Video 7: Sensing – Actions

Alva Noë: When I was a graduate student, I was a philosophy student, and actually just down the road. And I was steeped in the study of vision with scientists. And I remember I sat down one day with a vision scientist. I'm sorry. Excuse me. I sat down with an artist. That's the crucial part of my little anecdote.

[00:00:22] I sat down with an artist. And he asked me, well, what is vision science about? And I said, well, in vision science, we try to understand how it is that we enjoy an experience of the three-dimensional, spread-out, colorful world around us when what we're given are tiny, upside-down, distorted images in the eyes. How do we see so much on the basis of so little?

And he pondered. And he furrowed his brow. And then he blurted out, nonsense. The interesting question is not how we see so much on the basis of so little but why we see so little when there is so much around us to see.

And I offer that as a motto for my presentation. I hadn't planned on mentioning it. But I was inspired by Natasha's setting up a sense of opposition between me and traditional cognitive science in her introductory remarks.

I have discovered, as I've gone on over the past years in my own work on the nature of perception and perceptual consciousness and consciousness, that I keep coming back to the insight-- or at least to what I take to be the insight-- that the artist was right in that exchange. And it's something I also wanted to mention because in some sense, I feel it positions me as a philosopher in relation to artists. I continuously find not that my theoretical or academic training gives me insight into what they're doing, but rather that conversation with them teaches me something about the phenomena that I'm interested in.

This is not what I wanted to talk about. My real topic today is I wanted to say something about organization and reorganization, and the relationship between those ideas and the work of art. Very much responding to remarks that Bruno Latour made yesterday and also Tomas Saraceno's work.

And so I thought I'd begin by remarking that Plato once said that it's easy to make a picture. You just hold up a mirror. And in one sense, that's an obviously batty remark. There's all the difference in the world between a mirror reflection and a picture. We make pictures. We stumble upon or find reflections in the natural world.

But in another sense, I think, Plato was onto something. And what he was onto was the idea that although a picture is one thing and a mirror reflection is another, it may be that our relationship to mirror images, and in particular our relationship to reflections of ourself in mirrors, is shaped by the existence of pictures. And I've been thinking about this a lot lately in relation to the work of the historian Anne Hollander, who many of you -- who you'll know, and who passed away just a few weeks ago.

Her book from the 1970s called *Seeing Through Clothing* is a brilliant exploration of this idea that pictures shape the way we see. So she had the idea-- and she just lays it out in this marvelous way looking at the history of art. She had this idea that the way we think about the dressed human body, the way we think about the clothed person is shaped by depictions of the clothed person in works of art.

And so when it comes to thinking about mirror images, she talks about the way in which, when we look at ourselves in the mirror-- you're standing in the vestibule, or in the bedroom mirror, or in the bathroom mirror-- you don't just see yourself in the reflection. You frame yourself. You make a picture. It's almost as if we make a kind of provisional portrait for ourself.

Now she wrote this book in the '70s. She didn't have the concept of the selfie. But I guess she could've said, it's as if mirrors are selfie opportunities.

And she has this beautiful phrase that she used that I think of often, especially in these last weeks after she died. "The picture is the standard by which the direct awareness is measured." We make sense of what we see when we see ourselves in the mirror with reference to the picture.

Now there are two ideas in connection with that that I wanted to share. The first is a kind of fairly straightforward one, the idea that pictures shape perceptual consciousness, that we use pictorial ideas in thinking about the way we see. And it's almost a point at the level of ideology, as if an ideology of what the pictorial world is shapes the visible world.

[00:05:28] But what interests me in this discussion of hers, which I recommend to you, is a further thought that is not often stated. And it is that because of this, because pictures shape or provide the standard by which the direct awareness of our own dressed bodies are measured, it turns out that our concern, our everyday, daily concern with dressing and with how we look is, in a way, directly a concern-- or maybe I should say indirectly a concern with art. As if dressing is, in the way that looking at oneself in the mirror is, a way of exploring possibilities in the pictorial domain.

And that's beautiful. That means that the teenager worried about the sagginess of the jeans is kind of working in the vicinity of art. Maybe it's not art. But it's important in a way.

