find interesting but that I think actually could embolden your work is the criticism that the IAT actually doesn't necessarily measure an individual's prejudice, but measures cultural prejudice, which I think goes along really well with what you just said about, "Oh, you're in-taking that information more and you just have that implicit association more quickly," so I guess what exactly are you measuring with the IAT? Are you measuring cultural... Just norms and prejudice, or are you measuring changes in the population over time? And how different are those two things?

0:22:09.9 Dr. Tessa Charlesworth: Yeah, it's a great question as well. What I think about when I think about the IAT and especially what we're measuring in our work when we aggregate across people, is really this kind of average thumb print of the culture on the average mind. So it really is some combination between individual influences and cultural influences. And another kind of analogy that I often think about when thinking about the IAT is sort of like a lens where an individual will each bring their own prescription, their own like visual prescription, their own visual lens to filtering the information from their culture. So they're all seeing the same scene, and they're all gonna be influenced by that same scene, but there are also some individual differences in the degree to which you incorporate and endorse that information.

0:22:57.2 Dr. Tessa Charlesworth: And so what becomes a really interesting empirical question going forward, and I think we're moving in that direction now that Keith Payne's Bias of Crowds model has been proposed, is to really try and partial out how much and where and when are there these kinds of more individual endorsed components of the IAT, and how much and when and and where are these... There are these more cultural components. When I think about the IAT and what it's measuring, it's definitely some combination between the two, and it becomes really interesting to think about how to empirically pull those two apart. Your question though about sort of like, is it cultural norms or is it other changes in the population, again, it's another empirical question, and I think the best way to test those is to look at other indicators of cultural norms or other indicators of media and expressions and representation and see how much our measures of IAT-related change or attitude related change are tracking with those societal level indicators.

0:23:56.8 Dr. Tessa Charlesworth: And if they're exactly the same and they're perfectly correlated, then we can say there's a really strong linkage, it's probably picking up similar... It's probably picking up almost entirely those kinds of cultural changes and exposure or cultural events, but if there's at least some dissociation, then we might have to ask, "Okay, what else could the IAT be picking up above and beyond things like media representation changes?" If it's with a more moderate correlation and controlling for measurement error, we might ask, "Okay, could it be something else like changes in the demographics of the population? Is this just young people replacing old people, for instance?" We wanna partial that out next, and then we ask, "Okay, cool. Could this be individual people changing their minds?" And then we partial that out.

0:24:43.9 Dr. Tessa Charlesworth: So I think it really, again, it's a long-winded answer because it's such a rich question that essentially gets to, why is change happening and what is the change that you're seeing? And for me, that's basically my whole research future is partial-ing out those various pieces that are contributing to this one indicator of change over time.

0:25:04.0 Ava Ma De Sousa: So if people go on to Project Implicit and do one of these IATs and it shows that they're pretty biased, what should they do? Do you think that's an indicator like, "Hmm, something's going on here. Maybe I should question this a bit." Or can they say, "Oh, I guess the culture is just kind of influencing me here."

0:25:21.9 Dr. Tessa Charlesworth: Yes, I think if you go on to Project Implicit and you get the likely result, because 60-something percent of people will get a result that shows that they have some degree of bias, and it says, "You moderately associate Black Americans with bad and White Americans with good," for instance. I think you can make two conclusions; one, re-evaluate your diet of media and of culture, and see what are the things that you might be inadvertently reinforcing for yourself through the choices that you make and the kinds of exposures that you have. So you might choose to re-evaluate your television-watching habits of, why don't you watch some black documentaries? Why don't you watch some cool Dear Black People, like comedies and stuff like that, as well as what you're reading, as well as the kinds of friends you're exposed to and those kinds of things. So you can make a cultural attribution, but that also doesn't mean that you're off the hook, that you can say, "Well, it just means that the culture needs to change," because ultimately, you are a consumer of that culture and you can look to different kinds of consumption patterns.

