>> Welcome to the podcast series of Raising the Bar, Sydney. Raising the bar in 2019, saw 21 University of Sydney academics take their research out of the lecture theatre, and into bars across Sydney, all on one night. In this podcast, you'll hear Jason Chin's talk, "Can We Put an End to Wrongful Convictions?" Enjoy the talk.

[Applause]

>> Jason Chin: Hey, everyone. How you doing? Nice, nice. So, yes, this talk is, I think at its core and my research is at its core, about people as you can see in that little one-sheet that's around, it's about people fooling themselves. And as humans, I think we're really good at that, all of us. Our memories are reconstructive, so it's very easy to, when you're trying to formulate an argument, call to mind examples that support your view. Our cognition is motivated. So, if you've ever tried to engage in a political conversation with someone from a different ideology, you can be having the same facts you're talking about but completely different points of view. If you have ever tried to talk a mate out of calling his or her ex after a night of drinking, you can – they always come up with really good reasons that it's – that they should call their ex, but they're usually not. So, what I look at is when scientists fool themselves. And this can be especially problematic when there are scientists who are giving expert evidence in criminal trials, and especially when the other side, which is usually the defence, doesn't have their own expert. So, you're just relying on one scientist who is fooling them self. We've learnt recently that this is actually much more common than we thought. Maybe in the media, you've heard stories of something called the Replication Crisis or the Reproducibility Crisis? So, in some fields, they've attempted to recreate the findings of studies published in the top journals, Science, Nature, the top psychology journals, top medical journals, and in one study, they tried to recreate 100 of them, and only found the same results as the original in 35, which is kind of scary and kind of problematic. And it's led to what I think, and what many people think, is a revolution in science. So, we typically think of scientific revolutions as based on some sort of discovery, like Darwinian biology or quantum mechanics. But this revolution is more about methodology and it's about openness and it's called by many people, "The Open Science Movement." And I think the idea is that if science is conducted more transparently, then other scientists can actually police those scientists. The public can police them. And it becomes much harder to fool yourself. That's what I'll talk about today. So, I'll go over an example of a case that I think is quite interesting and that illustrates this, and I'll return to it at the end, after explaining kind of the nuts and bolts of this whole thing. So, the case is – oh, and I think – so, if you're interested in fooling yourself, I forgot this, on that little one-sheet, there's examples of that. So, studies that we've asked people to evaluate them self and found that they're not very good at it. So, 94% of college professors in the U.S. say they do above average work, which is impossible, right? Only like 49% can. They did a study in an election ten years ago in the U.S., and 90% of people said they would vote, but only about 70% actually ended up doing it, and doctors diagnosing pneumonia on average,

gave 88% confidence of that diagnosis, but only 20% were actually correct. So, we're very good at fooling ourselves. And we maybe see that in this case that I'll talk about, the Crown and Warren Abbey. This went through the Ontario courts, beginning in about 2007. It was this guy, Abbey, Warren, who was a – admittedly, a member of a gang, and he was accused of killing a member of a rival gang. And there was not much evidence. There was no eyewitnesses. He did have a motive. This guy had sort of done an afront to his gang earlier. He did tell two of his mates that he did it, but that's, you know, not the greatest evidence. He could have just been bragging. And both of those mates had a deal with the Crown to give evidence. So, maybe not the most reliable. Really, the key piece of evidence was a sociologist they brought in. This guy named Mark Totten, and what he was concerned with is that Abbey, pretty soon after the killing, got a teardrop tattoo under his right eye. And Totten is a criminologist working in Ontario. He's sort of, I guess you could – he claims to be the leading gang expert. And he says, "It's clear to me that Warren Abbey's teardrop tattoo on his right cheek below the eye represents the fact that he killed a rival gang member, most likely in 2004." And the basis for this was that he had interviewed many gang members during that time, about 300 of them. And each one with a tattoo that had killed – that was in prison for homicide, said that was the reason that they got the tattoo. So, at the first trial in 2007, Justice Archibald, the Ontario Supreme Court, excluded Todd. And he said, "This is not scientific enough. This is just interview evidence. I don't know if these gang members were telling the truth. And I don't know if they're actually very representative of Totten's gang – or of Abbey's gang." Which seemed kind of reasonable to me. Without that sociologist, which was the key evidence, Totten – or Abbey was acquitted. He went free. But then the Crown appealed, which they can do in Canada, and they said, "Well, you shouldn't be holding this guy Totten to this you know, scientific standard. He's a sociologist which has different norms, and interviews are very common methodology in sociological research." And the Court of Appeal agreed, and they said, "Yes. Yes, Totten can give evidence." He did, and he was convicted in 2009, based on the sort of mainstream norms of sociology at the time. I think now is a good time for me to turn to my experience as a psychologist, with those norms. So, I'm actually 38, I should say, so a lot of this is not going to make sense. I'm actually quite old. Early 2000's, I went to graduate school, edit point, edit that out, Martin. Took for graduate school in cognitive psychology at the University of British Columbia. And I was studying memory, and the first job, the first sort of project my supervisor gave me was to study this memory effect called verbal overshadowing. And it's this idea that if you verbalise a face, as you're looking at it, if you describe it, it actually impairs your memory for it. And the theory is that our language, our words, can't really capture a face, which is a very holistic sort of experience. And we end up saying things that aren't quite accurate. So, when we go to remember it later, we have a poor memory for it, which is kind of counter-intuitive, which is why it got published and was sort of well-publicised. So, we said, "Hey, why don't we see if this actually generalises across cultures and languages that have different sort of words for faces?" Like, really cool idea. So, my first job was