And it's also interesting, because I said that the first point, namely that pictures shape the way we see, you might think of that as pictures govern us. Pictures dominate us. And certainly in our hyper image-dominated world, there's a political anxiety there about the control that the picture has over us.

But the beautiful thing in this other idea is that pictures may govern us, but pictures might also give us the resources to dress differently, to conceptualise dressing as something suitable for pictorial depiction. And so we're governed, and were emancipated, not by stepping outside, not by ignoring pictures but by using them to

our own ends and devices. And I'm tempted to think that-- and I haven't really gotten there yet. But I hope you'll see connections between what I'm saying and what Bruno Latour said yesterday. I'm really struck by the idea that one ought to extend Anne Hollander's thought to seeing more generally, in pictures more generally. So there's a sense in which pictures really do provide the standard by which visual awareness, not just of the dressed body, but visual awareness of the visible world is measured.

Our idea-- and you see this, I think, in natural science or neuroscience, but you also see this in every aspect. When we reflect on seeing, the first thing that comes to mind is a kind of contemplative act, something like the contemplation of a picture. And the second thing is our conception of the object of perception is, as Bruno Latour said yesterday, something like the conception of the still life. As if what an object is for visual perception is this thing which is frozen and is selected as suitable for contemplation.

And the one thing we cannot do, the one thing we cannot do is ignore the way in which pictures shape the way we think about the visual world and our experience to the world through vision. We can't just repudiate it. We can't say it's false. It's a falsification of what visual consciousness is, because it's deeper than that. Its organizing role in our lives is deeper than that.

And yet I think we can change the way we see. And we do, throughout historical time, change the way we see. Remember, pictures are at least 30,000 years old. And so is dressing at least 30,000 years old. This is not a new trend, a new preoccupation.

And I think it is-- and this is sort of the main thought I want to share with you. I think it is the distinct concern, the distinct value, the distinct job of art to bring about that kind of reorganization. I think it's the job of art. I think it's also the job of philosophy. I don't think it's the job of natural science. So I want to say that there really is something special for art to be doing.

Let me now take a step back and try to say something to make this a little bit more concrete. So let me give you a concrete example. This is all by way of giving some background to what I just said. I only have seven minutes. So I think a good place to begin is with breastfeeding.

Human beings are mammals. And like all mammals, we breastfeed. But unlike all the other mammals, we're very bad at it.

Our infants get distracted. They fall asleep. They start chewing. They in some other way interrupt the process.

So you've got this activity, this breastfeeding activity, of suckling, falling asleep, jiggling, bringing the baby's attention back to the task at hand, suckling again, getting distracted by a noise, and so on. It's this really delicate interaction that happens in time and in space between the feeder and the child.

[00:10:20] Let me call your attention to six features of breastfeeding. One is that, in an obvious, straightforward sense, it's natural. It's basic. It's primitive. It's biological.

Two, despite being basic, and primitive, and biological, it's also at least an arena for the exercise of quite sophisticated cognitive capacities. Not just doing and undergoing, but paying attention, and having one's attention directed, and perceiving, and guiding, and interacting with another.

Notice that it's really got a temporal dynamics. It's almost got the dynamics of turn-taking. It's one of the first examples of turn-taking that I can think of in human life. And it's been suggested-- this is not original to me-- that maybe it's not an accident that the linguistic species is also the people that negotiate breastfeeding in this kind of turn-taking fashion, almost as if breastfeeding is a kind of primitive conversation. It's laying the groundwork for the possibility of doing something much more complicated later on called conversation.

Here's the fourth point. Neither mother nor child is in charge. They don't author this dynamic.

Here's the fifth point. It has a function. Breastfeeding has a function.

Presumably, the function is feeding. Although some have argued we're so bad at it that maybe the function isn't feeding. Maybe we feed that way in order to achieve some other function, like attachment building, or relationships, or indeed just the participation in the activity itself.

And then the sixth function-- the sixth point, rather, which is a very important one-- it's the one I understand the least well-- is that breastfeeding is at least potentially very pleasurable. Potentially. It's not always pleasurable. It can be a very anxiety-producing event as well.