0:26:25.0 Dr. Tessa Charlesworth: And then I think above and beyond that, you can also do a second thing, which is to evaluate how you yourself might be filtering that information. So, say, even if you have this perfect balanced exposure across groups and you're really intentionally trying to watch Black documentaries, continue to evaluate yourself and audit yourself in a way of like, when you're watching those documentaries, are you discounting the stories that you're seeing in ways like saying, "Oh well, yeah, that's just a one-off case. Barack Obama was just a really unique Black American, he's not representative of the broader community." So what are the kinds of cognitive habits and cognitive tricks that you are doing so that you're filtering that information in a way that might align with your biases? So again, it's kind of... Like I said, the IAT, because it can pick up on both that lens or that filtering as well as the culture itself, when you get a score that says something about the degree of bias you have, it's really an opportunity to look at both sources, the individual and the culture.

[music]

0:27:40.1 Beth Fisher: So when I talk about this stuff, this is way more Ava's area of science than mine, so I'm a bit new to it. But one of the things I was thinking about when I was listening to Tess' research about these implicit biases we have, so the biases we're not aware that we have, how that then plays out in discrimination, because I guess there's discrimination that people are aware of doing. I don't know if you don't support gay marriage, that's a clear, "I don't support gay marriage, I don't support gay people," and it's explicit that that's what they believe, but I was thinking, well, how does this implicit kind of, I guess, discrimination would play out and maybe in these ways that I'm not so clear. So maybe when also understanding in, I don't know, workplaces or schools, how that is for someone who's being discriminated against, but the people don't think they are 'cause they're not aware they have this bias, what that may look like. And maybe this is one of the complications when it's implicit, 'cause then how do you name if someone's doing something but they're so unaware of it 'cause it's implicit, how do they then learn from that or how do you name that? I don't know if you have any thoughts on that, Ava, or any examples, or 'cause it's your area, any studies that look at how that plays out.

0:29:00.0 Ava Ma De Sousa: Well, I think... And just for the record, Beth and I just discuss these. We come in hot live and we don't really talk about what we're gonna talk about, so this is just off the top, but I think one thing that's important to keep in mind about biases and that really changed the way that at least I thought about biases was the fact that, honestly, no matter what we do, the most anti-racist person, whether you're Black, whether you're Asian, whether you're White, whether you're Middle Eastern, you're going to be biased. Because as Tessa talks about in her research, we just absorb a ton of bias from our environment. So it's in everything that we read, it's in the way that parents interact with their kids, it's probably in the way that your parents interacted with you when you were a baby.

0:29:44.0 Ava Ma De Sousa: And I think that for me, some of the most interesting research has been about just kind of realising that you're biased, which can happen, let's say if... As Tess and I talked about, if you take an IAT and you see you're super biased against LGBTQ people or something, you might wanna think about why that is, and even if it's coming out in this sort of implicit way, you might wanna be able to just kind of re-frame, as she was saying, the media diet that you're consuming. But there is some cool social neuroscience research that looks at the ways that people can control biased impulses, and these are things that happen to me as well. Sometimes, I'll catch myself automatically saying "He" when I'm talking about someone's doctor, and they'll be like, "Actually, it was a woman," and I'm like, "Oh my gosh, look at me."

0:30:28.5 Ava Ma De Sousa: So that happens to all of us, even Beth and I who are in STEM, but so when you realise that, there's a few things that can happen, and there's three kind of main parts that people talk about in terms of what goes into realising that there's a bias that needs to be regulated. So first, our brands have to detect that there's a mismatch between our motivations to act in an egalitarian way and biased behavior or just biased thoughts, and the first step is implemented by a brain region called dorsal anterior cingulate, but then when you stop the biased response, that requires another region called the inferior frontal gyrus. And then after that, after you've actually stopped the response, so our first part was just realising that we're making a biased response, stopping that response and then actually replacing our response with an egalitarian one, that's linked to another region of the brain in the prefrontal cortex, but it's the left dorsolateral prefrontal cortex, so more like the top of the brain. And so knowing these processes and breaking it down, at least for me, when I read this research in terms of almost the computations that your brain is doing...

