

Models and behaviour change

A dialogue

This dialogue was constructed by Simon Christmas (King's College London) at the invitation of the UCL Centre for Behaviour Change, following an event with the title 'Models of behaviour change: how useful are they?' hosted by the Centre on 2nd June 2014.

Its content draws on the event itself, and on interviews with Jamie Brown (UCL), Nicola Christie (UCL), Anthony Finkelstein (UCL), Heather Gainforth (UCL), Graham Hart (UCL), Kate Jeffery* (UCL), Mike Kelly (NICE), Susan Michie (UCL), John Owens (King's College London), Alan Penn (UCL), Jeremy Watson (UCL) and Robert West* (UCL). Interviewees who were also speakers at the CBC event are marked with an asterisk.*

The dialogue brings together three imaginary graduate students – Evie, Paola and Yusuf – who attended the Centre for Behaviour Change event and have now met up to discuss further some of the issues raised.



- Yusuf It's great that we've been able to find this time to talk further about models: what they are, why we need them, and how useful they are in the context of behaviour change in particular. I know we all found the event hosted by the Centre for Behaviour Change thought-provoking. For my part, I came away with a real appetite to explore the issues raised more fully.
- Paola Me too. And even if we don't come to firm conclusions, we can at least try to get some sharper questions. Because I suspect there will be points we don't agree on.
- Evie I think you were looking at me then Paola! I know one thing all three of us agree on, though: that some of the most interesting stuff happens when we talk across disciplines like this.
- We're all so used to communicating *within* our own disciplines, with people who share the same worldview and assumptions. When we talk *across* disciplines, we have to re-examine the things we normally take for granted, make explicit our assumptions, and work extra hard to be really clear about what we're saying. And in the process, we learn so much.
- Paola Exactly. Provided we approach the conversation with an open mind, ready to listen and to be challenged – and also to offer our own challenges. That's what I love about talking to you two!
- Yusuf We'll certainly need open minds for this topic, I think. Listening to the three speakers at the event, I realised either how often the word 'model' gets used in my own field, or how rarely I ever stop to think what a model really is. Before the session, I'd never really thought about how people in other disciplines might use the term.
- Evie Or the fact that they might mean something different.
- Yusuf Exactly. But before we get onto differences, I wondered if there were some basic points we could agree on. I had two suggestions.
- First, it seems to me that all models are representations of something in the world.
- Evie Not all representations are models, though.

Yusuf No, I'm not saying that. But all models are representations. Perhaps we should go further and say that there's an isomorphism between a model and the world. For instance, that's true even of the kind of models that architects build, so that people can see what a building will look like.

Secondly, models always involve a degree of simplification or abstraction. They help us get a grip on a complex reality precisely because they leave out some aspects of that reality.

Evie That's linked to the idea of parsimony. You only put into the model things that really need to be in the model.

Paola I'm reminded of a map. A map that included all of the same features as the real world would be completely useless. In fact it wouldn't be a map: it would just be a duplicate reality.

Yusuf As the saying goes: the map is not the terrain. It's important never to forget that our models are just that – models, and not reality.

Evie That's very interesting, because it strikes me that there are at least two types of model, with different relationships to reality. In some cases, we're using models to understand things that already exist or happen in the world. I think the map is a good metaphor for this kind of model. In other cases, however, we're building a model of something we want to make exist or happen – in which case our model is more like a blueprint. I think these are quite different tasks.

Paola That's a really good point. It raises the question of purpose – what are our models for – and how that affects the way that we put them together. I wonder if the distinction you make, Evie, makes a difference to *how* you construct a model.

Yusuf I think there's another distinction we could make here. It seems to me that there is a *continuum* of types of model. At one end, we have models which are precisely specified, typically in the languages of mathematics and logic. At the other end of the continuum, we have less precisely specified models, which seek to express insights about how things work without doing so in formal terms.

Evie I have to say, my first reaction is that what you're describing there is not a continuum of types of model, but a continuum from good to bad, or from model to non-model. From my perspective, a model that isn't 'precisely specified', as you put it, isn't a model at all. You have to be able to describe a model in formal terms for it to qualify.

Yusuf Well, don't you think it depends on your purpose, like Paola said? After all, some models are not intended to be analytical. They are thinking tools, not analytical tools.

I accept there needs to be *structure* in a model. That's at the heart of the idea of isomorphism: the structure of the model reflects the structure of the real world. But I'm not convinced that this structure has to be describable in formal terms.

The Wider Determinants of Health model, for example, has a very loosely specified structure – but people find it extremely helpful. MINDSPACE is another example that's proved very useful in some contexts. Models like these can be extremely helpful in thinking through the kinds of factor that could be contributing to a given issue. It's a different purpose, so a different kind of model.

- Evie I don't dispute the value of things like that. All I'm saying is that for me they don't qualify as models. They are checklists, or frameworks, or communication tools, maybe. They may well capture quite sophisticated insights. But I don't think we should allow the fact that they're represented graphically, with boxes and arrows, to kid us into thinking that they are models – at least not in the sense I'd like to reserve for the term. What do the boxes really mean? What do the arrows mean?
- I think we risk losing an essential technical term if we stretch the use of 'model' in this way. Not to mention the risk that people start believing that frameworks and checklists like these have the same standing as formally specified models. Which they clearly don't.
- I mean, we could go in the other direction and say that the term 'model' can be used of *any* representation. For example, is a Picasso painting of a face a model? The question is not about the importance and value of Picasso paintings. It's about whether we lose something by stretching the term 'model' until it means pretty much the same as 'representation'.
- Paola It's interesting because, while you were talking, I was thinking of a third kind of model – one that is there to provoke new thinking. For example, you might use gravity as a model for flow of populations around and between cities. Or you might use an agent-based model to demonstrate how complex behaviour can arise from simple rules. The purpose of models like these is not to explain but to provoke. They challenge us to think of alternative ways of explaining the phenomena we observe in the world around us.
- Evie For me, those sound closer to metaphors than models. I mean, if I've understood you, you're not meant to treat them as literally true.
- Paola Definitely not. It can cause problems if people start believing that a model of the kind I'm describing is literally true.
- Yusuf Hmm. I thought I'd cracked it with my continuum! I suppose this is even more complicated than I thought.
- Evie It seems to me that what we need is some kind of framework here – I guess maybe you'd call it a meta-model, Yusuf! – of the different kinds of purpose models can serve, and the implications that has for the form they take.
- Yusuf I agree. A meta-model sounds great! Although actually, I think of it in terms of developing a common language we can use across different disciplines to describe and talk about these different kinds of model.
- Paola And some kind of consensus on what gets called a model at all!