The reason I'm going on at this length is I want to introduce the notion of a technical term of "organized activity." And I want to say that any activity that has these six features is an organized activity. And by the six features, I mean primitive, and cognitively sophisticated, and temporally dynamic and organized, and having a function but not being authored, and being a source of pleasure, OK?

Now in a longer version of my presentation-- and I had sort of mentally hoped to be able to give you lots of other examples of organized activities. But I'm sort of protected by time. Instead, I'll just make a few dogmatic assertions.

One is I think our lives are, almost everywhere you look, organized. They're complicated, structured nestings of participation and organized activities in this sense. We are not the masters of this. We find ourselves organized. Organization, in a way, is, I think, a crucial biological notion.

Again, I don't think that organized activity is something that you can understand individualistically. If you want to study it scientifically, I think that the relevant time scales are not going to be the time scales of the nervous system, not milliseconds. Nor are the relevant time scales the times scales at which conscious decisions and deliberate action are performed. It's some intermediate level, something like what has been called the embodiment level.

And it's biological. Any biology of organized activity-- and I think we need a biology of organized activity-- is going to be a non-individualistic and non-reductionistic biology.

Here's another organized activity. Dancing. I think it's intuitively obvious that it meets my criteria.

Dancing has a kind of spontaneous naturalness to it. And yet it is such a subtle communicative act of paying attention, interacting, watching, paying attention, noticing. Paying attention to the other, paying attention to yourself. If you're dancing to music, paying attention to music.

It's temporally organized. It's got a function, presumably, if you think of dancing in a traditional setting, in weddings, funerals, or in sort of social settings of celebration, or like dancing at the discotheque, courtship, or seduction, or just play. There are all sorts of functions that dancing presumably has.

And it obviously can be pleasurable, if you enjoy it. And it's not authored. When you dance, it just sort of happens to you. You find yourself dancing. Or the dance-- you are danced, one might say, by the dance.

Now what's really interesting to consider with this machinery that I've given you now is, what about choreography? Choreography, I think it's very important to notice, is not just more dancing. Choreography is not just the participation in an organized activity, dancing.

[00:15:09] Well, you might say choreographers make dances. But on the way I want us to think about a dance, you don't make a dance. The dance happens. What a choreographer does, you might say, is stage a dance.

Well, what does that mean? That means put a dance on a stage, or put a dance on display, or put dancing on display, or maybe put the fact that we dance on display, or put the fact that dancing organizes us in something like the way breastfeeding organizes us. And to give other examples that I didn't have time to discuss, the way talking organizes us. Choreography puts that on display and reveals something to ourselves about the ways in which we find ourselves organized.

And then, I have just a couple of minutes. But here's the beautiful point. From what I've just said so far, you might think you have dancing over here, a kind of first-order organized activity, and choreography over here, the representation or a display of it.

But of course-- and this is related to the point I made earlier about Anne Hollander-- the choreography loops back down and changes the way we dance.

In a world in which there is such a thing as choreography, there's no such thing as dancing immune to its image. We know what dancing looks like. In fact, we can't dance indifferently to what it looks like. There's many, many different choreographic models, whether it's Michael Jackson or Baryshnikov. There's many, many different pictures-- to go back-- which give the standard by which our direct experience of the phenomenon is mediated.

And so what starts off in some sense as a coping with the fact that we find ourselves organized by this phenomenon-- what starts off as an attempt to cope with that, to understand that, that yields art. That yields, in this example, choreography. But it loops back down and changes what we do at the first order.

Just to fix the idea, a very nice example is the case of writing. Writing is one thing. Speech is another. But we cannot speak as if there were no writing.

That's a marvelous thing. Linguists always talk about, we've been speaking for 30,000 years. Writing is this recent cultural thing. But our conception of our own lived life as speakers is shaped by our understanding of writing. And so this goes then to the idea of how it might be possible to go from organized activity, which is what I call breastfeeding and perceiving and talking and dancing, to art and to philosophy, which I think of as re-organizational practices, practices that really change what we are.

I've been given the time signal. But the final thought here is that-- I'm repeating myself. But I think this is so important. I cannot choose not to be organized.