0:31:33.2 Ava Ma De Sousa: Realising it, stopping it, replacing it, I think knowing those brain regions that are involved helped me be like, "Okay, those are things that I would need to do almost like therapy for myself," being like, "These are the things that we need to do to change these things." And so I think over 2020 and the rise and unfortunately fall of BLM, I think we realised that there's a lot of discrimination at least... Depending on how you took all of that, I know some people didn't really take that to heart and still argue that that is not real, but I think most of us do realise that there's a lot of discrimination that happens by accident in some ways, but I think when we call it implicit bias and when we call it by accident, that allows us to pretend that it is kind of like a feature of how we are, that it's just taking in this statistical information, so we're kind of almost being rational in it, and I think that that takes away responsibility. So I think research that shows the way that we can actually stop those biases really helps, for me, realise it's okay to have biases, but you just have to continuously work on not letting that control your behavior or the way you're interacting with people.

0:32:37.5 Beth Fisher: Yeah, so I did this thing recently, it's a mental health first aid training that we have in Australia, and I'm obsessed with it, and I'm telling everyone to do it. It's basically, rather than physical first aid, it's Mental Health First Aid, so you can help people if they're having certain issues and it's amazing. And I was telling all my friends about it, and one of my friends went and did it, and one of the exercises you do in the course is you have a list of mental health conditions, so like post-traumatic stress, depression, schizophrenia, all these different things, and you also have a list of conditions like cancer, diabetes. And they get all the research and you have to put them on match... So how debilitating are these? How much do they impair someone's ability to live a healthy life? And you have to match the physical condition with the mental health condition. So just say, for example, if you have depression, does it impair your ability to live a good life the same way having diabetes would, and is that the same level? And it's a really good exercise because it makes you realise how much if you have a mental health problem, that really... Even though physically, of course, we don't see it, it really does stop people being able to do things that they would wanna do.

0:33:44.8 Beth Fisher: And I was speaking to my friend about this, and she said... Obviously, she never has consciously been like, "People with mental health are fine and they're making it up and they can all be okay." And she said the minute she did this task, she really was aware of like, "Oh wow, I was really underestimating without realising it, the severity of these things for people," but she said she would have never identified with doing that because she is someone who cares and all of these things, but it wasn't until she did this exercise that she realised, "Oh wait, even though I was thinking I was understanding how these conditions affected people, I wasn't at all." Because when she came across that data and had to do this task, she had to completely update the way she was perceiving this. But she said, I guess that's the problem when it's an unconscious bias, [chuckle] you're unaware of it, but yeah. So I wonder if there's exercises like that you can do where it becomes very clear, like, "Oh, this is how I'm misperceiving this," I guess.

[music]

0:34:52.0 Ava Ma De Sousa: So one of the things that you talked about in your answers to the Project Implicit stuff was this idea of the broader culture that we're absorbing, and you mentioned already that you've been looking at the way that our text and different corpora that we take in can also tell us what's going on in the culture. So this is a technique that's called natural language processing that's been used more in computer science, that now some psychologists like you who are interested in cool methods have now been taking advantage of. So could you just talk a little bit more about what NLP is, or Natural Language Processing, and what word embeddings are and how you've been using that in your research.

0:35:30.2 Dr. Tessa Charlesworth: So these are tools, as you mentioned, developed in computer science that we're really trying to figure out, how do we represent word meanings? So, like a dictionary definition of a word using only numbers. So how do we go beyond the qualitative to make it quantitative? And the key insight was actually in the mid-1900s with this idea that any word meaning could be represented based on how a word co-occurs with other words. So the fact that the word "tree" often occurs with the word "leaf" tells us something about the word's meaning; tree is somehow related to the meaning of the word "leaf," for instance. So the key insight for social psychologists like me to apply some of those methods came in 2016 when now my collaborator, Aylin Caliskan developed this tool called the Word Embeddings Association Test, which essentially took those quantitative vectors of word meanings, so every word has this long 300-length vectors of numbers, and said, "Can we look at the relative association between words used to describe groups like old and young, and words used to describe attributes like good and bad," for instance. So it's exactly parallel in idea to the IAT, the Implicit Association Test.

0:36:45.0 Dr. Tessa Charlesworth: But here, all we're doing is instead of looking at reaction times, we're looking at the kind of quantitative vector representations of those words. And what was really cool from that paper that Aylin published in science was that the same magnitudes and the same presence of biases about race, about age, about gender that we had documented through Project Implicit Data appeared in our large-scale internet language and appeared in these massive corpora of texts, 840 million words. And that was sort of the first validation door-opening project that basically set all of our wheels in motion in this lab, looking at where we are very familiar with the IAT, to be like, "Oh my goodness, we need to start studying language as well, 'cause it just completely expands our toolkit. And so some of the first projects that we ran were, again, in the sense of validation studies, of, can we map on the biases that we know we should find in language as well? And for me, again, trained as a developmentalist, so I was intimately familiar with the kind of developmental work on gender stereotypes and how early those develop amongst children and how they are really deeply embedded in children's culture.