2

- Evie For the sake of today's conversation, I wonder if we can narrow the scope a bit – especially as what we're really interested in is the role of models in behaviour change.
- For example, Yusuf mentioned that architects build models of buildings, or maquettes, so people can get a sense of what they're going to look like.
- In a behaviour change context, however, I think what we're interested in is a particular class of models: models that seek to represent the dynamic behaviour of

- a system or set of systems. For example, that wouldn't include the architect's maquette; but it would include a computational model of the thermal behaviour of the building, which modelled how it absorbed heat, cooled down, and so forth.
- Paola I'm not entirely sure on that one, Evie. Buildings may be static in one sense of the term: but they differ when perceived and experienced from different points of view. So I could argue that the maquette is also a dynamic representation, because it allows the viewer to interact dynamically with it and adopt these different points of view.
- Yusuf I'm not sure about the distinction either, but for a different reason. The fact is that a lot of our models of health-related behaviour are actually regression models. These can be very precisely specified and very useful – but they're doing something different from a dynamic system model.
- Paola Sorry, Yusuf, can you explain that one to me? I've always assumed that the numbers on models said something about how much different elements of a system influenced each other?
- Yusuf That's a common mistake. In fact a regression model is a statement of co-variation, not influence. It's entirely constrained by the natural variation in the population. Something may be an important influence, but if it doesn't actually vary much in the population, then by definition its variation won't explain much of the variation in other elements of the system.
- For example, in the past, GPs all got paid pretty much the same. If you'd done a regression analysis, you'd have found that variation in levels of pay explained little if any variation in GP behaviour – for the simple reason that there was no real variation in levels of pay. But that doesn't mean pay is not an important *influence*. In fact, if you pay GPs lots of money to do something, they will do it.
- It's also important to remember that these are models of variation in populations, not models of what is actually going on in individuals. For instance, suppose you found that, in a given population, intentions to go to the gym explained 20% of variation in whether or not people actually go to the gym. That could be because intentions are having some influence on the behaviour of everyone in that population; or a great deal of influence on the behaviour of just some of the people; or anything in between!
- Paola Thank you. That's helpful!
- Evie I take your points. Maybe that wasn't the most helpful way to define the scope. But let me have another go: surely we're looking for models that enable us to make predictions regarding outcomes – at least probabilistically. Predictions that we can go out and test.
- It seems to me that that is central to the whole endeavour of science. The way I see it, a model is a way of formulating in precise terms what you *think* is going on in a system, that allows you to say: if I'm right about this, I ought to see the following things happen as well. Then you can go out and see if they really do happen.
- That's one of the reasons why I think it's so important that models are precisely specified. It's one thing to say that you think two things are connected in some way: another thing entirely to express the nature of that connection in a way that allows you to go out and test whether or not you are actually right.

- Yusuf That feels like a very important point when we consider behaviour change in particular. What we want to be able to do is predict what will happen as a result of some intervention – or at least assign probabilities to different outcomes.
- Evie Exactly. For example, that's at the heart of how models get used in cost-effectiveness studies. In an ideal world, of course, you'd have trial data: hard empirical data of what happens when the intervention in question is made. But if you don't have that, you can look at models of how key variables interact with each other to at least estimate how much bang you will get for your buck.
- In some areas, we have robust models to do this kind of work. For example, we can build models that link data about sexual behaviour in at-risk populations to data about the incidence of HIV, and use these to predict the likely impact on incidence of a given intervention – the earlier use of antiretrovirals, for example.
- Paola I was struck by the formula for cigarette consumption price elasticity. The change in consumption is 0.4 times the change in the price. I suppose that's a simple example of what you're talking about.
- Evie Exactly. The model allows you to make predictions, which you either go out and test or use to make interventions.
- Yusuf And in fact, those things are quite closely related. Every time you make an intervention, it's another test of the model. Even when we've tested a model many times, and have a lot of confidence in it, there's always the possibility that something new will come up.
- Paola The wobbly bridge!
- Yusuf Pardon?
- Paola When they built the Millennium Bridge, they were using well-established engineering models that allowed for all sorts of factors – including the fact that people tend to synchronise their steps as a suspension bridge bounces up and down. What no-one realised – because no-one had built a bridge quite like that before – was that people would also synchronise with the side-to-side movement of the bridge. The intervention – the bridge – showed up something that was missing in the model.
- Yusuf That's a great example.
- Evie I like this idea a lot: that interventions are also tests of the model that informs them. And it works the other way round as well: you test a model by making an intervention and seeing what happens. The difference is not in what you do but in how confident you are about the model.
- In a control system context, for instance, you have a high degree of confidence in the model, because the model is so well tested. So when the real world doesn't match the predictions of your model, you look for problems in the real world rather than in the model.
- Paola Could you give us an example?
- Evie Well, suppose you're running a brewery. You have a control system which takes the same inputs as the brewery itself, and models its internal states and outputs. Now suppose the brewery and the model start delivering very different outputs – for example, you start seeing very different temperatures or pressures in parts of the brewery. Because you've a high degree of confidence in the model, you look

- for problems in the real world: a faulty component, for instance, or maybe a bad batch of yeast.
- Paola So I'm trying to translate this into the context of behaviour change. I have a model of how I think people behave, and based on that model I make an intervention. For example, I develop a training intervention to improve hazard perception in novice drivers. I pilot my training, and it doesn't work in the way the model predicted. What does that mean?
- It could be that my model was wrong in some way – for example, it was missing some key variables.
- Alternatively, it could be that the model was correct but I did something wrong in the design of the training intervention – for example, I had people do it on a computer when there needed to be a group discussion as well.
- So how do I decide which it is in practice?
- Evie That's a really good question. I'm not sure I have an easy answer.
- It seems to me there's a lack of shared understanding and methodology around *how* models are tested – at least in the field of behaviour change. What kind of evidence is needed to say that the model needs to be refined, or extended, or developed? How do we gather and interpret that evidence in a systematic way?
- For me, this links back to issues around specification. In behaviour change, for example, the basic challenge is to build an evidence base of how techniques affect behaviour. The techniques are the independent variable, if you like, and the behaviour is the dependent variable.
- But as you say, Paola, you also have to take account of other variables, such as *how* the intervention is delivered. And then there's the setting, the cultural context, the population – these things can make a huge difference.
- Finally, which model you've used in designing the intervention is itself a variable. Because the question that sometimes comes up is: where is the evidence that interventions based on Model X or Model Y work better? And the answer is: there isn't any! In fact, we don't even have the evidence that theory-based interventions as a whole work better than interventions based on, say, intuition and gut feel. And that's because we don't have any systematic way of gathering that evidence or of ensuring that the model has been applied well in the first place.
- What we need is some kind of agreed taxonomy, a way of coding these different variables each time we make an intervention – not just the technique used and the behavioural outcome, but also the way in which the intervention is delivered; the setting, cultural context and population; and the model used. That way we can start to systematically track their relationships.
- Yusuf The kind of taxonomy that's already been developed for, say, types of intervention.
- Evie Yes.
- Paola That's going to be quite challenging in the case of something like culture. I'm thinking about the parallel with taxonomy in biology. Any taxonomy you come up with will be incomplete by definition, because new species will evolve in the future. In biology that's not such a big problem, because evolution happens so slowly. But culture changes much more rapidly, constantly creating new forms in the way that evolution creates new species.

