3._Collective_Intelligence_-_Collective_Decision_Making_and_Reasoning

I'm Geoff Mulgan from Nesta, and I'm going to give a little bit of an overview on some thoughts on theory and collective intelligence and some examples while your blood rushes to your stomachs and you quietly digest and hopefully don't fall asleep, which is what usually happens in these slots. We're going to email round this afternoon a paper which we put on the Nesta website yesterday on meetings. And I'll be very interested in many of your critical responses to our attempt at gathering some of the science, the experience, the examples of different meeting formats from conferences to parliaments to boardrooms. And this was done partly out of frustration. We couldn't find very good material out there to guide us on how we should organise meetings. And indeed, we often went to meetings which seemed to be a bit of a waste of people's time, knowledge, et cetera. So critical responses to that paper, very welcome, but I'm not going to say anything about it now.

What I am going to talk a bit about I hope builds on this morning's discussion. I guess I first became interested in-- well, for me the questions of collective intelligence, prompted by the sort of things you see on the screen. So on the one hand, IPCC is probably one of the world's most impressive attempts at orchestrating the brainpower of the world around a really important problem, climate change. Thousands of scientists in working groups producing scenarios, models, a number of different sciences involved, trying to distill guidance.

There you see the last meeting. Cop21 meets in Paris later this autumn. That was Copenhagen, 2009. Almost a parody of the opposite of collective intelligence. The world's most powerful people sitting in rooms, paper strewn everywhere, unable to make any decisions, unable to make any progress. And one of the questions for me is, if you were designing a proper IPCC, and indeed a Cop21, what would you do? Using the brain power in this room, where would you start?

And the other prompt is we are surrounded by all these tools which organise intelligence in different ways. Watson now being used for all sorts of things, not just playing Jeopardy on American TV, but also medical diagnostics. Here in London, we have Google DeepMind with a set of deep learning technologies able to learn as they go. And of course, we're all surrounded by sensors and semantic analyses, and so on. So the question then which flows from all of these is how do we use them? How do we help groups to think more effectively, both through people coming together and the combination of people and technologies?

And as some of you have said, there are really interesting live examples from Cambridge. The "Polymath" blog has now been going quite a few years posting very difficult maths problems, organising a community of solvers, and apparently getting to solutions faster, of better quality, which generate better follow-on questions than the more traditional model of sitting in your garden and writing papers for refereed journals. We have fields like chess, lots of experimentation with large groups advising on chess moves. And again, some evidence, uneven about the quality of decision making that can generate. And there are many, many other examples.

But if you pose this question, the obvious issue is where do you look for the appropriate theories, the appropriate knowledge, the appropriate evidence to guide you to understand which of these is likely to be most effective or not? What advice would any of you give to the organisers in Paris this autumn? A year and a half ago, I published an academic journal article attempting a bit of some versions of an answer. And I'm just going to give two tiny elements of theory from that and then a few examples of practise. This is in Philosophy & Technology journal, which probably no one reads. But that's true of many academic journals.

And there are two basic frameworks which I try to use, and which I've found useful. One was quite crudely trying to break down some of the elements of collective intelligence and see how they function within the setting of-- it could be a small firm, it could be a university, it could be a charity or a government. How do they organise observation-- seeing, hearing, collecting data, and so on? How, prior to that, they generate operating models of the world into which to put that data? How do they attend, how do they analyse, how do they create new things which don't yet exist? How do they remember, how do they judge, and how do they exercise wisdom?

And even if you take a plumbing firm in a small town or university, it is actually quite easy to analyse their tools and methods for each of these functions. Usually, we find in almost any organisation great imbalances in terms of their ability to do some of these relative to others. The investment banks are a great example, which had invested enormous amounts of money in analysis and even memory, but almost nothing in wisdom and judgement, and therefore came rather badly unstuck. Technologies have completely transformed how we do observation, and some elements of analysis, but very little on creativity or again, judgement and wisdom and so on.

So this is one frame which I think is quite useful for disaggregating the things we're discussing. And the other, just try to think about the loops of intelligence which you see, again, in any real group, whether it's a family, a company a city, a government, or a firm, where we have at the lowest level fairly routine automated thinking. Perhaps if you're a professor marking papers, you know how to do it. Your mind can probably do it almost on automatic, or walking down the street, driving, all sorts of things.

And then you hit a trigger. Something doesn't quite work. Something's unexpected, something's surprising, and other kinds of thought come in. You take corrective actions in response to surprises. Further triggers force you to create new categories to understand the phenomena you're trying to understand. And this is one thing the human brain is very good at is category generation, computers much less good at. And then other things trigger metacorrective new ways of designing the whole intelligent system of, again, it could be a firm, a university, a government, and so on.

And each of these levels requires significantly more energy, significantly more time, and more complexity. And yet we recognise as intelligent groups which can do all of those things. Being hyper intelligent at routinised, automated, predictive functions, we do not think of as very intelligent, even though many institutions do that.

The Army was being talked about earlier-- William McNeill's Keeping Together in Time. Armies are very good at that sort of routine, automatic function, as are, indeed, Premier

League footballers if they've been getting on with each other. But then, when something goes wrong, you need a completely different kind of battle strategy. You need a very different kind of thinking to adapt. So that's, again, very crude little bit of theory, but one which I think is quite useful for thinking about real life examples.

So where to take that then? One of the things we've been trying to do here, just gather examples of really good and apparently successful collective intelligence under these different headings. So creativity, there are actually lots of examples now of tools for harnessing creative problem solving on a very large scale.

This is Kaggle, which does it for data scientists, posing problems and allowing anyone to come up with solutions. One of their first solvers was a teenager in Mongolia who would never normally have been found as a solver. Most of the work is still quite individualised, but there's a whole host of platforms bringing teams together. And indeed, our Centre for Challenge Prizes here at Nesta structures both the posing of problems, individual solutions, but also bringing together teams, often of people who haven't met, to work collaboratively on problem solving. And I think these are much more effective for certain kinds of problem solving than traditional methods.