0:38:03.0 Dr. Tessa Charlesworth: So children's books, children's TV shows, children's conversations. And so as our first study using these kinds of methods, we wanted to look at how these biases that we could reveal in language might map on to the biases that we know exist amongst children. And so what we did is we basically took all of these massive corpora of child-directed language, so children books, children TV shows and children conversations with their parents, as well as the adult equivalent, and then assessed whether we could find gender stereotypes in those naturalistic languages.

0:38:37.7 Dr. Tessa Charlesworth: So we looked at whether there was a stronger association of women to home and men to work, and women to arts, and men to science, women to good, and men to bad, and then women to reading and men to math, so these are kind of like four very common stereotypes and attitudes. It turns out that in all of these corpora towards children, so all of these tech sources that children were consuming, we find evidence of these gender stereotypes, and so that was actually super surprising, super interesting in the sense of being able to replicate known psychological effects and known sociological effects with these large-scale language models, and now it's like once we had that, once we had additional confidence since we had additional skills in applying those methods, it then again, just opened the door to a vast number of inquiries going forward.

0:39:33.3 Ava Ma De Sousa: So in that project that was specifically looking only at children's corpora?

0:39:37.9 Dr. Tessa Charlesworth: Yeah. We had both. So we had children's corpora, and then for comparison, we also had adult-directed corpora, so we had Nickelodeon and PBS TV shows, but then also shows like CSI Miami and CSI LA and all of those things for adults as well, and in general, we found some differences when it came to trait level stereotype, so the degree to which an adult corpus might reveal an association of women to friendly, for instance, that was often stronger and more elaborated in the adult corpora, but when it came to those four big domains of gender stereotypes of math versus science and reading versus or math versus reading and arts versus science, those were consistent across the adult and the child corpora.

0:40:28.1 Ava Ma De Sousa: That's disturbing. [laughter]

0:40:29.4 Dr. Tessa Charlesworth: It is, it is. Especially when you think about... One of the corpora that we looked at was this corpus called the child's language corpus, which is a massive corpora of child and parent conversation, so naturalistic conversations that are recorded in a home, often of just things like what children and parents are talking about when they're brushing their teeth or making dinner together or whatever, those are...

0:40:57.2 Dr. Tessa Charlesworth: The mean age of children in those corpora, is two years of age, and we found biases in the kind of adult language, parent language for those corpora at that time, the mean age of children in those corpora is two years of age, so really, really young kids, their parents speaking to them at two years of age already contained the kinds of biases of associating say men with science and women with arts, so it was... It's in line with what we've already seen in the literature in terms of this early prevalence of gender stereotypes, it was just showing it at a scale that we didn't really realise was possible because it's now in this kind of real world context of naturalistic conversations. Naturalistic TV shows, millions of these, even there, when we expand the scope, beyond the lab, we're seeing the prevalence of these biases.

0:41:53.3 Ava Ma De Sousa: Yeah, that's really scary. I remember there's also, I forget the researcher who does it, but there's this researcher, I think a few decades ago, who took videos of people interacting with babies, they were told that it was a boy baby or a girl baby, and every time they'd hold the boy, they'd be like, "You're so strong." And for the girl, "You're so pretty." So I guess it's not surprising, but still, still disturbing. [laughter]

0:42:16.9 Dr. Tessa Charlesworth: Nonetheless disturbing. Exactly, and the fact in that study, we didn't have tons of power to be able to look at change over time, but the fact that we were able to see the book corpu... The children's book corpus were from the early 1900s. So it's kind of understandable that books would reveal these kinds of gender stereotypes, but the TV show corpus, that was from the 2000s, and you kind of have this hope that a lot of television shows might have kind of modernised in a way that they represent women over time and gender roles over time, but the fact that those were showing slightly weakened prevalence of stereotypes, but nonetheless, still present in the stereotypes and still present in the kind of gendered associations, to me, that was maybe the more depressing finding of maybe we've come less far than we had hoped we would over time.