Evie Well, I didn't say it was going to be easy!

3

Yusuf I'd like to bring in another angle on this. We've been talking about the need to test our models whenever we make interventions. Theory, we might say, needs to be continuously tested by practice, otherwise it won't be right! But I think there's another side to this. Practice should be informed by theory, otherwise it won't build on the accumulated learning we have built up over past interventions. We ought to be making behaviour change interventions with clear theoretical foundations.

Paola And do you think that happens?

Yusuf You know how I'm going to answer that question! No, I don't. Not often enough, anyway.

It's as if we've set up these two worlds: the world of theory and the world of practice. When in fact, that distinction is an artificial one: theory and practice should be in a dialectical relationship, two aspects of a single endeavour.

Paola I have one thought as to why that split between theory and practice might exist.

I was thinking earlier when you were talking about the idea that interventions are testing out the models they are based on... It seems to me this idea might make some of the people who fund behaviour change interventions very uncomfortable. Because another way of putting the same point is that there is always an element of risk that an intervention won't work. I'm not sure risk is something that people are always comfortable with.

I think there's this fantasy that, by doing research, you can remove the risk. You find out what will work, then 'roll out the solution'. People want research to tell them what to do, when all it can really tell them is what to *try*.

Evie And what *not* to do. Often it can be quite clear about that!

Paola Agreed! Either way, there isn't that dialectic relationship, that two-way conversation.

Yusuf You could well be right. But at least in the scenario you're describing there's a one-way conversation.

What worries me even more is situations where people simply ignore the theory altogether. I have encountered people who reject the idea that models have any role to play in the design of behaviour change interventions. 'I've got years of experience working in this field: what can your models tell me?'

Paola The reality is, of course, that even someone who claims to be anti-theoretical has some kind of model in mind when they design an intervention. It may not be articulated, even to themselves. But unless you make assumptions about how things work and why people do what they do, you have no basis for making any intervention at all.

For example, one you often see is the unstated assumption that people respond to financial incentives in line with economic utility theory. If you challenge someone on this, they will tell you it's 'common sense': but call it what you like, it's still a theoretical commitment, a model. And in this case, one that happens to be wrong!

- Yusuf I like this way of putting it. Rather than say that some interventions have a theoretical basis and others don't, it would be more accurate to say that *all* interventions have a theoretical basis: but in some cases, this basis is clearly articulated in the form of a model, and in others it is implicit.
- What we should be doing is bringing these implicit models into the open so that we can subject them to proper scrutiny. Does the evidence support them? Will the interventions we're planning test them? Are they even internally consistent?
- Paola Doing that can be very hard, because these implicit models are things we think *with*, not things we think *of*. They're like paradigms that sit in our head, structuring our interpretation of everything else.
- And as long as they remain that way, they can be remarkably impervious to evidence. For example, there's a common view among architects that enclosure creates community. When they apply the principle and it fails – enclosure fails to create community – the response is sometimes: 'Right, well clearly we didn't enclose enough. Next time we need to enclose even more!' The possibility that enclosure *doesn't* create community is not even considered.
- Yusuf It would be like the engineers who built the wobbly bridge refusing to consider the possibility that the fault lay in the model they had used.
- Evie This is why the discipline of creating an explicit logic model is so fundamentally important when developing an intervention.
- Yusuf Is that what I'd call a theory of change?
- Evie I think so. There are different names, but the principle is the same: that you specify clearly how the intervention is expected to work, based on what you already know about how things link together in the world.
- Building the logic model is a way of forcing yourself to make explicit any assumptions you have about how the intervention is going to work. It's a way of making yourself go back to the existing evidence: what grounds have you got for expecting things to happen in this way? And it's a great way of identifying gaps in our knowledge.
- Yusuf Because, let's face it, when it comes to behaviour change there are usually plenty of those!
- Evie Absolutely. So what we need to do is be clear about those gaps, clear about our own hypotheses, and clear about how this intervention is going to allow us to test them out. Which is another key function of the logic model: it provides the framework for evaluation.
- Paola I really like this idea that one of the functions of a logic model is clarity about the things you *don't* know.
- However, I can also see why people might resist doing that. It goes back to the point about risk I made earlier. If you're about to spend a pot of money on trying to change behaviour, and if your personal performance is measured on the success of the intervention – not to mention your professional pride – then it can be a bit alarming to be told that the planned intervention is in fact testing the hypotheses on which it is based!
- For me, it's all about understanding what that phrase 'evidence-based' really means. When people talk about an evidence-based intervention, I sometimes get the

- feeling that what they really hanker after is an evidence-*determined* intervention. They want the evidence to tell them what to do, and remove all the risk.
- Evie I think what we should really say is that interventions should be evidence-based *and* evidence-generating. Yes, in so far as we have evidence, the intervention should be based on that; and in so far as we *don't* have evidence, the intervention should be helping to fill the gaps.
- Even if you're consciously setting out to innovate, this formula still applies. The whole point of innovating is to try something new – something for which, by definition, there isn't an evidence base. But that just increases the need for your intervention to be evidence-generating. Hence you still need an explicit logic model – a clear articulation of the way you *think* your innovation is going to work.
- Yusuf That's very good. And obviously, when you talk about interventions, you're seeing the evaluation as *part* of that intervention. Whereas in practice it's often seen as something over and above.
- Evie You're right, I do see the evaluation as an integral part of the intervention. And again, applying the discipline of articulating a logic model helps to ensure that it is.