I cannot choose not to be organized by language. I cannot choose not to be organized by pictures or pictoriality. I cannot choose not to be organized, when I'm a dancer, by the ways we have dances and the ways we do dance. And that's a source of stress and anxiety. We just find ourselves organized nest within nest within nest.

I think philosophy and art are responses to that, not by dogmatically saying, stop, you don't have to dress that way. You don't have to dance that way. But by simply, from within, renegotiating what it is to do it. And thus, we get something like a liberation or an emancipation. Thank you very much.

Natasha Schüll: OK, we're going to go right into our next talk. And then we'll come up and have some Q&A. So Josh Tenenbaum from MIT.

Josh Tenenbaum: OK, everybody hear OK? OK, thanks. My name is Josh Tenenbaum. I'm a professor here at MIT in the Brain and Cognitive Science Department, like some of my colleagues who you've seen earlier in the symposium.

It's a great privilege to be here. Thanks to the organizers for making this and having me. It's sort of an undeserved privilege to be going towards the end where I can sprinkle my

talk with comments and reflections back on some of the other speakers, including the last ones in the session and also back to last night.

So in keeping with the sensing theme, I'm going to be talking about some version of sensing and also common sense, going back to Bruno's talk and something I think we're all deeply interested in. And very much in the same spirit that you heard from Tommy Poggio yesterday, I'm also involved with the Center for Brains, Minds, and Machines.

[00:20:10] This is a reverse engineering talk. It's a science and engineering talk. But I hope it will have something to say to the humanities and the arts.

I'm going to be talking about what I call reverse engineering the common-sense core. And I think you can make a contrast here, much as Tommy did last night, with the situation that we all are familiar with in technology and the technologies that have transformed our daily lives. What you could call AI technologies, things like these pedestrian detectors, or these amazing question answering systems, self-driving cars, even, now. Technologies that anybody, including the founders of AI-- people like Turing, who've been mentioned here-- Leibniz, you could go back there-- would have to acknowledge our remarkable achievements.

Yet we all know none of these are truly intelligent. So what's missing? That's what we're trying to understand. What's missing?

And for me, my focus here-- and it's also the part of the Center for Brains, Minds, and Machines that I'm working on-- is this, what I like to call the common-sense core. Now this isn't my idea. This is an idea that's emerged over the last couple of decades from a number of different scientists, cognitive scientists broadly construed, including linguists. I think some philosophers, developmental psychologists, vision scientists. There's a lot of connections here to what Alva was talking about, both his earlier life and his current life in studying perception and phenomenology. And computer scientists of many different stripes.

I think this is the key idea, that from the very earliest ages-- and I mean that from early infancy, maybe even from breastfeeding days-- human thought and action is organized around a basic understanding of the world in terms of these key concepts-- physical objects, intentional agents, and their causal interactions, or what you might call intuitive theories of physics and psychology, a kind of intuitive physics and an intuitive psychology. And by "intuitive theory"-- this is, again, sort of a term of art in cognitive science-- I mean systems of concepts and abstract knowledge that, much like a scientific theory, are not just a collection of facts but are principles that can be applied to an endless number of new situations that you find yourself in.

Let me illustrate this by actually talking about some sensing problems. These are problems that, again, many different cognitive scientists are interested in. Problems of, say, what the pictures on the left sort of illustrate, what I think of as intuitive physics in seeing, understanding, and action. And the ones on the right are kind of intuitive psychology.

And you look at these. It's not just that you can, as current engineering systems can do, detect the people or detect the objects. It's not just what is where. But it's understanding what will happen and what can be done. The possibilities for action and prediction and counterfactual reasoning, what you could have done differently.

In that workshop scene up there, you don't just see a crowded world of objects. But you know that the table is supporting the other objects on it. If you were to remove the table, the other objects would fall. If you were to bump the table, some of them might fall off. Maybe if it's just a gentle bump, the tire, the round tire that's leaned up against it precariously would roll off, but the others would be fine.

Or that scene over there on the lower left of the house being constructed with the wood. You don't look at that and see any of the nails, the things which are attaching the pieces of wood. But you know they're there, because otherwise it would be falling over. So how do you perceive the invisible nails?