0:43:14.0 Beth Fisher: So another thing I was thinking about when you guys were talking about children's literature and these ideas about men and women and warmth, and how those ideas are so involved in so many books that we read and things like that, there was also how women are not good at math. I know all those kind of classic things that we think about. It was funny because as Ava and I obviously are in STEM and we don't believe these things, but I was thinking there's so many things I do within STEM that kind of relate back to me believing those things, so I definitely feel like I should turn down my warmth.

0:43:53.7 Beth Fisher: It's interesting, and I feel that when I'm in this environment, I need to not be maybe as caring and these kind of things because I won't be seen as smart, like regardless of what my work is doing, and I guess I'd never made that connection that, "Okay, that's probably to do tied with being a woman," and we have that association with being more warm, I have also been told before [laughter] that I should tone that down, so it's also something that I've been told by someone in this environment that maybe it's not as good to be like that.

0:44:27.7 Beth Fisher: But in other work places, 'cause I've worked in other places, not as a scientist, that hasn't been a problem at all, I've worked in retail environments, I've worked in a local council and no one ever said, "Oh, you know what, I think maybe you should tone it down on the love." [laughter]

0:44:42.5 Ava Ma De Sousa: I think it does sound like picking up these sorts of qualities that are implicitly linked to femininity, like warmth, and maybe not making that link, I think that just shows this implicit nature of these biases is that you don't even realise that it's really tied to your identity as a woman, but potentially you might still feel that kind of threat if you were asked about it. There are studies like this, if someone were to tell you, "Oh, you shouldn't act that way or be as warm or show as many emotions as you normally would," that maybe you would also feel more threatened because you're a woman and kind of feel the weight of that stereotype, there's a lot of research, like classic social psychology research on this idea of social identity threat about women in STEM and math, specifically.

0:45:27.1 Ava Ma De Sousa: So it's a Claude Steele classic work from the '90s showing that when you give women a math test and remind them of the stereotype that women are supposed to be bad at math, that makes them really anxious and that makes them fulfill that stereotype. So it could be kind of the same case if someone tells you something about being too warm, but there's research that also shows that it's actually really easy to create, in theory to create stereotype lift. They call it on the other hand, which is just telling someone this isn't diagnostic of your ability, that allows women who are originally really good at math to do just as well as they normally should.

0:46:01.9 Ava Ma De Sousa: I think those stereotypes are changing, and I think an interesting thing about the literature, although Tessa mentioned that these stereotypes haven't really changed, but one thing that has happened in the literature is that this effect of stereotype threat for Women in Math hasn't replicated, so it hasn't replicated in, I think, a German and Dutch sample, and lately it hasn't replicated in the US, and some people claim that that's 'cause social psychology is having a replication crisis and say, "This is bullshit, this never happened," but I actually think that it's probably because those things are changing although Tessa's research might suggest otherwise.

0:46:36.7 Ava Ma De Sousa: But I actually don't feel the way that you do, and this might come from me, never having had an issue, I think a lot of changing of behaviors and feeling threatened is when you learn that you get a punishment if you behave a certain way, and I feel like I've never been punished for being warm or for having certain... I guess you'd call them feminine qualities, but I don't really like that term, but for being very social, and I feel like I put a lot of effort into trying to be overtly enthusiastic and nice, especially in emails, my emails have way too many exclamation marks.

0:47:09.8 Ava Ma De Sousa: But I've never had any negative feedback from doing that, and it feels more not authentic, I guess, whatever authenticity means for me not to be acting that way, but I think an important thing to know also is that Beth is in a much more kind of STEM-y and bro-ey place in a way than I am because she's in a philosophy department and does more... Also computationally heavy research.

0:47:35.4 Beth Fisher: I also wanna be clear that where I'm at now, I love. [laughter] Just so that's clear, this was past experiences, and now I'm in an environment where I do feel like I can be 100% myself. And what's been really nice is being in that environment, my work has really improved, so now I'm in a place where I don't feel like I have to tone down at all, I mean, I hate even using the phrase tone down, but that's how I really felt, and now I'm in a place where I don't feel like I have to do that at all.