4

- Paola It strikes me that there's another distinction we could make here, one that we missed in our initial discussion of different types of model. It's a distinction between generic and specific models.
- When you develop a logic model, that sets out the way in which a specific intervention is expected to bring about a specific change in behaviour in a specific context.
- On the other hand, there are generic models, like COM-B or PRIME or the Theory of Planned Behaviour, which seek to articulate the relationships between a range of factors and behaviour in general across a wide range of contexts.
- Which raises a question. Evie, you've set out a very clear case for why we need the specific logic models when we make interventions. But what are the generic models for? If I'm designing an intervention, why should I pay any attention to them?
- Yusuf That's an interesting question. I think the answer lies in what's *behind* those generic models. I mean, they didn't just appear from nowhere. Those generic models capture an understanding of human behaviour that has been developed over a century or more of scientific investigation. They're far from perfect: but starting with them makes a lot more sense than just saying: 'This feels right to me.'
- No-one would do that in, say, physics. But for some reason, when it comes to human behaviour, we're not always very good at building on what we've already learned.
- Paola There's the makings of an interesting response there to the kind of person you mentioned earlier who says: 'I've got years of experience: what can your models tell me?' The response would go: 'Yes, you have years of experience; but these models capture the accumulated years of experience of many different people.'
- Yusuf I may try that next time. I'll tell you if it works! The principle is right though: generic models provide a way of accumulating our experience and learning.

- Evie I think we could go a bit further than that. Generic models aren't just an output of the process of accumulating learning. They're also what makes that process possible. If you're going to accumulate learning across many different studies, you need to be able to review them systematically. If different studies use different models, then you have a problem: how do you compare them?
- Yusuf This is an important issue. Because in practice, even when people *do* base interventions on existing models – generic models, as you've called them, Paola – they often fiddle around with them, and alter bits to fit them to their own purposes. And that makes systematic review hard if not impossible.
- Paola It strikes me there are two potentially competing agendas here. For the person who wants to bring about behaviour change, fiddling around with models to fit one's own purposes sounds like it might be a good idea. For the person who is trying to advance our understanding of human behaviour through systematic review, it's a huge problem. I wonder if this tension accounts for some of the difficulties – and if so, how it might be resolved.
- Evie You may have a point. But I also think that the problem comes back to the poor specification of many of the generic models of behaviour that we have. Not all of them, but many. If a model is not well-specified, then you *have* to fiddle with it before you can use it for behaviour change purposes – because you need to produce a well-specified logic model for your own intervention.
- Yusuf I'm not sure whether the challenge here is with the models themselves or how they are *reported*.
- Evie Sometimes there's no way of knowing. But you're right, from a behaviour change perspective we need much clearer standards for reporting models. If you can't report your model in a way that allows other people to use it *without* having to fiddle with it, then it's of no use.
- Yusuf As a teacher of mine used to say: 'If you can't say it, you haven't thought it.'
- Evie Something like that. It's an area where I think one could learn a lot from looking across disciplines as well. How do different disciplines report their models? What can we learn from each other?

5

- Yusuf Another issue I have is the sheer proliferation of behaviour models. There are so many of them. As you said, Evie, part of the job of a model is to create a common language that allows us to compare and systematically review many different studies. Instead, we're all talking different languages. It's like the Tower of Babel!
- Paola Though we did say at the beginning of this conversation that you might need different models for different purposes....
- Yusuf That's true, Paola. But the reality is that many models are just exercises in relabeling psychological principles that have been around for many years. It's hardly surprising if people who are designing interventions throw up their hands in despair and pick the first one that appeals to them – for instance, because it reflects their untested assumptions about the topic.
- Evie Or else they take a pick-and-mix approach. It's like those sourcebooks that designers use: you flick through them looking for something that sparks your

- thinking. In so far as designing a behaviour change intervention is a creative act, I can see how that could be really important. But it doesn't help with building a science of behaviour change.
- Yusuf And another thing: when people develop new models they're not clear enough about the ways in which they are building on previous models. Typically they will make reference to other models, but they don't provide real detail about how they have built on or developed those models. So it can be really hard to map the linkages, the overlaps, the differences.
- For example, take a construct like 'self-efficacy'. Lots of models use this construct, and typically the authors will cite Bandura. But when you look more closely, different models are using the construct in different ways, without making that explicit. Some focus more on control, for instance, others more on self-confidence.
- The result is a set of models which aren't commensurable. They don't make our common language richer, so that we can talk about and understand more; they just add new dialects, so we understand each other less well.
- Paola Poor Yusuf! You're quite exercised about this one.
- Evie And you're starting to sound a bit like me. Better specification of models!
- Yusuf It's true, it frustrates me. It's crazy how often people come up with new models. And yes, Evie, it's partly about better specification. But I think we also have to look at the culture we operate in. Why do people generate models? How are they rewarded for doing so?
- In my view, people should not be allowed to generate new models unless they can show they've done a thorough review of existing models in the area they're dealing with, and established the shortcomings of those existing models.
- There should also then be an agreed framework for the reporting of the new model, including real clarity about its heritage: the other models it has drawn on, the ways in which constructs from those other models have been developed, and why.
- But making this stick would call for changes in career incentives, institutions, professional mindsets. The journals would definitely need to get behind it. It would call for behaviour change, in fact!

6

- Paola Meanwhile, right now, we need to find a way of helping people make better use of the models we do have. I was struck, Yusuf, by the sympathy you expressed for people who design interventions and are faced by this plethora of models.
- So I wonder if we could provide clearer guidance to people about how to select and use models. After all, even if our models met these high standards of reporting, there would probably still be more than one of them. We would probably still need different models for different purposes.
- Yusuf I'm certain we would. In fact, I think the idea of striving for a single, all-embracing model is probably incoherent.
- Paola Why so?