Then there's observation. We've, again, both funded and published a lot on new methods of seeing the world. This one is Peta Jakarta, citizen-generated flood data to help a city know itself, to know where there are problems, where there are resources. And every time there's a disaster now, from Haiti to Nepal, a lot of these citizen-generated data services essentially orchestrate collective intelligence far beyond what the state can do. And if you're interested, we publish lots of examples of these, which seem to work very cheap, generally quite effective at improving observation.

Analysis is a bit more complicated. Earlier this year, we published a thing looking at social needs and problems through citizens' advice bureaus and new ways of combining patterns of what are making people unhappy, are causing problems, and then analysing those data in different ways to help responses. And this example, which will be being shown in London in a few days' time, is a New York predictive algorithm for fires. 60 factors gathered in an algorithm to predict which buildings will burn down, so the fire service can target those buildings to prevent fires, rather than going to put the fires out. Predictive algorithms widely used now in primary health care probation, all sorts of fields, and clearly enhancing collective intelligence of all kinds, especially if they then get a data loop back and learn over time.

Memory is another really interesting one. We host here a thing called the Alliance for Useful Evidence, which gathers together with what work centres-- the memory of fields like teaching or policing. This is a very different one from last year for the centenary of the First World War, which was Historypin, which tried to aggregate on online platforms memories, in terms of writing or photographs or film of the First World War. Completely different approach to a museum, but again it's a collective intelligence approach to collective memory. And there are a lot of other examples of that kind.

And then judgement. We talked earlier today about democracy and politics. The dissent programme, which we've been running here from Nesta, has been trialling these democracy civic participation tools in Barcelona, Reykjavik, Helsinki, Madrid. And it is possible to see

in each of these a very different way in which the citizens of a city can take part in nominating problems, generating solutions, commenting on policy, seeing things through into implementation. And in political parties like Podemos, we see a very different way of organising the collective intelligence of members.

We're not short of examples now. Almost none of these existed four or five years ago. What we are short of is really good scrutiny, analysis, and interpretation of their elements of intelligence, and of their first, second, and third loops, when they hit surprises and things don't work and so on.

The other space we're trying to look at is hybrids. This is one which, again, we host here at Nesta and is formally launching later this year called Dementia Citizens, which will bring together, in time, hundreds of thousands of people with dementia and their carers to pool real time data through wearables on what is happening in their lives, in order to run real time experiments on things like music or dance or diet with control groups to turn a large community of people with dementia into a knowledge-generating community, a true collective intelligence. And that's a hybrid of quite a few of the elements I've already described. And a similar thing underway on Parkinson's at the moment.

Again, really exciting, lots of practise. But the practise, I think, is going rather ahead of the theory. So the final thing I really wanted to pose was, what kind of theory do we need to help the people designing these things, or interpreting them, or improving them, or assessing them? Do we need shared theory? Do we need shared concepts? Or is the very traditional disciplinary approach the way forward? Do we need multi-disciplinary collective intelligence research, bi-disciplinary, or are we looking for the creation of new disciplines?

And if you pose the question in the way that I've posed it there, which is a very long sentence, what structures, processes, and cultures, and what combinations of people, artefacts, and technology aid what types of problem solving, creation, memory, et cetera? I'm not clear where to look for for theories, or even combinations of theory, to answer that question. And that's, in a way, I think the challenge for today.

My guess is there's an element to this which will be not unlike economics. There are different costs and returns to different kinds of cognition, from the automated to the very complex. We've already talked a lot about what psychology can offer in terms of dynamics of interaction, impairment, enhancement, and so on. All sorts of schools in computer science understanding some kinds of reasoning, the new deep learning methods, but also analogical thinking, which was not something computers used to be good at, but are becoming perhaps better at.

There's a whole tradition in anthropology. Mary Douglas, not far from, here a pioneer of grid group theory, I think a very powerful framework for understanding how institutions think, but almost wholly unused by psychologists. And I've yet to meet anyone in computer science who's even heard of it. And there's a lot of theory from anthropology which could be integrated. And I guess what really interests me is how we combine theory with, as it were, experimentation in the wild, in real life.

And the reason for publishing the paper on meetings is that every time we have a meeting, we are, in a sense, running a live experiment in collective intelligence. So how can we get a more conscious, reflexive approach to understanding what works, what doesn't, for different kinds of meeting, whether it is problem solving, decision making, knowledge sharing, et cetera? And how can we, in a sense, proliferate these experiments, whether the sort of work [? Stefan ?] is doing with patients' organisations, working with city governments to help cities think, and so on, but looping back into a body of well developed theory?

So I hope that's all straightforward. I hope I've annoyed everyone from every discipline so far by what I've said to avoid you falling asleep. We are now going to get on to the more meaty stuff. And I do hope, as I turn over to Hugo Mercier, that we will have arguments, and well-structured arguments. And Hugo will talk about why argumentation is vital to reasoning, is deep in the human condition, et cetera.

But I'm still not quite clear what are the most productive arguments we should be having here. What are the camps, as it were? And most fields of knowledge advance by having well-structured arguments with two sides. And hopefully, in the end, one side wins. And one of my challenges to all of you is what will be the most productive argument in this field? And that's maybe a thing we can come back to ask Hugo [? spoken. ?] Over to you.

We'll talk to you today only about reasoning. So there is a slide at the beginning about intuitions and reasoning, because I am a cognitive psychologist. And so what we do is we chop up the mind in little bits and try to study them one by one. And I'm in charge of the reasoning bits, that's the thing I do, which is not so bad. And I just want to give you an idea of what I mean by reasoning, because many people have different views of what reasoning might be.

And the view I hold is relatively commonsensical, in that what I call reasoning is basically using reasons, when you produce reasons, when you evaluate reasons. So for instance, if you're deciding what car to buy, and you're pondering the pros and the cons of various cars, you're thinking of reasons for buying such-and-such car. So there you're reasoning. If you're thinking who to vote for, if you're comparing the platforms of various candidates, thinking of reasons why you should vote for one person rather than another, you're reasoning. If you're arguing with someone, if you're exchanging reasons with someone-- arguing in the positive sense of exchanging reasons, not in the sense of having a shouting match-- you're reasoning as well.