0:48:01.9 Beth Fisher: I can walk into the room and be 100% who I wanna be. And now you can just see my work has just improved and that's probably because I'm being more warm or loving towards people now, my research isn't as good, if anything, it's the opposite, I'm not stressed about trying to control that at the same time and trying to focus on work.

0:48:22.5 Ava Ma De Sousa: Yeah, there's also interesting work about stereotype threat that shows that when people feel threatened that they actually have a different so-called level of control, so they're focusing basically on different things, so if you're threatened, if you think there's a snake in your apartment, probably gonna be focused on finding the snake and vigilant about the snake, so you're gonna be focused on like, "Are people judging me because I'm a woman?" and not focusing and not being able to zoom out to the big picture details, and there's research that shows that when you successfully take away some of that social identity threat, that people do better in school and they're able to zoom out and they're able to not feel like everything that happens that's negative is a reflection of who they are and their belonging in research too.

0:49:08.5 Beth Fisher: I think when you feel that your personality is wrong [chuckle] or the way you're behaving is wrong, then if your work has any kind of mistake, which everyone's work does, no one is perfect, we constantly need to learn and improve from our stuff. But it feels more intense 'cause it feels like it's more of attack... I don't know, it's strange, on who you are. Whereas if you are accepted then it feels like your work is kind of separate to you and it's not when I'm getting com, like when the, when I'm, "Abby you could improve your work this way." It's not to do with like who I am, it's to do with my work and it feel, I'm like better at receiving feedback now because I don't feel like I'm being attacked.

0:49:46.8 Ava Ma De Sousa: Yeah, that's actually exactly what... So Claude Steele, who did this research, this original '90s research on stereotype threat, he does work on self-affirmation as well, which is this process where you just get people to affirm their value, so it doesn't even have to be things that they're great at, but it can be, it can be something like, "Oh, I think I'm a good friend, or I really value my religion, I really value my family, and spending time with my family, I think I'm a good son," and being able to zoom out of things that allows people to do exactly what you said.

0:50:16.1 Ava Ma De Sousa: So to decouple the immediate threats that they have in their environment, from their identity and therefore from their belonging in a specific setting, and they did this research that I had to read recently for a class that kind of blew my mind where they did this in a group of Latino middle schoolers, and they had them do this intervention throughout the year, and their grades improved and that lasted even for some of them until high school, so just having that kind of affirmation, lifting that threat, because it really creates these negative feedback loops, right?

0:50:48.3 Ava Ma De Sousa: Where you're like, "I'm terrible and I don't belong here," and then that screws you over for the next day, you're gonna be more threaten the next day and you're probably not gonna do as well, and there is evidence that at least in 2010, when the study was run that over middle school Latino students, grades go down probably because of this threat that they're having, but if you're able to lift that and decouple that as you were saying, then there's really surprisingly long-lasting improvements.

0:51:15.2 Beth Fisher: That's really cool.

[music]

0:51:19.3 Ava Ma De Sousa: So we've kind of talked about this in terms of how to address your own biases, but because the NLP stuff is kind of like a reflection of where society is, do you think that this work also informs the fact that perhaps if we were to get in more media that is less biased, that that would allow for changes to happen, do you feel like you have any data that speaks to that, or would it just be like speculation?

0:51:46.9 Dr. Tessa Charlesworth: Currently speculation, but also in the works in terms of data. So yeah, it's basically the biggest question facing the Computer Science and Social Psychology collaborations right now, which is the relationship between these massive language inputs and actual attitudes or actual human behavior, and we can get towards that relatively simply as we are doing right now in our ongoing projects of the relationship at aggregate level, how do changes in attitudes that we measure through Project Implicit for instance, track alongside changes in, say, the biases in newspapers across the past 20 years, or the biases in Reddit across the past 20 years.

0:52:30.3 Dr. Tessa Charlesworth: And we're doing that right now, so I can give you an update in a year or when those data are finally out, but that's one question, which is, "How do these two things relate to each other at a structural or at a macro level?" but I think the equally interesting, and perhaps even the more relevant for intervention question is, "If we were to change someone's media consumption and we were to show that quantitatively and give them a new diet of books, basically, could we then see corresponding changes in that individual's attitude?"