- Yusuf Because the essence of models is to simplify. You have to leave things out. And what you leave out depends on your purpose.
- Paola A bit like the difference between a road atlas and an ordnance survey map. They leave out different things because they're designed for different purposes. A road atlas would not be much use for walking in the hills; but if you're driving across the country, the ordnance survey maps would be very unwieldy.
- Yusuf Exactly. But an all-embracing model would be one that could be used for *any* purpose. Which means it wouldn't be able to leave anything out. It wouldn't be able to simplify.
- Paola It would be the terrain, not the map! That's a really interesting argument. So however much we tidy up our existing models of behaviour, we're still likely to have more than one. Which brings me back to the question of guidance. Can we provide guidance to people on how to choose the right model when they are designing a specific behaviour change intervention?
- Evie It's a great idea. I think a key aspect would be to help people understand the *limits* of different models. For example, the Theory of Planned Behaviour gets used to investigate behaviours that clearly *aren't* planned. In that case, the limits of the model ought to be obvious: it's in the title! But I guess that's an indication of how clearly stated the limits need to be, if even the name of the model isn't enough.
- Yusuf I like the idea of guidance as well. Building on Evie's point, it would be good to be very clear about what kinds of things a model is *leaving out* in order to simplify. When you're selecting a model, it's important to make sure you're not leaving things out that are really important for your particular purposes.
- Paola So we don't want people yomping across the hills with a road atlas.
- Yusuf That's one way of putting it...
- Paola Building on that, I think we should also remember that, for some people, the kind of intervention they can use is determined in advance. They either have a car or they have walking boots. They don't have a choice between them.
- I think we sometimes come at the topic of behaviour change from the perspective of someone for whom any intervention is possible in principle. I suppose that could be the case for senior policy people developing a Government strategy – say on smoking reduction. But it's quite an unusual perspective. Most people approach behaviour change with an intervention already in mind. If your job is in a pharmaceutical company, for example, then you already know that what you need is a drug to help people quit. If you're part of a communications team, you know you need a different kind of approach.
- So in practice, most people selecting a model need one with elements they can actually hope to influence. They're not in a position to step back and ask whether they should be trying an entirely different kind of intervention.
- Yusuf That's a really interesting point, Paola. It's a reminder that, if we're going to develop guidance on how to select and use models, we need to do it from the perspective of the people are actually out there selecting and using.
- In my head, I keep coming back to this point about the need for a common language which we can use to describe models and talk about their strengths and limitations. When I first made that point, I had in mind a language we could use

across different academic disciplines. But actually, we should probably drop the word ‘academic’: we need a language that can be shared by all the different communities of thought and practice who use, test and develop models.

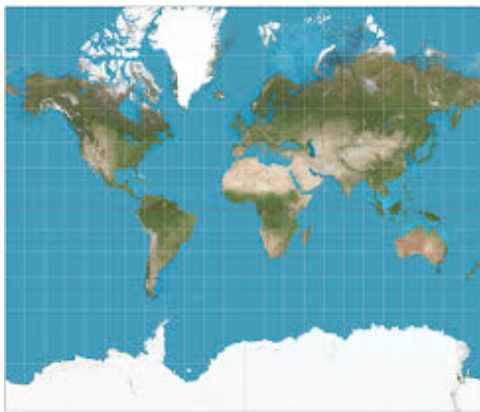
7

Evie What we’ve outlined here is the basis for a very rational and considered approach to the selection and use of models. That feels like it’s quite a long way from where we are at the moment, don’t you think?

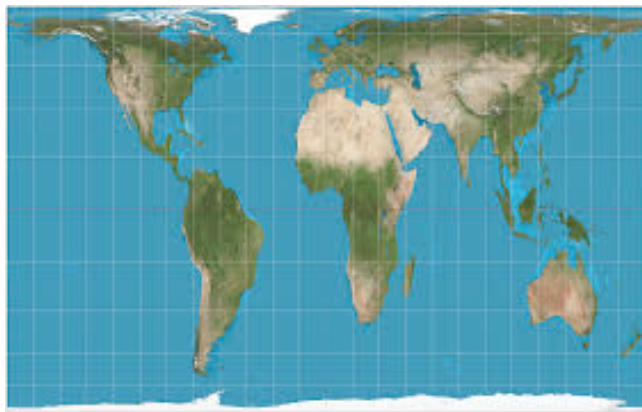
Yusuf It does. As things stand, I think people tend to stick to the models they’re familiar with. As I’ve said, I can sympathise with that. But it leads to problems: the more comfortable we get with a model, the more likely we are to forget that that’s all it is – a model, a simplification of a complex reality. We start mistaking the model for the real world.

Evie Paola’s map metaphor is quite useful again here. Maps of the earth all involve a degree of distortion of reality, because they’re projecting the surface of a sphere into two dimensions. We’re all so familiar with the Mercator projection that we tend to forget it’s not the reality. For example, we start believing that the areas of countries are as shown on the map. It’s pretty shocking the first time you see the Peterson projection, and realise just how misleading a map can be if you use it in a way it was not intended to be used.

Mercator projection



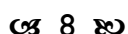
Peterson projection



Paola That’s actually an example of another potential problem, Evie. Different projections can become laden with political significance – for example, a tendency to overplay the importance of countries nearer the poles and downplay the importance of those near the equator. That wasn’t the intention of the projection, but it’s a product of *uncritical* use of that projection.

Something similar applies in the case of models. We’ve agreed that models simplify reality by leaving things out that are not judged important for some specific purpose. It just so happens that that’s also what political and cultural bias do: leave things out that are not judged important. That doesn’t mean that models are biased. But like maps, if we start using them uncritically, if we become over-familiar with them, then we risk allowing them to become laden with bias.

- Evie We should always be dating our models, never married to them!
- Yusuf That's a great way of putting it. Because I don't think familiarity is the only issue. In my experience, people become very personally attached to particular models. The model becomes part of their identity. You have to be very careful how you tell someone that the evidence simply doesn't support the model they are using.
- Paola I think there's another interesting dynamic that can come into play specifically among researchers. Evie, you made the point that models allow us to accumulate learning, by enabling us to compare and systematically review different studies. But it strikes me that there is also a potential trap in that.
- For example, suppose the first person to do a study on interventions to tackle a particular issue uses the Theory of Planned Behaviour. The next person to do a study in the same area wants to be able to compare their results, so they use the Theory of Planned Behaviour as well. Each subsequent study has a strong incentive to stick with that model. But no-one actually tests whether the model is actually applicable to the issue in question. For example, is the target behaviour actually *planned* behaviour at all? Did that first researcher choose the right model?
- Evie That's so true. Whereas if we had the kind of guidance we've been discussing in place, that first researcher could be encouraged to explain and justify their choice of model; and subsequent researchers might be encouraged to review that choice critically, rather than simply going with the flow.