By contrast, intuitions are everything else. So when you see someone, you have an intuition about how trustworthy or competent that person is. When you see some food, you have an intuition about how tasty it's going to be. When you see the weather outside, you have an intuition about whether it's going to rain. Most of our decisions and behaviours are driven by intuitions. And there is nothing wrong with that. Reasoning is just this one very specific bit of the mind that is in charge of using reasons.

And the thing I'm going to talk to you about today in particular is why we reason. Because if we adopt this definition of reasoning, we're the only animals to reason. Other animals don't think of reasons why they do such-and-such things. They just do whatever they do, and they're fine with that. So why do we reason?

By far the most popular theory can be called the individualist view of reasoning. And I think it's been held fairly consistently by philosophers and psychologists alike, and also by most people. And the view is that when you're reasoning, it is in order to improve on your decisions, to improve on your beliefs. So this is Rene Descartes on the left and [? Denny ?] [? Kenman ?] on the right, two illustrious representatives of this school. So I think at least that's what we hope to achieve when we're reasoning on our own. And that's a fairly intuitive view.

And so in order to defend this view, psychologists of reasoning have devised problems, such as the one I gave you earlier. So this particular problem was introduced by someone named [? Leveque, ?] so we named it that. I'm not going to bother you again with it, except that I will tell you the answer, which is not that. So in this problem-- and that's why psychologists like it-- there's a relatively intuitive answer that most people think of when they read the problem, which is we can't tell, and which happens to be wrong.

The correct answer is yes. So I'm going to assume that not everybody has talked with everybody, and so not everybody has figured out that yes is the right answer. Why yes? Because Linda is either married or not married. If Linda is married, then the statement is true, because Linda is looking at Henry, who is not married. If Linda is not married, then the statement is true, because Peter, who is married, is looking at her.

I'm going to go over that again, because otherwise everybody is going to spend the rest of the talk thinking about this. That's happened before. So Linda can be only either married or not married. There can be no other states. If she is married, the statement is true, because she's looking at Henry, who is not married. So you have Linda, married, looking at Henry, not married. If Linda is not married, then you have Peter looking at Linda. Peter is married, Linda is not married, the statement is true. So the statement is true in all of the two both possible worlds, and therefore the correct answer is yes.

Everybody who really gets the right answer does it thanks to reasoning. And they can tell you why yes is the right answer. They have reasons to justify their answer. And so it seems as if we have a perfect illustration of this individualistic view of reasoning. So you have a mistaken intuition, and then thanks to your solitary reasoning, you can realise this intuition is misguided and get at the right answer. And so you get to this kind of trivially true conclusion, which is that reasoning, indeed, can help the lone reasoner reach better decisions and arrive at better beliefs.

The word that people tend to forget is the "can." and what I mean by this is that most people don't. So the right answer in this problem, which requires no knowledge of logic, no special skills, everybody can do it, is about 10% correct answer, 20% on a good day. It's pretty low. And you have consistently similar answers for all sorts of problems that have the same property of having an intuitive but wrong answer. And so it seems that reasoning can indeed help the lone reasoner correct mistaken intuitions and arrive at better beliefs, but in most cases it doesn't.

And the reason why it doesn't is actually even more of an indictment of reasoning. Because you could say, well, reasoning is trying to get at the right answer. But there's a way psychologists say is, we have limited working memory, which is a way of saying that people are a bit stupid. And so we try to get at the right answer, but reasoning just can't do it. But

actually, when we look at what reasoning does when people are facing just about any task-- not only task like this one-- is they only, or overwhelmingly, think of reasons why their intuition is right.

So some of you might have thought, well, we don't know if the nice marriage, we don't have enough information, the answer depends on Linda's status. And so instead of doing what reasoning should be doing if it was really trying to help you eliminate misguided intuitions, which is impartially look at the different answers and critically evaluate the reasons that you find, reasoning only finds reasons supporting your initial intuition, whether it's right or wrong. And it's not even very careful about the reasons it generates, because all of these reasons are wrong, because they support a logically invalid answer. And very few people realise this.

And so what's striking is not so much that reasoning fails, because you could think, well the task is hard, we have limited abilities, blah blah blah. It's that reasoning fails because it does the exact opposite of what you would like it to do if it was really aiming at improving individual cognition. So I think the mismatch-- I mean, I've never seen such a striking mismatch between what something is supposed to do and what its features are. It couldn't get worse.

And so I think that this observation-- and I mean, the fact that reasoning is this what we call the my side bias, this confirmation bias, this tendency to find arguments that support our point of view, is consensually admitted among psychologists, and yet very few have drawn the conclusion that maybe their view of the function of reasoning should be rethought.

And so what kind of alternatives could we suggest? Sorry, could you give me an idea of the time at some point? Because I didn't check when I started. So the alternative that Dan Sperber and I have been working on-- we have called it the argumentative theory of reasoning. And the gist of it is that the main function of reasoning would be to argue.

So then again, not to argue in the sense of having a shouting match, but to argue in the sense of exchanging arguments. So we disagree about something. I'm going to give you reasons for my point of view, you're going to evaluate these reasons. You're going to give me reasons for your point of view, and I'm going to evaluate these reasons. So I'd be happy-- I won't have time now, but I'd be happy in the questions to say more about why we think that's a valid evolutionary story, and in particular how that relates to potential limits on how confidence can help us communicate.

And so I'm going to draw a series of predictions from this theory, the first one being that, counter-intuitively enough, maybe, if that is the function of reasoning, then we should have a my side bias. Because when you produce arguments, if the function of the mechanism that does this is to convince others, then you want it to have a my side bias. You want it to find arguments that will convince someone else. You're not going to win a debate by finding arguments for the other person's point of view. So the my side bias, far from being a bug or a flaw of reasoning, is actually a design feature.

