Knowing how the other thinks: the brain and influence in international confrontations

Title: Knowing how the other thinks: the brain and influence in international confrontations
Date: 13 October 2015 (12:00-13:30)
Duration: 53 mins

OK, so thank you very much. So I said to Nick that I was happy to introduce myself to reduce his workload, you see, 'cause he's very busy, so now he can do some emailing in this one minute. And so I'm Nick Wright, and I was supposed to give this talk a few months about [0:00:18] to introduce myself to everybody, but for a variety of reasons it was cancelled. In fact, it was entirely Theresa's fault 'cause she came to give a talk and so it was cancelled for that reason. And then there was the summer and so on.

So it's really – but it's really partly to introduce myself and just say that I'm – so to describe a little bit about what I'm doing, and say that I'm very happy to collaborate with other people and work with other people and – you know, I've been involved in a wide variety of collaborations and still am, and want to do that obviously here at Birmingham.

And so essentially my work falls into three – there are three strands to my work. So the first strand is looking at taking core concepts from the decision sciences, so psychology and neuroscience in particular, and applying them to decision-making in international confrontation. The second strand is to look at decision-making between cultures and so for example we've got some experiments just about to start in China and in Iran, looking at decision-making in both those countries, with leading psychologists from those places. And the third thing is to look at – to apply things to real-world policy applications, so the point isn't just to understand the world; the point is to change it, to slightly misquote Karl Marx, and…but it is to do with policy, and so I'm very linked in with both the United States Government and the United Kingdom Government, and so, for example, we are hopefully, with any luck,

just about to start a project which will go on for the whole of next year looking at grey-zone warfare [0:01:59] the Pentagon and the Special Operations Command and I'm very happy to talk with people about that. I know that [Danny - 0:02:06], for example, is interested in ambiguity, an area in which I've done quite a bit of work myself and that's basically one of the main areas they're interested.

OK, so those are my areas, and now I'm going to talk a little bit – sort of give the presentation that I was going to give today.

So, in order to conduct deterrence operations or to successfully manage escalation or reassure others, you have to anticipate how they're going to decide to respond to your actions. OK? And that's why having a realistic understanding of human decision-making is critical to deterrence or reassurance, for example. And so today I'm going to talk about two key insights from neuroscience. These are two core aspects of human decision-making that we've really been elaborating over the last 20 years or so. So the first of these is neural prediction error, which I'm going to describe in a little bit; and the second of those is our social motivations, and again I'm going to talk about these in a little bit; and I'm also going to discuss four simple rules for using neuroscience.

So one of the key things, and I'm sure Nick gets this every time he gives a talk about psychology or neuroscience, is, you know, why does it matter for the real world, and so on and so forth? And so you always get asked the same questions, so I have four simple rules for using neuroscience. So the first of these is 'are we sure enough of the neuroscience?' So just as Robert Jervis said, you know, you can always find the historical case to support almost any contention you want to make in international relations. He's entirely correct, I'm pleased to say. And in addition to that, you can also find almost any – you can find a psychology experiment that will support almost any contention you want to make. OK? So what I do is I try and take core, well-supported findings. OK?

So the second thing is 'does it matter in the real world?' Even if something matters in the laboratory, does it matter in the real world of states? You know, of highly-intelligent people often making decisions supported by huge bureaucracies and so on? And so you need empirical evidence, and there are two main sources of that. One is to look at, for example, in international – thinking about decision-making in international confrontations, is to look at historical cases across diverse contexts, so that's very much the sort of route that Robert Jervis went down in his classic work, *Perception and Misperception in International Politics,* published in 1976. And the second thing which isn't so relevant to that but is relevant to some other things is things like randomised, controlled trials or…randomised, controlled trials or more – you know, that type of thing. And we've done a – I'll talk – I'm coming to that very briefly at the end of the talk when I talk about Israel/Palestine.

The third question is, is it worth adding? So even if something matters in the laboratory and if that same thing does matter in the real world, you know, why does that matter at all? You know, why should anyone take any knowledge of that? So one of the best people was a chap I worked with in Washington, so he was in the Chinese nuclear weapons sort of…Chinese nuclear weapons – well, complex, I guess it would be. And I don't know if you met him when you were there, Josh – Li Bin? Professor Li Bin at [Ching Warren - 0:05:40] at Carnegie. And he said, you know, you're a decision-maker and you want to get the area of a piece of paper and so you could obviously – you know, width by height gives you that very quickly. And he said, you know, obviously, you know, you could maybe – if you want to get the area of this bit of paper now, what you could do is, you could work out the area by doing lots of little squares and adding them all up – whatever – but basically adding that in is pointless. You're still going to basically do width by height, if you just want to get a quick answer, and that's the only really useful thing to do. So everything we add in from psychology has an opportunity cost. OK? So everything we add in from psychology or neuroscience has an opportunity cost. So if I had to – so, you know, if I'm doing something on ISIS, everything I add – say, 'Oh,

you've really got to take account of this psychological aspect', that means they're not taking account of something to do with social media, for example, which could be, you know, more important. So you have to be aware of that and try and simplify concepts.