to, as you do in any study where you're trying to extend something, you first replicate the original finding. So, I – as you do as a scientist, you get ethics approval for it. You gather all the materials. You recruit subjects. And you have them go through this process. It took about six months. I got the data in. I analysed it and there was like no difference between the two conditions, which kind of makes sense, right? Like it's not really clear that verbalization really would affect your memory for a face, one way or the other. Maybe it does impair in some respects because you have poor language for it, but maybe it also like bolsters your memory because you're rehearsing it, right? So, that helps memory. So, that was discouraging. And so, we said, "Okay, what's going on here? Let's try to do a better job. Let's have him describe it for five minutes, and maybe that longer period of time really forces them to come up with words that aren't really accurate, and that will help us find verbal overshadowing," which has already been found. Nothing. Okay, well about a year and a half has gone by now, and PhD programmes are six years in Canada. And like, okay, this is getting a little close for me. Maybe one problem is we're using the original photographs, which were from like 20 years ago. Hairstyles have changed. Like, people look differently now. Maybe people are kind of like thrown off by that. So, let's get some modern photos. So, I got ethics approval. I got some grant money to go take photographs around campus. People gave consent. And I used the modern photos. Did the study again, about two-and-a-half years later. Nothing. And I guess most of you probably aren't scientists. I don't know how many [inaudible] is. It's very hard to publish negative findings. So, if you don't find anything, you can't really publish it, right? Maybe you should be able to? I truly think that. So, my friend Azeem [phonetic] who was – edit that out. My friend, he's going to get mad if I talk about him, he did a really cool study about how people who are prime with religious objects, are more like generous later. He was in the New York Times. So, he's getting published. He had job offers and I'm just – I've got nothing. And so, I started taking the LSAT and applying to law schools. I finished the law degree. I finished the psychology degree. Barely published anything. Kind of did like a mediocre PhD. And went to law school. Never having replicated verbal overshadowing. This is now, 2013, 2014. There's a story, like in the New Yorker, about this replication crisis, that all these well-studied phenomenon can't be recreated. And I was very interested in that because that happened to me. And I was like, "Okay, I feel a little bit better now." Like I – maybe that's the reason I failed, right? Like, it sucks to be a failure and I felt like a failure for a long time. And it's my bias. It's how I fool myself. And so, I started looking more and more into this, like in my off time, I was litigating cases in a commercial law firm. And this crisis actually motivated lots of really, quite interesting research, into these kind of grey areas of scientific practice. You might hear of them referred to as questionable research practices – researcher degrees of freedom, or P-hacking. These are all things that aren't expressly forbidden in scientific practice, and actually done quite frequently, with the biassed results quite strongly. So, a couple of examples. One example is to exclude outliers in your data on a sort of ad-hoc basis. So, I collect some data and I see three of them are quite low. Like

quite of – like three of the subjects said that they had really bad memory for this thing. Like, okay, maybe it's you know, this doesn't fit my data's narrative. It's not what I want to see. But also like, maybe they can like – weren't paying attention. So, that's like maybe a good reason to exclude them because they're not doing what I asked them to do. Another one is to run the study, check the results, see if you've found what you're looking for, and if you didn't, add like ten more data points. And then check it again. Oh, it wasn't quite what I was looking for. Add ten more. Oh, now it's there and you publish. It's not quite intuitive, so I'll explain like why that does sort of bias the findings. So, here's the example I give.