The second prediction is that reasoning, we think, should be selectively lazy. So when it comes to the way people produce arguments, maybe they shouldn't bother too much. Because

when you reason in a discussion, if the first argument you produce doesn't work out, you can always give another argument. And you can actually benefit from the counter-arguments that the person will have given you, so you really address that specific person's counter-arguments.

By contrast, when you're evaluating someone else's arguments, you ought to be quite careful. You want to reject bad arguments so that you don't accept bad ideas, but you also want to accept good arguments. Otherwise, argumentation would be pointless. And so our prediction is that reasoning should be selectively lazy. It should not be really paying much attention to the arguments we're generating, and we should generate relatively weakish, generic arguments, at first anyway. But it should be careful to the arguments other people generate.

And so to test this, we did a really sneaky experiment, in which had three phases. In the first phase, people were asked not to reason at all. They were asked to rely on their intuitions. So we gave them a small reasoning problem. So the problem itself really doesn't matter that much. You tell people that in a shop, none of the apples are organic, and they are asked to draw one of the five possible conclusions. So let's imagine a participant who answers that it follows that some fruits are not organic, and they do this five times in a row. So then again, trying to answer as quickly as possible so that they give an intuitive answer.

And then we introduce the prediction of arguments. So we're going to ask people to produce arguments to justify or to explain their answers. So we remind people of what they have just answered, and we ask them, can you give us a reason for this answer? And they do so. We do this five times. And what we observe is in line with the predictions that the overwhelming majority of people stick to their initial answer, and that includes about half of people who got the wrong answer at first. So people just generate an argument that justifies their intuitive answer, whether it's right or wrong. And indeed, among those who do change their mind, they're not more likely to change their mind if they're wrong than if they're right. So solitary reasoning produces absolutely no benefits in this example.

And then we asked people to evaluate other people's arguments. And so then again, we give them the same problem-- so all of this is really one minute after the other. It's really all in succession. We give people the same problem, we remind them of the answer they have just given, and then we tell them, look, here is what someone else has said. They have given this answer, and they have given this argument. So the problem we have here is that we could have fudged the results. We could have fudged the experiment so that we give very strong arguments for good answers and very weak arguments for bad answers. So we could have cheated.

So what we did instead is for one of the problems-- sorry, obviously we can ask people if they want to change their mind after seeing the arguments. And we do this only four times. On the last problem, we lie. And so we tell people that they have answered something they have not answered. So we tell them, if you have answered some fruits are organic, we tell them that they have answered something else. And then we give them back their own answer and their own argument as if they were someone else's. And then we ask them, do you want to change your mind?

And so what happened there-- and we do this only one time, because we don't want too many people to catch up. More than half of the people did not realise they were evaluating as if it was someone else's the argument they had just typed one minute before. And so if you look at these people who did not realise they were evaluating actually their own argument, most of them rejected it. So the very argument they had typed as a good justification for their answer one minute before, now they think when it's someone else's argument, maybe it's not so good. So obviously that only happens to other participants, not to us.

And the thing there that was still nice is that people were still more likely to accept their own good arguments than their own bad arguments. So we do find this a symmetry in the laziness of reasoning. People were producing reasons without being very careful. They got more careful when it came to other people's arguments, or what they thought were other people's arguments. But they are still good at discriminating, so they're more likely to accept good arguments than bad arguments. And indeed, this is one of the predictions of the theory, that if reasoning evolved for arguing, then we should be quite good at evaluating other people's arguments.

So because I don't have that much time, I'm going to pass over argument from authority slide. A lot of studies show that I'm right. And I'm going to move on to a funnier-- I mean there's really no other way to transmit the message. This is really the outcome of the literature. All the experiments that have looked at this have found the same result. I've done that for my Ph.D. Believe me, I'd be boring you to death if I were to review all of that.

And so the first prediction is that if you put together on the one hand, our ability to produce arguments and to justify our points of view, and on the other hand, our ability to evaluate other people's arguments and to discriminate between good and bad arguments, you should get good things happening. Because everybody can defend their own answers, but then people who give bad arguments, the bad arguments are shot down, which they wouldn't be if people were reasoning on their own. And then you are exposed to arguments for views that you would not have been able to find. You would not have been motivated to find these arguments. And so things would work relatively smoothly.

So indeed, we have a series of experiments. And this is not new with us, but we've replicated and extended some of these findings. So to take a very simple experiment, we gave a large classroom the problem I gave you earlier. And first, people had to solve it on their own, and they got about 20% correct answers. That was a good day.

And then we put them in groups of about four people, not to break the other Dunbar's number. And we got about 60% correct answer, which means, actually, that every group in which at least one member had the right answer converges on the right answer. And so that is true even if you have one person who faces three people who all agree on the same wrong answer and are more confident. In some cases, the majority is more confident they're the person who is right, and yet she still always manages to convince everyone of the right answer. And when I say always, this is really like 95%. As far as psychology goes, this is not bad.

And so then we wanted to scale it up again a bit. And so then we took a classroom. So this is the classroom. Every square is a participant, a student. The black squares are just empty seats.

And so you have the physical layout of the classroom there. This is what happens after five minutes of solitary reasoning. People have been given the problem, they have thought about it on their own for five minutes. The red squares as you might imagine, are people who get it wrong. The green squares are people who get it right. And then we let them talk with each other. We let them talk with their neighbours, with eight neighbours at most. And every minute, we record who is thinking what. And this is what happens.

[AUDIO OUT]

So that was another good day. I'm not sure the data of the bad days. So that works really well. And what this shows, which is quite interesting, is that once you have been convinced by someone, you've understood the argument enough that you can convince someone else in turn. So the thing really spreads. So in some cases, it doesn't work. And just to give you a taste of what our poor psychologists have to face, one of the participants who did not want to accept the right answer told me that the answer was no, because if Linda is married, she's in love. That person was not married, I believe. And so if Linda is married she's in love. If she's in love, she's blind. Therefore, she cannot be looking at Henry.