And then, finally, 'why neuroscience?' So why can't we just use psychology, for example? And this is a key criticism that a lot of people get, you know, from a psychological perspective. They say, 'Well, you can just use psychology', or 'you can just use economics'. Why – what does neuroscience buy you over and above explanations purely derived from the behavioural level? And I'll come onto that in a little bit. So why is it more [0:07:01] than psychology?

So, the first example I'm going to talk about is the neural phenomenon of prediction error. OK? So this is one of the core findings in neuroscience over the last 15 or 20 years, and it's a very simple concept but it is central to the way that humans learn, understand and make decisions about the world. OK? So, basically, when you receive – when an action occurs to you – when an event occurs, the psychological impact that that impact has is critically modified by how unexpected that action is. OK? And the brain very, very carefully measures this unexpectedness across a wide variety of different brain systems, and so you can measure this neural phenomenon of prediction error – you can simply call it sort of the identified event minus the expected event. OK? And so this explains – this is – you know, as I said, it's central to how humans perceive the world, how humans understand the world, how humans respond to rewards and punishments, risks and so on, and – but, now, can it explain a wide variety of historical cases across diverse concepts – across diverse contexts? And so I'm going to give you one simple example, although there are a wide variety more.

So we can think about things, for example, in a simple 2 x 2 table. So an event can either occur or not occur, and an event can either be expected or not expected. OK? And so we can see then how prediction error – so the

prediction error associated with those events will differ in each of those boxes. And just [0:08:55] the bigger the prediction error, the bigger the psychological impact that an event has.

And so let's think first of all about – and so, as I said, there are a variety of different cases one can look at, but I'm going to talk here about the effects of the psychological impact of strategic bombing. So here, for example, we have – the case for an event was not expected and the event occurred. OK? So in strategic bombing, in World War I, there were zeppelins dropped bombs on London – German zeppelins dropped bombs on London, and this caused – had a large psychological impact. This was highly unexpected, and it had a large psychological impact of people running into the streets of London; caused panic on the streets of London. People were assaulting officers of Royal Flying Corps, they were – you know, calls to close down factories and so on and so forth. And because it had a large psychological impact, the key interwar airpower theorists said, 'Well, if a relatively small number of bombs can have a large psychological impact like this, imagine if we do – we use a lot of bombs again and again and again and again.' OK? So this – so, you know, it will cause a huge psychological impact and this will [0:10:10] the opposition and we will win the war without having to fight a conventional war – or much more of a conventional war. And so [0:10;18] and so on and so forth. And this was dominant in air forces across the western world.

So what actually happened? So what actually happened was that when German air-raids occurred on London in World War II, now they were well expected. They did bomb with much larger amounts of bombs again and again and again and again, but now they were much better expected, they had a much smaller prediction error and so they had a much more muted psychological impact than was anticipated.

And then finally we can think about – we can also think about when an event is expected and it doesn't occur. OK? So now an event is expected and it doesn't occur. And so it has a large prediction error, OK? So, again this has

a large prediction error, but now this prediction error can actually have the opposite effect. So, for example, if you're expecting bombing and you don't receive it, then this can be perceived as a positive thing. And an example here is the United States used long, prolonged bombing – they bombed continuously during long campaigns in the Vietnam War, and they used bombing pauses as signals, and there is some evidence, at least, that they were taken as such by the North Koreans. Signals of conciliation.

OK. So that's prediction errors, and we can also – so we can demonstrate that predictions errors explain events in historical contexts, and what we can also see is that prediction errors subsume a wide variety of key strategic concepts.

So one example is up here. So an event is not expected and it occurs. And of course this subsumed much of what people talk about in terms of strategic surprise. OK? So strategic surprise, which is central to, for example, essentially thinking of – if you ever read Sin Tzu's *Art of War* or, you know, a wide variety of different people…or Clausewitz. So strategic surprise is just one instance of prediction error. OK? So prediction error subsumes the surprise. And surprise is central to both, for example, US and Chinese doctrine. So here we have – for example, here on the left we've got – this is the – what's called the *Joint Operational Access Concept* from – published in 2012. So this is a key document in how the US advise a – essentially – it doesn't say the word China anywhere, but this is a key document for how the US would fight a war with China and, as you see here, they say 'maximise surprise' – this is right at the top of the document: 'maximise surprise through deception, stealth and ambiguity to complicate enemy targeting'. Surprising the enemy is always a virtue in war. OK? So surprise is central to the way they think about conducting military operations. But, of course, surprise isn't always a virtue in war. It is if the only thing you're interested in is war-fighting, but if you're also interested in the signalling impact of your actions – so if you're fighting a limited war, then that's not necessarily the case. It should be used as a tool that you can manipulate appropriately. And you can also see

that the Chinese, in fact, arguably – I would say very strongly one can argue that they place a much – even a much greater emphasis on surprise. And so this is taking from *The Science of Second Artillery Campaigns*, which is an internal document from their second artillery, which is their ballistic missile forces, and so in *The Science of Second Artillery Campaigns*, as you see, they say, 'It is necessary to strike the enemy at the first opportunity, before the enemy has discovered our campaign intentions and actions. Surprise the enemy at – before the enemy, strike rapidly; catch the enemy by surprise.' And there is a very strong emphasis in a wide variety of different Chinese doctrinal writings on initiative – grasping the initiative, seizing the initiative, maintaining the initiative and using surprise right through to the end of limited wars or campaigns.