- Paola I'd like to introduce a new question into our discussion. Does it make any difference if the thing we are modelling is human behaviour?
- I mean, we've been using examples from a wide range of contexts and disciplines – bridges, breweries, maps. At the Centre for Behaviour Change event, the speakers discussed examples like a drinks vending machine and rat navigation. I'm just wondering if it makes any difference what we're studying.
- Evie It's an interesting question. I can't see what difference it would make, though. In principle, modelling human behaviour is no different from modelling the behaviour of anything else. Ultimately what you're trying to do in every case is understand a causal network, and how certain things cause other things to happen.
- So in the case of behaviour change, what you're ultimately trying to do is map interventions to behavioural outcomes as closely as possible, via various internal states of the system. The challenge is in principle the same.
- Having said that, I recognise that human behaviour is significantly more complex than, say, a typical engineered system. But that just means we need even more care and discipline in how we go about modelling.
- Yusuf I'm inclined to agree with you Evie. Modelling human behaviour is no different in principle, but complexity is a huge challenge. I also think it may be worth distinguishing two different sources of that complexity.
- First, there's complexity in the inputs to the system. Whenever you build a model, you're trying to home in on the differences that really make a difference – the key variables. The challenge with human behaviour is that things you thought weren't important have a habit of making a big difference.

- For example, I remember hearing this story about a new product launch. The marketing team piloted the product in Scotland, where it was a huge success. They then went on to launch it across the country – and it failed spectacularly. When they tried to work out what had gone wrong, they came to the conclusion that it all came down to the small print on the advertising campaign they’d used in the original pilot. For legal reasons, they’d had to add the words: ‘Only available in Scotland’. The success of the pilot came down to a feeling of national pride which the UK-wide launch lacked.
- Evie I think the *kind* of intervention you’re making also makes a difference here. If your intervention is putting a drug into a person, then the task of deciding which things to leave out of your model is much easier. Still not straightforward, but easier than if your intervention is, say, a piece of advertising.
- Yusuf I agree. And then, on top of that, there’s the second kind of complexity: complexity in the system itself. In behaviour change, we’re rarely talking about a single, linear pathway between intervention and outcome. Instead, there is typically a multiplicity of possible pathways between the input and the output, many or maybe all of them non-linear.
- Even at the level of the individual, behaviour is the product of multiple mechanisms operating in parallel and interacting with each other – rational, emotional, instinctive. And that’s before we take into account the interaction between the individual and their social and physical context. Not to mention the differences between individuals.
- For example, take a relatively simple intervention like a GP giving advice to a patient on losing weight. The relationship between this intervention and the subsequent behaviour of the person receiving the advice is extremely complex. There are multiple possible pathways between the advice being given and the patient actually acting on it.
- Paola That’s really interesting, Yusuf, and it’s sparked off a couple of thoughts in my head. One relates to ‘the interaction between the individual and their social context’, as you put it – but perhaps we can come back to that one later.
- The other thought was about the layering of different mechanisms even within an individual. It just occurred to me that at the heart of our approaches to modelling are these ideas of simplicity and parsimony; but the things we are trying to model are the products of evolution, which often produces redundancy and complexity. Take the human eye, for example: no engineer would build it that way, but that’s how it’s ended up. So this thought just popped into my head: when we model human behaviour, is our desire for simplicity well-placed?
- Evie I’ll have to think about that one a bit, Paola. But I think there’s a difference between wanting a model to be simple and wanting reality to be simple. For me, the value of simplicity in models is that it helps us start to get a handle on complexity. A simple model is at best partial, but it still gives us a platform from which to explore the complexity further. The problem comes when people want the real world to be simple as well. They latch on to simple models as ‘the answer’. Maybe it links back to the points you were making earlier about risk.
- Paola That’s really helpful, Evie. I’m reminded of that quotation – I think it was Einstein who said it: ‘Everything should be made as simple as possible, but not simpler.’ The difficult bit is judging what’s possible and what’s *too* simple. I think we’re starting to

touch here on what might be *called* moral questions around the use of models. It seems to me it's not simplification that's the problem so much as the intent of the person simplifying..

I think you're also right to link it back to the topic of risk. Because the upshot of the two kinds of complexity Yusuf has described is that, with human behaviour, you always have the risk of unintended consequences, unpredictable and emergent outcomes. Which can be very uncomfortable if your job is to predict, plan and control.

Which makes me wonder.... Evie, you said a few minutes ago that the challenge for behaviour change is to map interventions to behavioural outcomes as closely as possible, via various internal states of the system. But given all this complexity, maybe that's an impossible dream.

I also wonder if this might be why some of the models in the field of behaviour change are less well specified.

Evie I'm not sure that follows. Yes, human behaviour is extremely complex, but that doesn't obviate the need to approach the world systematically.

Let's remember that complexity is also a feature of many physical systems. Take the weather, for example. Yes, there are always unexpected and unpredictable outcomes, but you can still create models with high predictive value if you collect the right data and model it in the right way.

The same is true with human behaviour. There are some areas where we have very clear evidence of overall patterns of behaviour in response to types of intervention. Like the formula for cigarette consumption price elasticity you mentioned earlier, Paola. Or the models linking sexual behaviour to the incidence of HIV which I was talking about.

So human behaviour is complex, but in some respects it is also very predictable. In fact, you could argue that it's the predictability of people that makes social life possible at all: we all have fairly good tacit models of the behaviour of others which guide decisions about our own social behaviour. Essentially what I'm proposing is a more formalised, explicit and rigorous version of that everyday understanding.

I agree we don't have formal models like the ones I just mentioned in as many areas of human behaviour as we'd like. There's plenty of work still to do. But that doesn't mean it's not possible in principle.

Yusuf I think there's one other issue you're missing here, however, which is time. You mention the example of the weather, but the predictive value of meteorological models declines pretty quickly over time.

The problem is that the gap between many behaviour change interventions and the intended outcomes is so long that it's arguably not even possible to talk about them as 'intervention' and 'outcome' any more. For example, suppose you're trying to establish the impact of Personal, Social and Health Education in schools on later sexual behaviour. It's a bit like trying to use a meteorological model to predict the weather in three years time.

I'm struck that the examples you give relate to quite immediate effects. For example, the relationship between price and consumption may not be a simple one, but I can believe that a change in price has a fairly immediate effect on consumption behaviour. Whereas something like education...

In fact, it's not just that the models are unlikely to be able to make predictions over these long time-spans. I'm not sure how you'd even gather data to build the model in the first place. Take the example of education again. Suppose you want to understand the long-term effects of educational choices on the economy. How do you gather the data? The dynamics of change are too slow for you to be able to sample the system. Instead, you have to start looking at proxies and predictors of outcomes: so, for example, you use exam results as a predictor of the longer term economic outcomes you really care about. But as soon as you do that, you introduce noise into the system. You're not starting with the high quality data you need to build a model of the kind you're describing.