[LAUGHTER]

And so after delivering the smartass prize of the day, I moved on. And so sometimes, you don't get 100% because you have people like this. It's kind of fun, but kind of ruins your data a little bit. But sometimes it does work out quite well.

And so on the whole, argumentation works, I think, surprisingly well. So this is kind of an extreme case, because on this type of problems, if you have one person with the right answer and she can argue with others, she will convince everyone all the time. Actually there is a real world analogue to this, which is mathematics. So in mathematics, this is how things work.

So the strongest example of this is probably Godel's incompleteness theorem. Many of you will probably know this. So when Godel introduced his theorem in the early 20th century, in the '20's, I guess, the whole of mathematics, really the goal was to find the foundations of mathematics. And Hilbert, Russell, people who are usually smart, usually influential, had devoted their lives, literally, to finding a foundation for mathematics. And then this nobody comes about and tells them that it's impossible. And in a matter of days-- or months, because they didn't have the internet-- everybody accepts it.

I mean, it's staggering the speed at which these things work in mathematics. And if you look at physics, things are a bit slower, because the standards of evidence are a bit lower, necessarily, but things still work really well. And then if you look at us, It's a little bit slower, but arguably that is still some relatively rapid progress.

And so argumentation works really well in that it allows the best answers to spread on logical and mathematical problems very well. It works very well as well on looser problems. So if you have inductive problems that don't have one right answer, but still have better answer than others, that works well. It works well in the workplace. There are a lot of studies of the efficiency of teams and whatnot. It works extraordinarily well in schools. And there are

literally hundreds of papers on collaborative or cooperative learning showing the benefits that argumentation can bring to learning in schools.

It works in the lab, so that is a very nice. Ethnographic work by Dunbar's namesake, Kevin Dunbar, done on how important meetings are to spreading good ideas in science. It actually even works in juries. Some people sometimes can have a relatively dim view of juries, but actually, it seems that jury deliberation works pretty well.

It works among citizens, so many of you here today will know about deliberative democracy. And people who have done experiments in deliberative democracy have found quite positive outcomes on the whole. So it's harder to tell who is right, because we don't know who is right, but usually there is progress in terms of people better understanding their perspective, better understanding other people's perspective, and converging on the middle ground. Things work pretty well.

And one last data I want to show you is kind of like a meta-experiment we ran in which we asked participants to do not exactly the task I gave you but something very similar to it, and then they had to evaluate how many people would get the right answer working on their own and working in small groups. And so what I'm going to show you here is the ratio of group to individual performance. So basically, if it's one, it means that people think that groups will do as well as individuals. And this is what people think.

So most people-- and we asked people in different cultures and different occupational groups-- people think that the ratio is about one. There is no advantage to groups whatsoever. If you ask a psychologists of reasoning, people who have spent their lives studying this kind of problems, they will tell you that the ratio is about two. And the real ratio in that case was about four. So people vastly underestimated the efficacy of argumentation in a case in which it's really well documented. And people-- especially in [? psychological ?] reasoning-- ought to know about it, and even they didn't know about it.

So the summary of today is that reasoning is for arguing. At least that's what we're claiming. Individual reasoning is quite a bit overrated, I think. And by contrast, argumentation is quite severely underrated. Thank you.

[APPLAUSE]

I hope I will be able to tell you what I want to tell you quickly and painlessly, because most of you will probably by now know almost everything that I want to tell you. I changed the title slightly from what was in the programme to "What Makes Collaboration Work" to address the question that was placed in the talk at the beginning of this session. And so I guess nobody needs to be motivated for listening to me at this point, because you have been listening to similar questions all the time. Collaboration is everywhere. This is an example of a Skype advert in a bank station in 2012. It says, "Why collaborate when you can work alone?" and the idea of the campaign is very obvious. Collaboration is good. Everybody should do it. We give you the option to do it freely and with anyone over the planet, so why not? Just use our product.

When we look at the history of social psychology and ask the question, should we collaborate at all, again I'm sure you'll now agree with me that there is a lot of evidence saying that yes, you should. Things even from Bible-- two heads are better than one. And also similar popular wisdom, saying too many cooks spoil the broth, and you shouldn't. And there is just as many old arguments as there is new ones. There is all sorts of ideas relating to madness of the crowds that you will see a little bit further forward. And there is, of course, Condorcet's jury theorem, which was mentioned a couple of times in the morning session.

I've just put-- you can almost say the exact phrasing of the theorem here. A group wishes to reach a decision by majority vote. And there is a correct answer, and each voter has a slightly better than chance of getting the answer right. Now if that probability is slightly better than flipping a coin, as the number of voters goes to infinity, the accuracy of the group decision becomes certainty. We also know that Francis Galton, from University College London, indeed, was one of the first people to actually make observations in line with Condorcet's jury theorem in his famous experiments. I'll not go through them either.

But my favourite case is the case of Charles Mackay and his great book, "Extraordinary Popular Delusions and the Madness of Crowds," which is still a bestseller on Amazon, by the way. And in his book, he chronicles the story of several important and very conspicuous cases of madnesses of the crowds, that up to that point, up to 1834, when he published the book, had been observed. His favourite cases are financial bubbles and Middle Eastern peace process, by which I mean the Crusades at that point.

And his insight is very helpful. What he says is that social influence is often at work. There is an inherent assumption in Condorcet's jury theorem, and that inherent assumption is the independence of individual opinions. Whereas social influence like campaigns, like any process of collective decision making within human groups, is fraught with social influence. And once we have that, by definition we can assume that Condorcet's jury theorem is not going to be working. And if we want to trust democracy or any process of that sort, Condorcet's jury theorem cannot be the right argument for it.

So perhaps the best empirical evidence for how social influence can make you do things that probably are not good is the Solomon Asch experiments in '50s, where he had groups of confederates-- psychology students doing their Ph.Ds-- pretending to be subjects in a social psychology experiment. And one subject was particularly recruited without knowing that he's the only one. And they were given very easy perceptual task to do-- name the matching line to the test on the left with the standards on the right. And in the critical trial, everybody, all the confederates, would give the wrong answer. And the question was what would the participant do.