So what are some policy recommendations? So I know that policy recommendations are often some of the least interesting things to people in a university, so I apologise for bothering your elevated minds with such prosaic things as policy recommendations, but what can we say that's of use? So we can think about where – how do you use prediction errors when you're making actions?

So first of all if you're making actions. So there you are. You're [0:14:52] into strategic command or you're – and you're trying to – or you're sitting in the National Security Council and you're trying to tell President Obama what you think the anticipated impact – political or psychological impact or signalling impacts are of, say, two or three different actions. OK? And so what we can see is that you can use prediction errors as a tool in signalling. So, first of all, when preparing potential – potential options for a decision-maker, for each option, this gives you a specific question to ask. How unexpected will it be for the adversary? And then there are wide variety of different things that you can look at that will – so this gives you a targeted question. So Nick's obviously very interested – Nick [Wheeler's - 0:15:35] obviously very interested in empathy, but this gives you an actual, practical, targeted question that then gives you the beginnings of a checklist for how you actually

do empathy. It's all well and good saying – being empathic, but how do you actually do that? This gives you a specific question you can ask as a [0:15:50]. And, for example, this might be that some domains are inherently at the moment more – have more unexpectedness associated with them, just as we saw, you know, air power did in World War I and the same would be true, for example, now of [0:16:05].

The second thing: if you can manipulate predictability, so you can use – so you deliberately choose to manipulate predictability. So, for example, you can use signposting. So in a Pacific scenario – a China/US escalation scenario, you might think, for example, that you – the United States may want to move a carrier fleet, but when should it do that? If it signals a day before that it's going to move its carrier fleet, then that is altering the predictability of that. It's reducing its psychological impact but you still have the same military – you still have the carrier fleet moving. And, for example, in war games, whether they – if the US – China/US war games, you know, if the US alerts, for example, parts of its strategic forces without warning the Chinese first, then this has a bigger psychological impact than if they warn them first, for example. And you can also think, you know, in – this can also be applied to a wide variety of other different scenarios. So if you're thinking about Israel/Palestine, for example, one of the things you want to do is try and increase – if you want to reduce the impact of actions in a potential escalation scenario, you want to increase the predictability of the actions, and that's an actual, practical thing that the international community can [provide - 0:17:26]. So I did some work for International Actors in Israel/Palestine last year and earlier this year, and this is something practical they can do. They can attempt to increase bandwidth and increase the predictability of each other of the actions between the two sides.

And, third, obviously I would argue that both the United States and China should change their military doctrine to avoid simply maximising surprising doctrine. They should obviously stress the idea that this is a tool that should be used to signal. Although, of course, you know, when they're doing war

fighting – you know, you have to do war fighting and so obviously, you know, they need to do that.

Secondly – so that's when making actions, but what about receiving actions? So the key thing is that prediction error's unavoidable here. OK? You will always have prediction errors. You always have surprise. So you need to manage the effects of prediction errors on yourself, so don't overreact, essentially, and the other thing to say is that prediction errors are – you know a key reason why we care so much about prediction errors in the brain is because they are a key learning signal. So they tell us we need to change and update our models of the world.

So, in summary, prediction errors are a very simple idea but we have a very wide variety of different sources of evidence that they're fundamental to the way humans understand, learn and make decisions about the world, and we can show that they're important across a wide variety of different historical contexts. And the take-home would be to understand prediction errors and use them as a tool to implement and interpret signals. OK?

So this is deliberately simple. That slide has literally been shown a number of times in the Pentagon and in other places because you have to be simple, because those guys are, like, going off to have some lunch or do something genuinely important, you see, so they don't have time, necessarily, to take on board anything more complicated.

And how does this relate to the four rules for using neuroscience that I talked about? So the first thing is, are we sure enough? Yes, we are. This is absolutely core, OK? Prediction error is absolutely central – it's a central finding of decision neuroscience over the last 20 years or so, and we're very cer[tain] – you know, we're very confident this is important. Secondly, does it matter in the real world? And, as I said, we could talk – look at historical cases across contexts. Third, is it worth adding? So you can turn it into both a simple and operationalisable format, and because it subsumes other, pre-

existing concepts, it doesn't add to the analytical burden. Third, why neuroscience? And so now – so what we can say is this – so I'm going to give you two reasons – two main reasons.

So the first is the idea of consilience. So consilience is just a philosophical – is a simple philosophical concept – just a really simple idea that if you have multiple, independent sources of evidence for a contention, then you're more confident that that contention is correct, OK? So it's a simple thing. And so the facts that prediction errors – we have this wide variety of neuroscientific evidence means for prediction errors – just give you – us an extra, independent source of evidence that prediction errors are likely to be important. OK? Amongst the many, many different competing, behavioural-level concepts.