0:53:03.8 Dr. Tessa Charlesworth: And I think right now, that stands as a million dollar question for researchers, there is some early indication that it should be the case, it should work. So there was a really cool recent paper where they took Fox News viewers, paid them to watch CNN for a month and then measured their attitudes, and they found that there was attitude change basically in line with this idea that if you change your media consumption, you change your diet, you go from Fox News to CNN, then you should have a corresponding change in your attitudes.

0:53:35.5 Dr. Tessa Charlesworth: It's an early study who knows if it replicates across many different groups and across time and across all these other things, but it's enough evidence for us to continue to advance this idea that media does have an influence on your mind.

0:53:49.4 Ava Ma De Sousa: So in the gender stereotype stuff, you mentioned that there was quite robust gender stereotypes that persisted but you've also looked at how racial stereotypes have changed over time using the same methods. So I was wondering if you could discuss that study and some of its main findings.

0:54:05.1 Dr. Tessa Charlesworth: So this was the second big paper building off of these methods. Once we sort of shown the validation in the gender stereotype domain for these particular child corpora, we asked, "Okay cool, could we do that same method but look at data now across 200 years continuously in book text for 14 different social groups?" So we tried to do this whole big expansion and ended up being like thousands and thousands of lines of code. But it was very fun. And the key takeaways there are twofold.

0:54:36.7 Dr. Tessa Charlesworth: One that we looked at different sort of metrics of stereotype change. One being this more sort of surface level change in the top traits that we associate with groups. And then the second one being more latent dimensions of sort of the underlying meaning of those traits in terms of how positive or negative those traits were.

0:54:55.3 Dr. Tessa Charlesworth: And so the one of the first main findings is that when it comes to the sort of turnover in the traits themselves, there's actually a lot of change over time for most group targets that we look at. And this aligns with the fact that we can go from calling a group lazy to calling a group helpless 200 years later, and on the surface, sure, we've changed in the kind of stereotype, but when we look one layer lower to the kind of latent meaning of what those traits actually reflect, there's actually much less change because both lazy and helpless are negative in their representation, and so we found that in general, there was this kind of dissociation between this more surface level turnover, or surface level change in the content of traits, and these more latent persistent undercurrents over time across 200 years.

0:55:42.6 Dr. Tessa Charlesworth: The second key distinction that we saw was between the types of groups we were looking at. So we looked at more racial and ethnic related groups. So Irish and Black and Asian and Native American as well as more non-racial groups. So old young men, women, fat, thin and so on. And those non-racial groups of gender and age had changed way less than our racial and our ethnic groups. And that's really interesting for a number of reasons, including relating to these old theories, or I shouldn't say old, but theories from the 1990s in which Jim Sidanius proposed that these kinds of racial and ethnic groups are much more culturally constructed. They're kind of what he would call these arbitrary sets.

0:56:28.6 Dr. Tessa Charlesworth: And that's because basically for a long time in our histories we didn't use race to divide groups. That's a much more recent thing that we use to divide groups based on their resources and those kinds of things. Whereas gender and age and social class have been really long-standing divisions that we use across historical societies to divide labor, to divide, you know, placement of people in different villages. And so they should be much more stable across history just because they have so much relevance to the kind of way that we've structured our society.

0:57:00.6 Dr. Tessa Charlesworth: Whereas race and ethnicity being a little bit more culturally constructed, a little bit more culturally malleable should be showing more change over time. And so we also found that sort of general conclusion that these racial and ethnic groups have changed more than others.

0:57:13.3 Ava Ma De Sousa: I'm not sure what the causal chain is, but is part of that also what we were talking about before with the fact that for some of these groups there's just been a lot more racial protests about, "Oh, we shouldn't be calling Black people lazy," and then is it just kind of a protection of the social order still to have that transform into something like helpless.

0:57:32.8 Dr. Tessa Charlesworth: Absolutely, and I think... Yes, so in our most recent paper that's currently being revised on my other screen right now, [laughter] we are looking at some of those factors of frequency of how often we talk about a group or how much division there is in the way that that group is talked about, across society, and that does seem to be a key predictor of whether or not that group is gonna change in these kind of hundred years of text.