>> You're listening to Raising the Bar, Sydney, 2019.

>> Jason Chin: So, imagine I have a coin. And I say, "This is a special coin. It's weighted the same, but 75% of the time, it gives heads." And like, "That's pretty cool. Do a test to demonstrate that." So, I have my research assistants do it, and it's 50 heads, 50 tails. Like, well that's not quite right. So, I go into the data and I see, well, Research Assistant 1, he's an expert flipper. Like, he's been flipping coins his whole life. He's very good at this. The other one's never done it before. Let's just look at his data. And he's 60%. It's like, "Okay, that's pretty good." Also, I noticed that in a few of these cases, it hit the ground and rolled and hit the wall. That's a little bit different than the other one. So, let's exclude those. Now, it's 70%. Like, this is looking like I have a special coin. So, then not quite there. So, let's run ten more – let's have – flip it ten more times. Five heads, five tails, not great. Do it again. Oh, there's a run of nine heads. Stop. And we've got the special coin. If you just saw the final result, you would think that looks pretty convincing. But if you saw everything that led up to it, it's not very probative. It's not the strongest evidence that this coin is special any way. And that's – I think that's – so, I'll ask you then. So, they gave an anonymous survey to ecologists, biologists, and psychologists, and they asked, "How many of you have ever used that optional stopping rule. How many of you have looked at the data and used that to determine if you should collect more data?" What percentage of people do you think, of psychologists, ecologists, and biologists, anonymously said that they had done that? Oh, okay. I thought you were going to go lower. So, this is not going to be very impressive. But it was 60%. And if everyone's doing 60%, then you're going to – that kind of explains what's happening here. So, what's science doing to fix this? It's, as I mentioned, there's a – really a pronounced move towards open science. The National Academy of Sciences, Engineering, and Medicine, published a report in 2018 called "Open Science by Design," saying that this is how science should be done now. One of the key reforms is something called preregistration. And this is actually pretty common in medicine, because in the late 80s, early 90s, clinical medical researchers were finding that most drugs that they studied, worked, or most published studies said that they worked, which doesn't – which didn't really map on with reality. So, and then there was some big controversies because they weren't working. Some big pharmaceutical companies were sued, and they started adopting this practice called preregistration. And what this

is, is before you do the study, before you collect data, you put on an online database, which is private at first. Here's what I'm going to do. I'm going to test this on 100 people. Here's my rules for what's an outlier. He's how I'm going to do the data analysis. And then when you go to submit it to a journal, that becomes public, and the editors and reviewers can say, "Oh, okay. You didn't do what you said you would – on what you did. So, maybe that makes us put less weight on this finding." Or maybe you did exactly what you said you were going to do, and that is – that's good. That's what you'd expect. Another one is open methodology. So, lots of people are surprised to know that historically, you hadn't been required when you publish an article, to put your entire methodology in the journal article. This is for historic reasons, because they used to be mailed out to scientists, and there was page restrictions. So, you could just very – in a very summary fashion say, "Here's what I did," and as you can imagine, that can hide all manner of sin, right? You can sort of gloss over the fact that you've had three different research assistants. Open data. So, people are also surprised to know that, still most cases, you aren't required to provide your data when you submit an article for publication. Only a few journals require this now. It's growing, which is good. I think the final one which maybe most of you or more of you are familiar with, is actually open access to the journals themselves. So, like if you don't have – if you aren't affiliated with the university, maybe you come up across this, and you've tried to – and you say, "Oh, here's an interesting article. I want to read it." You try to click it. It's like, "Pay $30," or something. Some crazy fee. So, this knowledge really has not been open to scrutiny by the public in a meaningful way. And there's a big movement towards that, both to make knowledge more democratic, but also to make it more reliable. So, yes. So, cut back to me – cut back to me five years ago, I felt vindicated. I felt good. I felt like I wasted some time. I wish this movement had happened maybe ten years earlier. But professionally, I was quite worried, because I was litigating cases where we had expert witnesses relying on scientific research that might just be – might be wrong. And that's kind of scary. I think this is maybe most – this is a bit different, but an example, and the example that I work with a lot is, forensic scientific research. So, that's on one of the slides. In the 90s, one forensic practice was developed that kind of changed the game for all of the other ones. That's the DNA analysis and the technology that allows you to take a DNA sample and pretty definitively, match it to a certain individual. And once that became possible, they began to, in the U.S. especially, they began to look at previous convictions in which some of the biological material was stored. And they tested it. And they found in many cases, that they had gotten the wrong person. And kind of ironically, the rise of that DNA technology, that forensic process, cast a lot of doubt on other forensic processes, because they learnt that in many of those cases, it was some other forensic practice that put the person behind bars. Things like fingerprint identification, finding a fingerprint at a crime scene, and matching it to a certain person, finding – that was only tested in 2011. We didn't know that worked and how well, until 2011. Bite mark analysis is sort of a notorious one. That's you find a bite mark on somebody,