And, secondly, we can talk about universalism. So the fact that prediction errors are important in Pasadena and the fact that prediction errors are important in London and in Birmingham [0:21:00] about showing prediction errors are literally just up the road here or across the railway track or whatever it is. You know, we know that it's important in these places and we also know that prediction errors are important across a wide variety of different animal species. OK? It's important in rodents, it's important in monkeys, it's important in bees, even.

So we know it's important across this wide variety of different animal species, and in fact – certainly – obviously not bees, but in mammals we know that it's the same neural structures that are involved as when humans make decisions about wars and [0:21:37] and prediction errors and so therefore *a priori* we can never be certain but we can be much more confident that it will all – that prediction errors will also be important, for example, in China, where we haven't yet got this type of research, although I'm obviously trying to collect that data.

And we can think about universalism in two ways – OK? Two important ways. So one – so here it is in this little diagram. So we can think about universalism both within and between cultures. OK? So by – between cultures – so we can, for example, think of some kind of trait, like the impact of prediction errors, and it's likely to be normally distributed in China, 'cause most of it – or there'll be a distribution in China and there'll be a distribution of the effects of that in the United States, for example, and so the idea of 'are things different between cultures or the same between cultures?' is saying, you know, are these two distributions over – to what degree do these two different distributions overlap? OK, so is China on average different to the United States on average, for example? So are there things that are universal between cultures?

#But then also we can think about universality within cultures, and so the example I always give 'cause it was just – it was just incredibly striking was I went to a meeting a couple of years ago, taught – and on which two of the – there was a panel with two people on. So there was Bill Gates and Bill Clinton, and these guys are just extreme guys. They are both – and Bill Clinton in particular, he is an extremely articulate guy. He is the most articulate – even more articulate than anyone I've met so far in Birmingham. He's that articulate. He was the kind of – every one of his answers was perfect. It was exactly the kind of thing where, you know, you go home on the bus and you're thinking, 'Oh, I wish I'd said – you know, and everyone had gone, "Wow, that's amazing".' And that's how he sounded all the time. And this was just some, like, random event in Washington. He is an extreme guy. Now, I don't know what personality dimensions he is extreme along, but he's an extreme guy. He's an outlier. And so you have to say, 'Is there universalism – like, are – is, for example – do prediction errors matter even in these extreme populations?' So these odd people here? The odd – like Xi Jinping is – you know, he's a pretty extreme guy, probably. I imagine. Along a variety of personality dimensions. And so will Barak Obama be. And so do these same things matter in these extreme populations? And so for example the fact that prediction errors matters across a wide variety of different

species/across a wide variety of different human beings that we've looked at makes us much more confident there will also be universalism within cultures as well as just between cultures.

And just to say that another thing that people always say is – they always say, 'Ah, but, you know, also how do we know that even though you say that Bill Clinton's really clever and – but then we've looked at clever people and monkeys and all this kind of stuff.' You know: 'How can we know these things really matter in leaders?' And that's true. How can we know these things really matter in leaders? And, you know, ultimately, there's no really good evidence either way for almost anything.

But there's one good source of evidence for at least one sort of psychological concept that I can – that I've looked at where we can look in a leader, and so we can, for example – so the reason why prediction errors are so important is that – prediction errors are important because we don't see the world just by looking at the information of the world as it comes into us. We basically see a model of the world all the time and we're particularly interested in deviations from our model of the world. So the key thing is that we see models of the world all the time, and…and I won't go into it all now but the gist is that our models of the world tend to see patterns in the world and that's why people, for example, tend to see conspiracies. And there's a nice literature on conspiracy theories and all sorts of things to do with conspiracy theories. And so a question would be, for example, when – you know, are leaders like, say Bashir al-Assad who says, 'The external conspiracy is clear to everybody' when he was talking about, you know, the events of the Arab Spring and what's gone on in Syria. And all of these people constantly evoke conspiracy theories and so on. You know, are they – here we have Prime Minister Erdoğan responding to the Gezi Park protests again, you know, talking about conspiracy theories and domestic collaborators and all this kind of stuff. You know, there are so many, it's absurd. When you look at the – [0:26:13] you look at the thing, they're all obsessed with conspiracy theories, or they say they are. And they could just be doing it because it resonates with the Middle-