Or here's a scene from one of our local cafes, Area Four over there, which the barista described-- this as sort of a dish bin with a bunch of coffee things in it-- as a disaster waiting to happen. You know instinctively that if you were to go pick that up and bus those dishes, you first have to rearrange them, or the whole thing is going to go tumbling over onto the floor. So how do you know that in a glance without even really thinking about it consciously?

Or these scenes over here on the right. I think these are organized activities, more advanced ones, but the same kind of thing that Alva was talking about. You don't just look at these and see people. But you see a whole kind of intuitive psychology. You get a sense for what they're thinking, what they want from each other, what their roles and expectations towards the others might be.

I really like the street and crossing guard scene down there. Think about, when you look at that, what's going on inside that crossing guard's head. You know where she's looking, what she's expecting to see, what she might be worried to see, what she's looking to see, what she's thinking about the people on the other side of her who she can't see, and what they can see and what they can't see. What might somebody else, a driver of a car coming be able to see or not see?

It's this kind of understanding of the agents in the world around us and their mental states that's at the heart of common sense. And it's the sort of thing which, on the technology side, I think everybody knows is a key frontier.

This is a quote from some of the tech leads of Google's self-driving car project where they say, "The thing we've got covered is the sensors. We're not missing better lasers or scanners or images. But it's the common sense here." As they say, "We do a good job of detecting pedestrians at the side of the road. But we don't yet have built in the kind of intuition for what a pedestrian might do."

That's the kind of thing which, as stressful and anxiety producing as it might be to be a teenager or the parent of a teenager learning to drive, you're not so worried about this. This is the thing that just comes automatically from being a human being. So how can we understand this?

[00:25:09] Now as I mentioned, I've been very influenced by developmental psychology. And I think that you can look and see this most clearly in young children playing with blocks, playing with each other, playing with each other playing with blocks. These are some of the basic organized activities that, again, a little bit more advanced in life than breastfeeding.

But I think you see many of the same kinds of themes. I think they basically fit all those six criteria. And they are natural activities. Might not have been blocks back in the Pleistocene. It might have been stones. We'll see a little bit about stones later. But it doesn't really matter.

Or when you look at a child learning about physical objects for the first time, like magnets or silly putty or a touchscreen device, contrast this with what you might be familiar with in the state of the art of machine learning with big data, where you have hundreds and thousands of examples and so on. A child here is going to learn much more than any machine system gets from just one example.

Just the right combination of the first time you encounter one of these touchscreen devices and you touch the screen at the right point-- none of you can really see what I'm doing. But we all have these in our pockets. Just the right spatio-temporal coincidence of how and where you touch, and when you touch, and what happens on the screen. The first time you saw that, it takes you by surprise. But then you know what's going on. And you have at least some idea of this new causal mechanism in the world.

How does that work? Here, I think, are a couple of movies I want to show, which really highlight the challenge in the most-- sort of lay it bare in the most compelling way to me. I think of them as kind of minimal sensing examples. But yet they fully engage your common sense.

So the first movie is one on the left here. And again, these are kind of organized activities. The one on the right, I think is a better and more famous example. But this is from two developmental psychologists, Southgate and Csibra, from a study of common sense in 13-month-olds.

Now when I play this, you'll see, hopefully-- you can see this is a blue ball and a red ball rolling on a green background, and some other objects there. But hopefully you don't just see it as a blue and a red ball rolling. You tell me, how do you see this? How would you describe this? What's the activity?

Yeah, chasing. The blue ball is chasing. And the red ball, is it doing anything? Yeah, running away or fleeing, or something, right.

It's a bit of a competitive interaction. But it's a kind of organized activity, maybe sort of a dance. But very goal directed.

Now here's another question for you. Which ball is smarter, the blue ball or the red ball? How many people say the blue? How many people say the red? OK, everybody says the red. Now why is that?

That's a judgment about a certain kind of mental character of mental states, right? Well, to understand what's going on here, to see those goals of chasing and fleeing, you have to understand also something about the other kind of mental state, the belief mental states. The blue ball has some correct beliefs. It seems to know where the red ball is.