The interesting result, which has been replicated in all sorts of ways with variations of all sorts of quantitative aspects of this experiment, if you like, is that in that experiment, in Solomon Asch's original, 33% of the critical trials, the participants did actually give in to the crowd. 75% of the participants did it at least once. And only 25% actually just persisted in their argument, which probably was the good thing to do in this case, although you would call it an intuitive task, I would say. Right? Yes.

So in my attempts to combine these two questions, the question of independence of opinion and the role of social influence, I had devised a sort of classroom experiment, a demonstration that I would do during my talks or at the beginning of my talk, and then make the demonstration relate to the content of what I'm going to show you. But given that my computer broke down, or the VGA connector broke down, or whichever. We don't really know what at this point. I actually have had to skip the experiment, so I'm just going to show you the data from previous groups.

So what people did in my experiment in settings very similar to this-- I would do this at the beginning of my talks, so hopefully, although people knew what the topic is, they wouldn't exactly know what I'm going to tell them about. And hopefully you will agree with me that the results that I'm going to be emphasising is independent of the fact that my listeners were also my experimental subjects.

We had a number line task. And in this number line task, we present a line on the screen, and there is a beginning and an end. To each side we attribute a number. And there is a cloud of dots-- somewhere you probably can see it. Ah, here we are. There is a cloud of noisy dots here. And the question was what number does the centre of that blob correspond to? The correct answer in this case is 259.1 But what I asked people to do was I gave them pieces of paper. And on each piece of paper-- half of the people in the room would be given paper A, and the other half would be given paper B. This would be done totally randomly.

And people in the group A would first give their responses without anyone talking to anyone else. I would ask everyone not to cheat, and I would ask everyone not to talk to each other. And people in group A would first write down their answer without talking to each other. Then after that, I would have everybody in group A read aloud their numbers, individually, one at a time, until everybody had heard everyone in group A1's opinion. And then both people in group A and group B would also be asked to write down their numbers after having heard everybody else's opinion.

So in this case, you can see we had the independent responses of half of the subjects in the experiment recorded first, without anyone influencing anyone else. These people would be in the Condorcet/Galton camp, or let's say this question. The second time, they would write their answer, and the people in group B would be the ones who have been influenced. These would be the Mackay and Asch camp. Now the question for me was, we know the right answer. Can we look at the answers that they have given us and understand the relationship between independence and influence through what happens in these two numbers? And interestingly, this very simple, one-shot experiment does show you very interesting results.

So what I'm showing you here is the collective mean error. So if a given subject had given me a number, I would subtract that from 259.1 and divide it by 350. So positive numbers mean they have been overestimating that number. And as you can see, in general, people in this case had a tendency to overestimate the correct number on that line. And what is interesting is that as we go from A1, which is the independent opinion, to A2, and especially to B, we see that that bias to overestimate the number increases. In other words, the existence of the bias in group A1 seems to be exaggerated as we go from A1 to B, and everybody seems to actually be more convinced that this number should be a big number, rather than a small number. So that's one of the interesting aspects of social influence.

Another one is demonstrated when we look at the distribution of the results. Here I'm plotting the distribution of the errors for A1 and for B. And what you can see here immediately is that the distribution of responses across 200 subjects that I've been testing in different talks is much wider. And note that the x-axis of the two scales is the same. So the larger width of the distribution on the left actually shows you that there is a lot more diversity in individual opinions when those individual opinions have been elicited independently. The very loner individuals on the far left are contributing to the overall accuracy of A1, when averaged across everybody. So even that very misguided under-estimating subject is influencing the group in a very good way.

Whereas the people in group B, when we look at the distribution of their responses, what we see is that their responses are now much tighter. They're actually mostly one of two responses, either 275 or 300, which are those two sticks that you see much, much bigger than the other ones. And the interesting point about them is that this distribution shows us what social psychology has also found for quite a long while, which is herding behaviour. So what we have here is that people just copy each other after they have heard each other's opinion. Again, maybe in very similar to what we were just discussing.

So what we have is that social influence can exaggerate collective biases and can reduce the diversity of opinions. So this is probably the first thing that I want you to keep in mind. But we can also flip the question and ask it the other way around. If we assume total independence, if we generate situations for independent decision making, is that enough for groups to do better than individuals? In other words, are two independent heads always better than one?

And in order to answer that question-- one Ph.D. was on perception, so when I started doing these tasks, I thought, OK, we are going to do a perceptual experiment. Because what you know from perceptual psychology is that perception is noisy and incomplete. The best example of noisy, incomplete perception is probably refereeing in football. And having done my Ph.D. in England, I could think of these three. It is interesting that English football has always been fraught with very controversial decisions in refereeing and catastrophical consequences.

So the question you could ask, even on behalf of football, is if you had two, three, four, more than one referee, would you reduce these errors of perception, decision making? And most importantly, for the quantitative case that we were interested, by how much? The important point here is this quantitative aspect-- by how much performance is going to get better when you have two referees as opposed to one, or three as opposed to two, is going to be influential in deciding what model of collective decision making we are going to be accepting as the right one.

So the way we do this is that we put two people in the same room, we connect their computer screens to the same computer, and we show them stuff including noise and signal, and we ask them to do simple perceptual decisions for us. So if we go back to the first talk today, what I want to tell you is that in terms of intensity and extent, here we have a social interaction that is extremely tight on extent and very, very shallow on intensity. But the nice thing about it is that it still can tell us something interesting, I hope, about social interaction.

So this is what my subjects do in the experiment. There is a two-interval forced-choice task. They are looking at their computer screen, and they will see two separate presentations of a visual stimulus on the screen. Each one is about 100 milliseconds. There is one second in between them. In each one of these stimuli, there is a fixation point in the centre and six items around. One of those six items has higher contrast. The whites are whiter, the blacks are blacker.