East street, and indeed it probably does. So in 2011, for example, *Q* found no Muslim country where over 30% believed Arabs carried out the 9/11 attacks. Only 9% of Turks and 12% of Pakistanis believe that. OK? So clearly they don't think – they clearly believe there was some kind of conspiracy theory. It could be the Erdoğan and Assad and so on are just expediently using the idea of conspiracy theories to sort of resonate with their populations. But we do have one good source of evidence as to whether leaders are expedient or being earnest in their belief in conspiracy theories, and that is that – So Saddam Hussein, he recorded all of his conversations. So just as Nixon did in the White House, you know, installed by the Kennedy administration – recorded his conversations, Saddam Hussein had all his conversations recorded. So you can hear – you can listen to – I think from the [0:27:20] from the late 70s, you can hear everything that went on in Saddam Hussein's office. So his end of telephone conversations and chit-chats where he was talking about killing people and stuff. All sorts of fascinating bits. And anyway, the long and the short of it is that, even in the privacy of his own office, Saddam Hussein did basically believe there was an international Jewish conspiracy. So doesn't seem likely that he was being expedient or saying it 'cause he thought people would like it; it's that he basically thought – he really thinks, or thought, seeing as he's dead (unless that's a conspiracy) that there was an international Jewish conspiracy. And so, for example, there's a great book that has [parts of this - 0:27:58] and they're releasing the tapes. I think they've got up to the late 90s now, so [0:28:06] back in 1981, 'It is Zionism. It is Zionism that is guiding them.' This is Iranians, apparently. Bear in mind by this point it's the Islamic Republic of Iran. So the Islamic Republic of Iran is being guided by the Israelis – or by the Zionists. 'The Zioinsm is taking the Iranians by the hand and introducing them to each part one by one, channel by channel—' and so on and so on and so on. And then he even talks about conspiracy and goes off on a big conspiracy rant.

So the point is that here we have something that is a psychological idea for which we have a wide variety of different sources of evidence – the fact that people believe in conspiracy theories and so on, and here we have at least

one source of evidence that, for example, at least one leader really does believe in conspiracy theories and he's not just using it as something expedient.

OK, so social motivations: so I've talked about prediction error and another key set of findings from neuroscience over the past 20 years or so is related to social motivations, and I'm going to talk about three briefly today. So the first of these is fairness, the second relates to accommodative conciliatory gestures, which Nick and spent this morning talking about and writing about together, which was great fun, and I'm also going to talk about in group/out group.

OK, so the first thing to say is that humans pay high costs to reject unfairness, and that this is based on our biology. OK? So I'm going to give an example. So I'm going to play a game with Tim. I'm going to hypothetically play a game with Tim. So imagine I have – I'm given £20 and I can decide how to split that – or I'm given £10 and I'm deciding how to split that with you. OK? So I can decide how to split it with you and I can offer you, for example, a 9/1 split. So £9 for me and £1 for you. OK? [0:29:59] £10 that I've received, by magic, and you can say either, 'I accept that', in which case I get £9 and you get £1; or you can say, 'I reject that', in which case we both get nothing. OK? And what people tend to do is they tend to say, 'I can't believe you're taking £9 for yourself and giving me £1. That's so unfair. I'm going to reject it.' And people reject the money. OK? So that happens about – people reject offers of less than a quarter about half the time. OK? People tend to reject that. Unfair offers. Unfair [0:30:35]. OK? But, of course, that's bonkers. So if you are rejecting an unfair offer, what you're actually saying is – what Tim would actually be saying, and I didn't give Tim, you see, the right to answer, but if he'd rejected it I would have said, 'Tim, you're bonkers, because your choice is between £1 or £0. OK? You shouldn't care about what I'm getting. You should care only about what you're getting. But that's not what human beings tend to do. They tend to care about these social motivations that also weigh on their decisions. And so we know that – so here's just from a paper I

published a few years ago: we know that the exact amount of an equality in these types of gains is carefully encoded within the brain (and I can talk about that if people are interested) and we also know even that monkeys will reject inequality so – and I've played this before, but it's so fun, this, I can't resist it.

So imagine this is Tim now. [Laughter] So here we have the monkey. See, the monkey is given a little task with a pebble and he gets given a piece of cucumber, and it's like, 'Oh, that's nice. Oh.' Then the other monkey does the same task with a pebble and he gets given tasty grape. So this monkey's like, 'Oh, right.' So then this monkey does the same task for a bit of – and he gets given a bit of cucumber and he's like…. [Laughter] 'How—? What are you doing? To the other guy? Look, this is outrageous. The other guy got a tasty grape and I just got this…!' Look, he's banging his little hand on the thing. Tim. No, not really, Tim. I'm sure you'd be much more elegant and sophisticated. You'd write a really angry email. Only joking. And so basically – and so the gist is that – look, there you go. He doesn't— [Laughter] And so the gist of this is – and this is a great video. I'm sure it's probably got a million views or whatever on YouTube. And the point here is simply that that – these capuchin monkeys, they like cucumber. They like cucumber. There's nothing wrong – they'd much rather have cucumber than no cucumber, but when they see the other monkey getting the grape, they're like, 'Why am I getting a bloody cucumber when the other guy's getting a grape?' And it's not – they like it. They like cucumber. I stress again, they like cucumber but they're prepared to pay a cost when they see that the other guy has got a tastier grape instead.