and match it to somebody's teeth. Firearms analysis. So, you find a bullet, and you match it to a certain gun that shot it, based on the sort of patterns in the gun. So, that all happened. Academics for a long time, when those exonerations were happening, were raising awareness of this. It became – lots of articles were published. Finally, in 2009, the sort of landmark article, sort of the landmark report of the National Academy of Sciences, which is on the slide, was published, reviewing all of these exonerations, reviewing all of this evidence. And it – the most famous line is – I'll just read it. I need yours. "The bottom line is simple. In a number of forensic science disciplines, forensic science professionals have yet to establish either the validity of their approach, or the accuracy of their conclusions, and of course have been utterly ineffective in addressing this problem." So, for most of these practices, they were never tested. We never knew if they worked. This inspired research. The forensic community finally got off their asses and started doing this research. It turns out that they couldn't demonstrate bite mark analysis works. It's wrong about as much as it's right, because the skin kind of swells and bruises in unpredictable ways. Fingerprint analysis does work pretty well. Like, they make errors 1 in 306 times, which is good compared to bite marks, but you know, if that's the only evidence against you, and they're wrong 1 in every 300 times, it's – that's a tough sell for me. And so, they're finally doing this validation research. And my worry, like my big concern, is – and what I research a lot about is, if they start doing this research with – in the way science has typically done things, it's not going to be much better. And it might actually be misleading. So, if I do these validation studies, but I don't preregister it, so I had the chance to – and this has happened, in the study by the Miami police, just exclude some observations because they didn't quite make sense, so they excluded 26 – so, it was a fingerprint study. I think they excluded 26 of the examiners responses because they said – they were so bad, that they must have made a clerical error. Well, if you don't preregister, like we can't know if that actually happens. It took the U.S. Presidents Council of Science and technical advisors to sort of swear that would happen there. So, if these forensic professionals start coming in the court, with these validation studies that are, you know, full of researcher degrees of freedom, and P-hacked and what not, I think that just kind of gives this sort of – this misleading patina of science, to stuff that really isn't very scientific. That brings me back to the Abbey case. So, interesting thing happened in Abbey. In 2012, so, he's been in gaol for three years after the retrial. This guy Totten, starts giving evidence for the defence in certain cases. So, now he's up against quite a party, the Crown, who has the resources to actually go through – go through his research and try and sort out if it's actually very strong or not. And I guess they have some interns and some people going through them, and they find out that in each study – he's done like five studies on gangs in Ontario, and in each one, he has a different definition of gang member. And he uses the definition that fits his results best. So, if he wants to show someone's a gang member, he uses a very liberal definition. If he wants to show someone's not a gang member, he uses a very conservative definition. But if he uses his more restrictive one, that he's been using lately, of the 300 people who were

gang members in the research he quoted in Abbey, only 20 were gang members. So, you know, a very impressive field of research – a very impressive body of research becomes much less impressive. Finally, this gets appealed up again in 2017. The Court of Appeals says, "You know what? That should have been excluded. It actually isn't very impressive research. And there's going to be a retrial." The Crown is embarrassed, and they drop the manslaughter. He pleads out to that. And he's released ten years after this whole thing happened. He's been in gaol a lot of that time. All based on what are really mainstream – it's mainstream sociology at the time. Like, there were no requirements that they use a prespecified definition of what a gang member is. And so, I think the danger that Abbey demonstrates is that, if forensic science does become mainstream science, it's still not much better, and maybe actually a little more worse. I think there's probably some lawyers in the audience. I'll end with this. In law, we kind of treat openness as an ideal, right? We have open courtrooms where the public and the media can come in. We have judges publish their reasons for their decisions. Science, forensic science especially, is increasingly like an important cog in that criminal justice system, yet it's conducted in a very opaque way. So, what I think that forensic science should do, is to follow suit to become more like, law in some ways, and open itself up, and maybe that's a good way for it to help stop fooling itself. And that's all I have. Thank you.

[Applause]

>> Thank you for listening to the podcast series of Raising the Bar, Sydney. If you want to hear more podcasts from Raising the Bar, head to Raisingthebarsydney.com.au.