Evie That's a really interesting point. That's going to be a much bigger issue for certain kinds of intervention than others. So the impacts of something like price may be much easier to model than the impacts of something like education.

Yusuf And I think there's a real danger that people start believing the interventions that are easier to model are also ones that work better. Because I think people sometimes mistake greater confidence in the outcome for greater effectiveness. We're back to Paola's point about risk: as human beings, the desire to feel like we're in control sometimes trumps the desire to actually get things right.

9

Paola I feel like the distinction we made between specific and generic models may be relevant to this topic of complexity. Evie, the successful models of human behaviour you've talked about relate to quite specific areas of human behaviour. When I then look at popular and well-evidence *generic* models, like the Theory of Planned Behaviour, it seems to me that they're rather different from what you're describing.

You see, I'm struggling to connect those generic models to the complexity of real human behaviour. Yusuf, you talked about a multiplicity of non-linear pathways. But it seems to me that our generic models of human behaviour typically show single, linear pathways.

Yusuf I'm not sure that's entirely fair. Many models explicitly include feedback loops. And they're simplifications as well: they're trying to get a handle on one just part of the complex picture. As Evie put it, they're platforms for further exploration.

However, I do think there could be value in taking a closer look at the different techniques that have been developed for modelling complex phenomena in other disciplines, and asking how we could make better use of some of those techniques in the field of behaviour change. I'm not sure we have a good enough grasp on what's in the modelling toolbox.

Because the reality is that lots of different disciplines are wrestling in different ways with problems that are too hard to model analytically. To take just one example, what could we learn from continuous flow dynamics and the way it uses iterative methods involving cellular automata?

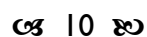
And that's just one example. I'd love to see a project that set out to catalogue the different type of modelling technique available, and their differing applications.

Paola Perhaps we're back to the third type of model I suggested at the outset – the provocations to thought.

Evie What I preferred to call metaphors! I agree, though, there could be lots to learn from looking more closely at the way in which other disciplines model other kinds of complex phenomena.

So sticking with the example of continuous flow dynamics, consider the flow of air across an aircraft wing. At the trailing edge of the wing, you have lots of chaos and unpredictability, meaning it's impossible to model what's happening on a moment-by-moment basis. But if you then pull out to the aircraft as a whole, you can be completely confident about the overall lift on the wing. The regularities only become apparent when you pull out to the bigger picture.

I think there's an interesting metaphor there for human behaviour. Maybe if we look at individual behaviour it's a bit like the trailing edge of the wing: you just can't predict what's going to happen. But if we pull out to the population as a whole, we may start to see patterns.



Paola That's a really interesting parallel. And it flags up another important issue for me – this question of levels of analysis. The models you were talking about, Evie, are precisely models of behaviour patterns at the level of a population. They're not saying anything about what's going on at the level of individuals.

Whereas most of the generic models we have focus on the individual and what is going on between their ears. They're the products of a psychological perspective on behaviour, not a sociological one.

Yusuf I'm not sure I agree with that. I mean, it's a point that people often make: but I'm not sure it's actually a very fair one. Let's start with this complaint that the models focus on what's going on between people's ears. I think this is just a straightforward misunderstanding of many of the models we have.

For what it's worth, I suspect it arises from an ambiguity in the term 'context'. On the one hand, we use the word context to refer to all of the variables that we leave out when we build a model. Context is the thing we abstract from. On the other hand, when we're talking about human behaviour it is often useful to distinguish attributes of the individual – things like beliefs or intentions – from attributes of their social and physical setting. And the latter often gets called context as well.

Historically, psychology may sometimes have been guilty of confounding these two things, and treating anything that wasn't 'between people's ears' – i.e. context in the second sense – as something that could be left out of the model – context in the first sense. But I can also point to a long tradition of social psychology that has taken context in the second sense very seriously in trying to model behaviour. Many of the models we have clearly reflect this.

Of course psychological models involve abstracting from context – in the first sense of the term. That's what every model does. But that doesn't mean they're ignoring context in the second sense. In fact, I don't see how anyone who's serious about understanding human behaviour *could* ignore the importance of social and physical context.

Paola But models like that still work by turning social phenomena into attributes of the individual. Take social norms, for example. In a popular model like the Theory of

Planned Behaviour, these become *subjective* norms, an individual's *perception* of a social reality.

Or take gender. People collect data on biological sex and call this gender. In fact, of course, even sex isn't binary as is usually assumed; but at least it is clearly a property of the individual. Whereas gender isn't: it's a product of the relationships between people. Gender only emerges as a meaningful construct if you look at the level of a community or society. If you try to reduce it to an individual characteristic, it stops being gender.

This is the other thought that was going through my head earlier, Yusuf, when you mentioned 'the interaction between the individual and their social context'. I suppose my starting point is a belief that the social represents a distinct level of reality, a distinct level of action and meaning, which needs to be modelled in its own right. It's more than just the sum of a series of contextualised individuals.

My concern is that if you only model at the level of individuals, you end up with a very partial picture. For example, I saw a study of a sex education intervention which targeted young people just before they became sexually active. At an individual level, the intervention had some clear impacts: levels of understanding increased, for example, and boys were less likely to pressurise girls into sex. But when it came to the hard medical outcomes – terminations of pregnancies, for example – there was no effect. And that, I would contest, was everything to do not with the individuals but with social processes.

- Yusuf I still don't really see what the problem is. I can account for that outcome at an individual level by noting that the theory of change which was behind the educational intervention clearly didn't pay enough attention to these contextual variables, and attributed too much importance to things like knowledge and attitudes. Which is a very common mistake, I agree. But it's not one that you're bound to make just because you model at the individual level. Making the individual your unit of analysis is not in any way the same as saying that only individual variables matter.
- Paola It's interesting. To me the point I'm making seems quite obvious; but I can see that to you, your take also seems quite obvious. I suppose we're bumping up against some quite fundamental philosophical issues here.
- Evie That's certainly how it sounds to me. I think you're getting into a question about reduction: can one level of analysis be reduced to another? And even that question comes in different flavours: ontological reductionism, epistemological reductionism.... Maybe we need to invite a philosopher of science along the next time we discuss these issues.
- Paola I agree. And yes, the thing I'm resisting is what I perceive as a tendency to reduce the social to nothing more than a collection of individuals. Which maybe is being unfair to you, Yusuf?
- Yusuf I'm not sure. Maybe I'm not really getting your point. It's another reminder of how important it is to keep talking about these things across disciplinary boundaries, and working hard to develop a common language.
- Paola Absolutely. For what it's worth, I think I'd actually distinguish *three* levels of reality when we try to understand human behaviour. So there's a micro-level, the level of interacting individuals; a meso-level, the level of communities and institutions; and a

macro-level, the level of an entire society. Modelling at each of these levels would bring something different to the overall picture. The dream, I suppose, would be to connect all three levels, and see how they interact with each other.