And then just to say that they've played a lot of these types of games up the Amazon and in Papua New Guinea and so on and seen similar sorts of behaviours. And so, now, how we think about that in the context of deterrents…so here we have the Deterrent Operating and Joint Operating Concept, which is a key US doctrinal document that describes how they think about deterrents. And the central concept, or the core concept of the DOJOC is essentially this, which is that you have a choice, so here you are, and you

can choose. You can either act or show restraint, and there are costs and benefits associated with acting and there are costs and benefits associated with restraint, and the idea is – this is the adversary's decision calculus, and the idea is that you influence the adversary's decision calculus so that you get them to do restraint rather than act. OK? So they have a little see-saw picture but it's essentially exactly the same as that. I just put it in that format 'cause it looks a bit nicer. OK? So this is how people think about deterrents. And this is actually how – [0:34:00], they use a very similar basic idea.

And so now let's think about what that would predict for the ultimatum [0:34:07] terrible predicting how humans actually behave. So here, for example, we had the ultimatum game, and we can put in a similar sort of format. I wish I had a pointer, so I apologise if I [0:34:17] stuff. But here they can either choose to reject or accept, and there are costs and benefits associated with rejecting and there are costs and benefits associated with accepting. And if you only thought – so here, for a 9/1 split, for example, if you only considered the material costs and benefits, then you just cannot predict how human beings will actually behave in the ultimatum game. OK? You have to add in the social motivation of unfairness in order to predict how human beings – to understand how human beings are going to react in the ultimatum game, for example.

And we can see this across a variety of different historical cases, or it's a plausible, possible explanation across a wide variety of historical cases. So I'm sure that this isn't going to work – the audio, but here we have Javad Zarif, the Iranian foreign minister, giving a – this was his YouTube video at the end of 2013 when he was trying to – when they were doing the first nuclear talks. Yeah, I told you it wouldn't work. Right, I'll start again. Oh, that's irritating.

[Video plays:] *—negotiable. Is there a price tag? Imagine being told that you cannot do what everyone else is doing – what everyone else is allowed to do. Do you back down? Would you relent, or would you stand your ground?*

So basically what he's saying is that you're being told you can't do what everyone else is doing. Would you back down or would you bear costs to reject that unfairness? OK? So these are his words. This was, you know, produced in English and it was basically intended for a western audience, and indeed, you know…it's very hard to explain otherwise. So, you know, they've basically – the costs are estimated – this was even about 18 months ago. The costs were estimated in the Iranian nuclear programme in terms of loss of investment or revenue – they've borne costs of, say, $100 billion, OK? For this nuclear programme. And it's very hard to explain that nuclear programme in terms of any kind of material benefit it brings itself, so only – actually, 14 nations enrich uranium, and the Iranian nuclear programme is dreadful. It's a dreadful programme. Not only is it built in a really dangerous bit of the world, but it also takes one year to enrich as much uranium as Europe's top facility produces in five hours. It's not a good programme. You wouldn't buy that programme if you had that kind of money. So it's not for material reasons. Now, I think you can then say, 'OK, is this just because they want to build a nuclear weapon?' and that's a perfectly plausible explanation, but another also plausible explanation is that there is a rejection of the unfairness to other countries (for example, Israel; for example, Pakistan [0:37:09] have done these things, and India, for example) and they're rejecting that unfairness. And even if they don't necessarily believe it themselves and we don't know whether Javad Zarif or the [0:37:21] are being earnest or expedient, but it is a political – it resonates with the population, and that is a political reality that's based in human psychology. And we can also see that, back in the early 1950s, basically the British offered Mossaddegh – offered the Iranians a 9/1 split – actually, it is beautiful 'cause it is a 9/1 split in that they offered – we were going to keep – by 'we', the British were going to keep 90% of the profits of BP, the Anglo-Iranian Oil Company and they were going to give 10% of the profits to the Iranians. And the Iranians said, 'No, that's a terribly unfair split.' And then, in addition to that, the Americans then gave Saudi Aramco a 50/50 split, which made it look even worse. And basically, despite the fact that Mossasddegh was told that they would be crippled economically if they rejected the offer, they rejected the offer, they paid huge costs to do so

because of the British embargo, and then of course, as we know, a couple of years later there was the MI6/CIA coup.

And we also see how these things factor into Chinese decision-making. So there are the – there is a long history of what they term 'unequal treaties', so these were treaties that were signed unfairly or forced upon China unfairly by a variety of Western powers and Japan, and this – because they were these unequal treaties – they're called unequal treaties – because they were these unequal treaties, they're now entitled to restitution. And we see how this plays out in more modern conflicts. So back in 1969, there was a very significant border conflict between China and the Soviet Union in which scores died on each side and it got to the level of nuclear threats being made, so it was very significant. And basically this whole argument was about the revision of an unequal treaty The Russians had forced upon the Chinese back in the Century of Humiliation and it was actually over a set of meaningless, totally strategically unimportant islands in the middle of a river. So there's a set of islands in the middle of this river, and how do you divide these islands? OK, but the Russians basically said, 'We want all of the islands', and the Chinese said, 'No, no, we'll run the line through the middle of the river and we'll split them 50/50. And the Russians said no and then there was this whole to-ing and fro-ing about it, and so explicitly the actual bone of contention was over the splitting of these islands where the Russians wanted all the islands and the Chinese said, 'We want a 50/50 split of the islands.' They ended up being split 50/50, but obviously a long time later.