The task of the subjects, which they first do individually, without talking to each other, by pressing a button, is to say if that higher contrast item was in the first or in the second interval. It's not about left or right, it's not about top or down, it's not even about the intensity of that oddball, just which one of the two intervals did it appear in. And then the computer tells the two subjects who are sitting there, by colour codes, which one of them has given which choice. And if there is a disagreement, they are prompted to talk to each other and make a joint decision. And then at the end that we give them the feedback for who was right and who was wrong.

The nice thing about this is that the history of perceptual and cognitive psychology is full of extremely sophisticated ways of interpreting these behaviour. And the way that we do this is that we plot their responses in terms of probability of reporting the second interval as the target interval. That's what I'm plotting on [AUDIO OUT] here. And the x-axis is the contrast difference at the target location, second interval minus first interval.

So when I have large numbers on the x-axis, large positive numbers, that means the target was in the second interval, and it was very easy to see. The difference between target and distractor was very easy. That's why on the very far right of the graphs, you see that people almost 100% of the times report the second interval as their target. Quite the opposite when the difference is negative and big. That means the target was in the first interval. That's why C2 minus C1 is now negative. And it was very big, and that's why almost 0% of the times my subjects are reporting seeing the target in the second interval.

Now the important point here is that by bending my data according to these different levels of contrast difference and then fitting a quantitative model, in this case a cumulative Gaussian to it, I can estimate two very useful behavioural parameters. The on that I'm going to focus on here is the slope of the rise of the curve that is fitted to the data. You can see that the slope of the rise of these curves is a very good estimate of the subjects perceptual sensitivity-- in other words, the subject's ability to distinguish signal from noise.

If the subject is perfect this slope at the centre would be vertical. We would have a step function that goes from zero, suddenly to one. If the subject is just doing it randomly, either by closing their eyes or by not really taking the experiment seriously or whatever, then we would have a flat line that goes at 50/50 all the way. So the slope will tell us how good the subject is at telling the signal from the noise.

Now I can ask-- I can rephrase the question, are groups better than individuals, this way by asking, is the slope of the red line, which is representing the groups here, higher, more vertical, than the slope of the blue and the yellow, which are the two subjects. And I can do that across many, many groups. This is just an example group. And I can then look at the question of what actually makes the difference between successful groups and unsuccessful

groups. If a group has been successful, that means the red line will be steeper than both yellow and blue. If the group has been unsuccessful, that means the red line will be flatter than both yellow and blue, especially flatter than the better of the two members.

Take it at this point, that I can quantify the individual subjects' success in doing this task by their slope. And I can call them maximum and minimum, referring to my better subjects and my worse subjects within a group of two. In this case then, I can define collective benefit as the ratio of how good the [INAUDIBLE] is doing versus how good the better member of the individuals is doing.

And also, I can define a metric of similarity. The similarity in this case would be the ratio of the worst person to a better person. The reason why I do it this way is that this way, similarity has a fixed range of going from zero, when the worst person very inferior to the better person, to one, when the two of them are identical.

What we find in these experiments, the critical finding, is that when we look at these ratios, what we understand is that the similarity is the main predictor of collective benefit. In other words, as Chris already told you in the morning, what we see here is that as we go from very low similarity to very high similarity, there is a very predictable and very replicable, very robust relationship between whether the competence or the sensitivity of the two people working together is similar with each other and how much they are going to benefit from working together.

So in a way, we can say that yes, we can have two heads are better than one. And we can have too many cooks spoil the broth. But the critical parameter that tells the difference, especially in this case, where all decisions are made initially independently, for two independent heads to be better than one, they have to be similarly competent. So I guess we can summarise this part as two equally competent, independent heads are better than one.

And then finally, I can get back to the question of what makes cooperation good or bad. And I can tell you that my data shows you that there are two things that are important-- diversity and equality. If we have those two things, then we can do better than one. Individual opinions that are independent from each other are much better when we put them together. And individual opinions that come from equally competent partners are much more valuable.

But I'll just finish with the final question that I want you to also carry with you, not in terms of a conclusion from what I've shown you, but as a reminder that all the things that I've told you is just a very, very limited aspect of collective decision making. What I think is also very important, and hopefully, our lab's research is heading towards that direction, is that the literature, the background of collective decision making and social psychology has been obsessed with, are two heads better than one. How can we make cooperation better and better? What are the constraints that can increase collective intelligence?

What I think is also important is that collective action has consequences and benefits for humans that are not just limited to increasing accuracy. I've just listed a couple of them here. Collective decisions diffuse responsibility. Take the very simple case of murder. If you have one murderer, or if you have one person accused of murder, that person is accused of murder. In English legalese, I think if you have more than one, the term changes to conspiracy to

murder, which suddenly changes everything. The moment you have more than one person accused of murder, everything changes. In a lot of cases, even in much older, for example in Islamic Sharia law, it's almost impossible to prosecute more than three people on the charges of murder.

Similarly, collective decisions reduce regret. This is a work that has been done by Antoinette Nicolle and Chris in Wellcome Trust Centre a few years ago, where they showed that if your decision is part of a collective, who have unanimously chosen to go for a certain option, and later on it turns out that was the wrong decision, people express much less regret, they show much less brain response to bad outcomes. And they seem to take the bad outcome much more gracefully when they realise that everybody else thought the same, which also brings up the issue of collective decisions help you justify your choices. And also they help you later on. If it turns out that you were wrong, your obvious argument would be everybody else thought the same.

So I will stop here, only with the idea that I've given you a couple of parameters that I think are important for increasing collective intelligence, perhaps. But I think collective intelligence is not just about getting better accuracy. Thank you so much.

[APPLAUSE]

Go back to the question you made, Geoff, in your talk. One of the cleavages there, to me, is really the question of individual collective and cognitive or emotional [INAUDIBLE]. Because I mean, something that was very interesting in the last talk is [INAUDIBLE] that collective intelligence is not just about intelligence, so to speak. It's not just about cognition. It's not just about efficacy of cognition, but is about investment. It's about the investment people make as part of a group. And investment is particularly important for them because they stand to lose something if they are collectively stupid, rather than collectively intelligent. So that element is crucial to develop our understanding of collective intelligence.