I think this matters when you look at what it takes to achieve behaviour change. Focusing too much on the micro-level, the level of individuals, leads in my view to an undue focus on interventions at this level as well. In reality, behaviour change typically involves interventions and changes across all three levels.

Take smoking, for example: at the macro-level, you had changes in the law; at the meso-level, changes in social norms and workplace practices; and at the micro-level, changes in the attitudes and behaviour of individual smokers. And all three levels interacted with each other.

Yusuf I think that's a good analysis. I also feel there's nothing in there that is problematic for the kind of approach I subscribe to. At the same time, it's clear to me that you feel my approach is too individualistic, that it fails to do justice in some way to these different levels of analysis, change and intervention. So I guess we'll just have to keep on talking about this one.



Evie There's another type of philosophical question which I think we may need to wrestle with. One point it seems we all implicitly agree on is that the study of human behaviour and behaviour change can and should be approached scientifically. A lot of our discussion of the role of models, and the interplay of theory and practice, has taken that point for granted.

But that raises a couple of questions. What do we think science is? And what bearing does that have on the way that we understand models?

Yusuf That sounds like a rather abstract point to me. Can you say a bit more?

Evie Well, let's think through a couple of examples. Suppose first that you're a Positivist. That means you believe that science is about observing the world, measuring it, and identifying regularities. If you take this view, then models are just ways of capturing the regularities you have observed.

If, however, you're a Popperian, then you turn this view on its head. From this perspective, models are expressions of patterns you *expect* to see. They're hypotheses, which you then go out and try to disprove by testing them.

A Critical Realist might take issue with both these views, and question whether there are any neutral, objective observations of the world. Models are something closer to interpretative stances – and you can never entirely step outside them and make 'objective' measurements. Your model shapes your observations, just as your observations shape your model.

Paola Maybe if you're a Kuhnian, and believe that science advances through the overthrow of paradigms in scientific revolutions, you'll see another role for models. It goes back to that third type of model I mentioned at the beginning – models which provoke thinking. The roles of models like these might be to throw up anomalies, unexpected results, the kinds of things that can lead us to question paradigms and initiate scientific revolutions. Maybe models play a role similar to

- that which the philosopher Ian Hacking ascribes to experiments: not testing theories, but creating new phenomena about which to theorise.
- Evie It all depends what your view of science is. What is it you think we're actually doing here? How do we gain new knowledge, and what's the status of that knowledge once we've got it?
- Yusuf And that in turn could shape how you go about building and using models... I think I understand the point you're making now.
- Evie It's a question, really. I'm not sure whether it does make a difference. But if we're thinking about building that common framework for talking to each other about models, we need to explore the possibility.
- Paola If we're doing that, then I think we should also re-examine the central role we've been giving to prediction. Evie, you suggested we focus our conversation on models that enable us to make predictions we can test, and you argued that that is central to the whole endeavour of science. Narrowing the scope in this way has been really helpful for today's conversation. But I've had this nagging doubt at the back of my mind: is prediction really the right goal for a science of behaviour change?
- You see, I could argue that the fixation with prediction is a relatively recent phenomenon – that we're all still in the thrall of Newtonian astronomy, with its beautiful mathematics. But planetary bodies don't think. People do: they reflect on their own condition and environment, they have *their own models* of the world, which they also adjust in response to experience, and which influence their behaviour. Human behaviour exhibits a level of complexity that, I'd argue, is qualitatively different from anything else we encounter in the world.
- And even if I'm wrong on that specific point, we all seemed to agree that the complexity of human behaviour is such that it's very difficult to predict anything beyond short timescales and automatic processes. So I'm still left wondering what's left of a science of behaviour change that puts prediction centre stage.
- I'm not saying we should throw up our hands in despair, just that we might want to reconsider what we're trying to do. It seems to me that, even if we can't predict, we can still *explain*. We can understand, forensically dissect. We just need to recognise that those are different tasks, and stop fixating quite so much on prediction.
- Evie I don't know if I really understand this distinction you're making between prediction and explanation. I mean, I know the words have different everyday meanings. But how do you make that work in a scientific context? How do you know your explanation is correct if you can't test it in some way?
- Paola I suppose we might have to look at something like evolutionary science to get an answer to that question. After all, there is plenty of debate about different models of evolution, conducted on a clearly scientific basis. But so far as I'm aware they don't make any predictions.
- Or better still, perhaps we should go and talk to historians. They are clearly in the business of explaining behaviour based on the evidence: but again, no predictions as such.
- Yusuf As ever, Paola, they're interesting challenges you make. But I think I'm with Evie on this one. After all, what we're talking about here is not just explaining behaviour

- but *changing* it. If you can't establish some kind of predictive connection between intervention and outcome – be it ever so uncertain and probabilistic – then how are you going to decide what to do? What's the point of explaining what's happened in the past if it doesn't help you do things better in the future?
- Paola Well, I supposed I'd turn that on its head and say, if in fact we *can't* predict the future, what do we achieve by *pretending* that we can? We're just consoling ourselves with an illusion of control.
- Besides, it seems to me there's a very real sense in which we can learn or fail to learn the lesson of history, without being able to make predictions about what will happen next. If nothing else, history tells you something about what *could* happen. We can't predict earthquakes, for example; but we know where to build buildings that can withstand them.
- Maybe that's the kind of role models can play too. Not as predictive tools that give us control over the future, but as mechanisms that allow us creatively to explore different possible dynamics, and prepare for a future we can't control.
- Evie And where does that leave the idea of behaviour change?
- Paola I don't know! We'll have to pick that one up the next time we meet. But it's good to end on a note of controversy, don't you think?
- Evie Absolutely. Plenty to take away and think about, and plenty to talk about next time.
- Yusuf Agreed. I feel like we've made a lot of progress, but I'm still left with a great many questions.
- Paola If nothing else, I think we've demonstrated the value of discussing those questions across disciplines.