We also see that fairness shapes the form of events, so for example that they asked for an apology during the EP3 reconnaissance-plane incident in 2001, and I don't have time to go into that. And so we've seen how we can think about unfairness in US doctrine, but how does it fit in, for example, to Chinese doctrine? So in Chinese doctrine there's a – the Chinese published their doctrine not in doctrinal documents in the way that they do in the west but in a variety of different things, and so this is a key military textbook called *The Science of Military Strategy*, published in English, and this is a quote taken –

so this is the first chapter after the introduction, where they describe how the international system works and what motivates states. And so they talk about 'national interest is an aggregate of objectively physical requirements and spiritual requirements; national interest is the cardinal basis to determine the alignment of the state's military strategy'. And then the outline: what does it – of what is national interest comprised? And there are six things. So, as you see, [0:40:56], territory, security, blah-de-blah, and the sixth is national dignity. So they explicitly have in there the concept of national dignity as one of their core components of national interest. And they define it. 'National dignity means the states deserve status and prestige as an intangible national interest. National dignity as manifested by states deserve [0:41:16] and quality in international contact. Superficially loss of national dignity is an emotional national humiliation instead of material damage but the loss can gravely harm national security and developments.'

So it's just to illustrate that while the US doesn't include these things in their military doctrine, the Chinese do include these things in their military doctrine, and in fact the Chinese have a much more psychological take on a lot of international conflict-management.

And I'm not going to talk – oh, I'm not going to talk about that. Or policy recommendations. I can go through these 'cause they're of no interest to anyone here, so it's fine. I'm only joking. OK.

And just to say that this also – thinking about some of the social tasks – social motivations also is interesting when you think about nature and nurture. So as I said before, we can think about differences within cultures and differences between cultures. So what about differences between cultures? So the first thing to say is that, actually, people do reject unfairness and this is pretty consistent across different cultures, so this is just an – this is a thing where they went to lots of tribes all around, you know, Papua New Guinea and what-not, and basically the gist is that the bigger the blob, the more people rejected offers and basically people reject unfair offers the more – the

smaller they are. OK? So if I offer you a 9/1 split, you're more likely to reject that than if I offer you a 7/3 split. OK? And that's seen across a wide variety of different cultures. But then you can also see very interesting things which you don't necessarily know, and this is why empirical work is important. So this isn't specifically related to unfairness; this is related to something called The Public Goods Game, which is basically – so the idea – so what they did is they looked at when people free ride. OK? So free-riding is when there's a common good and I take stuff from the common good but I don't give back. OK? So I, like, you know, drive a really gas-guzzling car and emit loads of pollution or whatever and I don't, you know, contribute to the common pot of carbon reductions or whatever. And basically what they did is they got people to play games across a wide variety of different countries and they basically – this was a science [0:43:19] and they published a series of nature and science papers for extremely august journals, and once they'd shown that basically people are prepared to pay money to publish free-riders. OK? Even though that's not a rational thing to do. you shouldn't care if someone free-rides – or you should care but you wouldn't pay to punish them. People do. They pay to punish free-riders. And they showed this, and this had all been done in Northern Europe and Nottingham, just up the road, and in Germany and so on, and so they were like, 'OK, well, this is like a – just a key thing about human nature.' But then they did it across many different parts of the world, OK? Including in China. And so, for example, in China, it's basically you get the same thing – that they – this is all the green lines. This is the – punishment of free-riders is all the green lines. And they basically showed that people are prepared to pay money to punish free-riders. But then they looked at a variety of other cultures, and basically what they discovered is that down here – OK, yes, it's true. Everybody – all around the world people punish free-riders – that's true. But in addition to that, over here, what you see is, for example (and I'm so sorry, Theresa, and I always mention this [0:44:26]) but in Greece, for example, they pay money to punish people who actually contribute more than average to the common pot. So these are people who are, like, going out and doing more digging on the common irrigation network or whatever, who are like going out and helping more than

other people. These are people who are contributing more to the common pot than average, and in Greece people will pay to punish those people.

F Because they're suspicious.

Yeah, exactly. What are they up to, those helpful guys? And the thing is, when you say this in northern Europe, people just – it's, like, mind-blowing, but – and nobody can really explain – I don't know what happens in Hungary. I'm sorry. And – but basically – so the point here is just that we know that – you know, so things can be common across much of the world, and that's true (say, for example, the [0:45:22] unfairness) but there may well be interesting additions – I'm not saying that everything is the same all across the world, and how things manifest may differ according to different cultures, and you may well see extra phenomena such as, for example, the Greek and in fact much of the Muslim world where there was a lot of this anti-social punishment in addition to the pro-social punishment. And just to say this was published before the financial – just before the financial crisis, so that can't explain things in the Greek case.