0:57:58.8 Dr. Tessa Charlesworth: So I definitely think that frequency of discussion or the kinds of intensity with which we approach a group in terms of protests and those kinds of things that will have an effect on producing greater change could be determined exactly how much that contributes above and beyond all the other features that we've already talked about in terms of the degree to which that group is more body-related versus more demographic-related, the more it relates to things that are controllable versus non-controllable, but I think, again, these methods are the first quantitative stuff that we can take to actually try to partial out some of that variance.

0:58:31.3 Ava Ma De Sousa: And were there any groups that actually did have a flip in terms of positivity or negativity? I'm thinking Asian-Americans, 200 years ago were seen as very low status, more kind of associations with dirty-ness, which I guess has kind of come back during COVID, these really different associations to what we see now, which is more of a mechanistic dehumanisation where it's like "Oh, you guys are so good at math," and maybe they're more robotic, but really high competence, which I think is really different to 200 years ago, was that reflected in the data and were there any other groups where there was something similar where there was a really big shift?

0:59:09.6 Dr. Tessa Charlesworth: So the representation of poor changed in valence, it was actually the only one that became more positive over time, the rest of the groups that did change in valence, and there were only five out of the 14 groups generally moved towards more negative representations over time. None of them actually flipped, interestingly enough, went from a positive representation to a negative representation over time, generally, it was like they kind of started out as already a little bit negative and just became more negative over time. [laughter]

0:59:37.8 Dr. Tessa Charlesworth: Yes, another positive finding for you to take away, [laughter] but the other really interesting thing that you mentioned is not just the valence of the Asian-American stereotype, to take that as an example, which might be quite stable over time, as we are seeing in our data, but there's that additional dimension of the content itself going from in the early 2000s in our data often represented in terms of Orientalism and mysticism and interesting mixed valence of reverence, but also othering.

1:00:08.1 Dr. Tessa Charlesworth: And then now being the top associates when we look at the words are often actually a lot of cultural connotation, so things like Charmane or Jade, or words that are referring to the group more factually in terms of the things that they're just associated with in our culture, rather than necessarily to do with their placement in terms of traits or adjectives. So there's a lot of richness when you start to look beyond just reducing it to valence to actually what are all of these content-related dimensions that it's changed across, so it's the idea of mysticism or orientalism that's fallen out.

1:00:45.3 Dr. Tessa Charlesworth: But now we've got all these new food-related or Jade-related words that come online that I think the Asian American scholars, people studying Asian American history will be much better equipped to try and understand all those really rich qualitative data, but it's now all openly available for people to make those inferences based on each of these 14 groups and how they've changed in these really unique contentful ways.

1:01:11.9 Ava Ma De Sousa: The final question that we give people to answer is, what are the things that you're really excited about, which you have been hinting at [laughter] during this whole interview?

1:01:20.2 Dr. Tessa Charlesworth: So much is next, I'm so glad I have my entire career ahead of me to continue to address these questions, I think for me, the biggest question is really triangulating across these methods and across the real world, so right now, and you kind of hinted at this in some of your questions of right now as a social psychologist, I've so often been trained of thinking in the lab and thinking about attitudes as the be-all, end-all of what we're interested in.

1:01:46.1 Dr. Tessa Charlesworth: But attitudes have relationships to behaviors as well as to our culture, and so trying to understand and triangulate how attitudes change, how that trickles into behavioral change, and how both of those might be influenced by cultural change to begin with, and media change, and changes in our language and our rhetoric, and so to me, triangulating across those three pieces of data, it's such a massive undertaking, but I think one that we are beginning to have the methods to really understand deeply.

1:02:14.9 Ava Ma De Sousa: Thank you so much, Tessa, it was such a pleasure to speak with you today.

1:02:18.2 Dr. Tessa Charlesworth: Yeah. Thank you so much for having me.

[music]

1:02:30.7 Ava Ma De Sousa: Minds Matter is mixed, edited and created by Beth Fisher, she's the Australian one and me, Ava Ma de Sousa. Our intro and outro music is Nobody Stayed for the DJ by Glassio. Our transition music is Back For More also by Glassio. We'll be back in two weeks with a brand new episode of Minds Matter. In the meantime, find all our episodes and show notes on mindsmatterpodcast.com.

[music]