And linked to that, what do you know of the motivations for people to go along with the group, knowingly? Is it fear of humiliation? Because there's actually no cost associated with giving a wrong answer. Or do people genuinely believe?

Well I think statistically, especially if you're in an uncertain situation, it makes very good statistical sense to go with the group. So that's one thing. Herding, in a lot of situations, is the best thing you can do. And it's nothing to be ashamed of, so to speak.

In the Asch experiments in particular, he had a condition in which he looked at that specifically. So the condition that you showed is one in which the answer is public. So you have all these confederates who said all the same thing, and then you're the last one, and you have to say what you think is the right case. So there's a lot of group pressure, because you know they'll see what you've said. And then you had another condition in which people were writing down their answers on their own. Everybody else was giving the answer publicly, but then you had some kind of pretence that people had arrived late, and so they couldn't fully participate and they had to write it down. And then no one complied.

[INTERPOSING VOICES]

So it is purely normative. It's like 10% percent. It's very, very low, the rate of trials in which people complied to it. In the case of the Asch experiments, we know it's truly normative. There is no informational conformity.

I found it interesting that in one of the [INAUDIBLE] we put gender. We would also put a gender of the participants. So I was wondering if you had done some studies related to social influence based on gender or other elements. I know that for juries, for example, there have been studies about when professionals were giving their opinion, that had much more power than people that were coming from different backgrounds, for example.

When I first read your paper and described it to someone, they said, he's just writing about men.

[LAUGHTER]

Women are never like that. [INAUDIBLE].

[AUDIO OUT]

Yeah, I mean there will be gender differences, in particular, to some extent sometimes in how explicit the arguments are going to be. And that's also true of many-- so in many cultures, people frown on direct confrontation. So they're not going to tell you you're wrong because this and this. They're going to tell you something that it's pretty clear they want you to understand as a reason for why you're wrong, but they're not going to spell it out, because that would be rude.

And I think you have gender differences, cultural differences, but in the end, we're still talking about argumentation. So in a way, it's less easy to see, especially if you're not in that culture, because you don't understand what people are saying. But for people in the culture, it's pretty clear that they're arguing in the sense of exchanging reasons for their points of view. So we haven't done experiments with comparing genders. But we have done experiments in Japan, for instance, which is a bit like this, in which people try to avoid being overly confrontational, very much unlike France. And so we found over all the same result.

In one experiment, we had this interesting thing in which, when we asked the group-- so people gave answers on their own, and then they gave answers in pairs, and then they gave answers on their own again. And so on the pairs, there was not much improvement, because people were just averaging their opinions, because that's the best way to save face and to not offend anyone. But then, there was an improvement when they were giving their opinions on their own. So then during the discussion, each of them knew who was right. But they didn't say it, to be polite. But they knew who was right. So in a way, you still get the [INAUDIBLE] benefits, because people know what is the right answer without having the social implications of saying outright, you are wrong.

Well, in my experiments, this is a kind of ongoing issue, especially with quantitative modelling of interactive decision making. Almost all of my work, tested in several different countries, can only predict and explain male behaviour. But female behaviour-- my models failed to capture what they were doing in Iran, in Denmark, in China. And it is very funny, so

I'm not offended if people laugh, because the regression line that you saw in the data, the red regression line, is actually derived from the quantitative model. And it holds in all of these countries that I told you if two people who are communicating are men.

In Denmark, female pairs actually fall mostly below that regression line. That means they underperform compared to what the optimum model expects them, given their own individual ability to do. In Iran, they actually overperform, so my model fails in different ways. So it's not just that my model fails in a simple way.

Excess collaboration that is, basically.

In the case of Iran, yes. And what I've seen in my data in Iran, which is very surprising, is that in Iranian women, their relationship to similarity seems to break down. What we have been thinking about as a consequence of sitting through a lot of these and watching the footage and looking at what people talk to each other about, is that I think perhaps the gender difference between how people resolve these uncertain collective decision makings should come down to what it is that we call the utility of what they are doing together.

I think it's fairly straightforward to assume that male dyads actually have a clear definition for us, not for them. I mean, we know clearly that their utility is to maximise accuracy. Whereas for females' dyads, I don't know what it is that they are trying to maximise. One thing that I can guess is that the process of collaboration and the process of listening to you once, and then listening to the other person, social inclusion, can be helpful. I've seen this happening in different countries when women promise future decisions to each other. They say we will go for this one, my decision, but I promise next time we disagree, I will go with yours.

In the context of a purely perceptual contrast discrimination, this is not something you want to do for doing better in this experiment. But they forego the accuracy in order to maybe transfer social capital to each other, which I think is basically just as sensible a strategy to have as maximising accuracy. They are volunteering for the experiment. It's not like their life depends on it. So if they have fun with it and do it the way they want, and that turns out to be consistent across countries, then that's for me to figure out why.

We need to move on in a moment, just maybe one very brief question to each of you. What, in the field you've set out, do you most want to know that you don't know?

[INAUDIBLE]

[LAUGHTER]

For me, the relationship between responsibility and accuracy. That's what I'm studying next.

Is that in terms of what you were just saying, how much at stake there is in--

My hypothesis is that people are ready to give up accuracy in order to reduce their responsibility. But I don't know what.

I guess one of the things I'd like to know is why people so tend to underestimate argumentation. It seems that people really don't have the intuition that when you put two people together, the one who's right is going to win. And it is overwhelmingly the case. So it's funny. There seems to be a discrepancy between what happens and the perception people have of what happens, which is a bit puzzling to me.

Someone should make a remake of 12 Angry Men with Henry Fonda, but where Henry Fonda takes the wrong position and persuades everyone.

[INTERPOSING VOICES]

OK. [INAUDIBLE]. Thank you.