OK, and just because we're going to talk about conciliatory gestures, I'll just say that a conciliatory gesture is where you make some kind of unilateral gesture to stop escalation or to de-escalate problems. OK? And what we know from a wide variety of different behavioural experiments where you can look at how cooperation or competition evolves over time, we know that basically humans make – actively make conciliatory gestures to manage escalation. OK? That's what we do in the laboratory and just to say we can also see that in historical cases, so for – you know, very many crises do not escalate to limited wars and very many limited wars remain limited, and so, for example, even the toughest, you know, cold warriors – so even here we have John Foster Dulles, who was an extremely tough guy, but even he and Zhou Enlai in 1958, for example, in the 1958 Taiwan crisis, between the two sides, they both made and received conciliatory gestures in order to control escalation and de-escalate that crisis. So it's just a point here that humans –

it's not – we don't prevent escalation by just not having escalatory factors in play; we actively manage escalation by using conciliatory gestures, and we can see that across historical cases, and you can enhance conciliatory gestures even by using prediction error; for example, Anwar Sadat, back in 1977, when he made the highly unexpected gesture of offering to go to the Israeli Knesset and speak there, so you can enhance your conciliatory gestures using the neural phenomenon of prediction error to increase the impact. And we see that during recent events with Iran.

And so what I'd say, just a – two policy recommendations with respect to conciliatory gestures: the first thing is that you should expect [0:47:53] conciliatory gestures as natural and common. OK? Which is not the case amongst a – the way a lot of people think about these things. And this is including as part of escalation management, OK? So if you read a lot of the way that people particularly in the US think about escalation management, they don't have active – the making of active – the active making of conciliatory gestures as part of escalation management, which I think is a significant problem. And you can also make conciliatory gestures more effective by increasing their unexpectedness, using prediction error.

And I don't really have time to say much about in groups and out groups except to say that this is an – like, so this is very old psychology. Like, you know, 50/60-year-old psychology that people naturally split into in groups and out groups, but it boggles my mind that people don't think about this type of thing when they try and understand the world, so especially – I'm sorry, Stefan, but a lot of Germans I know can't talk about nationalism, for example. Even the word 'nationalism' makes them feel uncomfortable, and they're like, 'Oh, no, [0:48:53] national identity.' But you cannot understand the world unless you – I would suggest unless you understand nationalism or, for example – which is basically an expression of in group/out group. It's a basic biological drive that conditions the way that the international political system can be. And manifests as nationalism within Europe, for example. I mean, the Scots may, you know, argue that it's all about fairness and equality but, I

mean, I think a lot of it's just old-fashioned nationalism. The Russians, at the moment – Vladimir Putin, I think you could make a perfectly reasonable case that it's old-fashioned nationalism. Narendra Modi is a Hindu nationalist. The National Yoga Day – I know Nick's going to go and do yoga tomorrow. National Yoga Day, for example: that, actually, if you know the ins and outs of it, is actually a Hindu nationalist day about trying to enforce ideas of what it means to be a Hindu, and there's all sorts of things about – for example, the German embassy, they were very upset when the – they refused free German money to teach German in India because they wanted to teach Sanskrit. And so, you know, you have to see – like, you can't understand Modi unless you understand basically that he's a Hindu nationalist. And of course everybody knows about – or should know about Chinese nationalism. You know, the Chinese Communist Party is no longer communist, so what do you have left in their name? They're the Chinese Party. That is the gist, and they are – you know, nationalism is an absolutely core component of what they do. And you need nationalism, so did a project recently on Israel/Palestine and, you know, if you're thinking about how the West Bank and Gaza are going to act as a unified government, you basically have to think about nationalism or national unity or national identity or whatever you want to say.

And then in terms of trust, you know, I don't agree with everything Nick thinks about trust, but even – you know, regardless of that, trust is fundamentally a psychological phenomenon that needs to overcome this in-group/out-group problem, and in fact we have – and this is where the randomised, controlled trials and so on come in, is that – so when we did the project recently on Israel/Palestine, one of the things is we know a lot about how you can build trust between different groups – just practical interventions that you can use, and so we can use those and we can use those to think, for example, about how different aspects of the Israeli and Palestinian security forces can coordinate and build trust and between senior officers and so on and so forth. But, again, you need that because you have to fight against this in-group/out-group psychological factor that shapes the international and national political scenes.

So, social motivations: there's fairness, which can limit deterrence/cause escalation; [0:51:33] conciliatory gestures are natural, common and enhanced using prediction error. (Woops!) In group/out group are critical to understanding things like nationalism and trust. And with respect to social motivations, the take-home message is that forecasting the others' decisions is necessary to conduct deterrence operations or manage crises and escalation, and social motivations help identify what drives both adversaries and allies. And so if you were sitting there trying – like, trying to actually do policy, this would actually be genuinely useful. Again, it gives you a set of simple – a checklist for empathising with both adversaries and allies – those that you wish to seek to influence.

So, in summary, two sets of insights: one prediction error and the other related to social motivations. And we can think about four rules for using neuroscience: are we sure enough of the neuroscience? And we are in both cases. Does it matter in the real world? And I think we can show for example historical cases across contexts. Is it worth adding? In both cases we can make things simple and operationalisable out of these insights. And then, finally, why the neuroscience? And as I said before, we can think about both consilience and universalism.

Thank you very much.

END OF RECORDING