But it also has some false beliefs. It thinks it can fit through those holes, which it can't fit through. And not only that, it persistently holds those false beliefs even with lots of evidence to the contrary. So that failure to learn registers on you as not being very smart.

And of course, behind this kind of intuitive psychological analysis is also a kind of intuitive physics. It only makes sense as that kind of goal-directed action, as chasing and fleeing, because of the physical constraints of the blocks, the fact that the balls can't pass through those other objects and that they know that-- they both know that, except for the holes part. If you were to remove those physical constraints, you would no longer see this as chasing and fleeing. You'd actually see it as more like dancing, which is a different kind of activity.

It's necessary to see this as a kind of efficient-- like, that the blue ball is trying to move along what it believes to be the most efficient path to the red ball. That principle of efficiency, as we'll see in a bit, is really at the heart of intuitive psychology. But it depends on physical constraints. So it's got intuitive physics wrapped up in it.

Now how many people have seen this famous Heider and Simmel video here? OK, you should all watch it. It's one of the most famous and most important movies in psychology.

It dates from the 1940s. So it's very low-quality stop-action animation filtered through several generations of analog and digital recording and re-transfers. So excuse the low quality.

But when you look at this, again, you won't just see some shapes in motion. But you do see characters in a whole story. So you could see two triangles and a circle. But it looks like there's a bit of a competitive interaction.

The big one's sort of bullying the other one, backing him up against the wall. Now he's a little bit scared and running away. The circle's kind of-- was watching and is now hiding.

And now the big triangle goes and goes to try to find the little circle. Cue the scary music. There's no soundtrack. But you can imagine it, just like with the interaction between perception and art there.

If it gets a little scary here, don't worry. It ends happily, at least for two of the three characters. You can watch the rest of it on YouTube. Just Google Heider and Simmel.

So what's going on in these videos? You see, as I've suggested, a kind of intuitive physics, intuitive psychology. Also, the one on the right there, think about what was happening at the beginning of the video. You see it as a kind of a fight or a competitive interaction because you see these force impacts. You see the big one banging into the little one. You see him backing him up against the wall that's immovable.

[00:30:11] So you see forces, which sometimes are not really there. These are just shapes being moved around from one frame to the other on a board and filmed from the top. But you see it in terms of those forces.

And then on top of those forces, you see all these psychological states and even other kinds of intuitive social sciences, like intuitive sociology and morality or ethics. And from what I like to call, in my version of what Alva said-- which is he really did characterize very well the sort of modern paradigm of cognitive science, asking not just in vision but in common sense more generally, how do we get so much from so little? I often start off my talks asking exactly that question.

Here you can ask that question. How do we get so much meaning from just so little? It's not just very simple images of a few shapes. But if you were to characterize them, like with a kind of mathematical description that you saw maybe in Josh McDermott's talk, like time series of signals, these are very much simpler signals than the sound textures that he was playing this morning.

It takes only nine or 10 numbers to describe these movies over time, nine or 10 numbers to describe the horizontal and vertical position and the orientation of each of these several shapes. Yet from just those nine or 10 dimensions going over time, much less information than in a musical score or in a cocktail party recording, you get so much. So how is that possible?

I'd like to come back in the question period to what Alva also said about art, which is, I think this is totally compatible with the mission of art being to say, how do we get so little from so much in other kinds of settings? But I'll come back to that later if we get the chance.

understand this knowledge in what I call "reverse engineering," a term that just means in the same kinds of computational terms we'd use to build an intelligent machine, a robot, both software and hardware. We're really interested in how this comes to be, its origins in the infant mind in some combination of nature and nurture. What's innate and learning from experience?

We're interested in how we can fix it when it goes wrong in certain kinds of developmental disorders like autism or others. We're interested in how we can expand on this kind of common sense in education and policy and the arts. And very much, I'm very interested, and many of us here at MIT, in how we can get this kind of thing into machines.

I'll just give you a taste of how we're doing this. In introducing me and introducing the session, people mentioned the idea of Bayesian inference and hierarchical Bayesian models. And people have used the word "priors" a lot. And I can't give you anything like a really technical treatment here.

You might have heard of these things called Bayesian networks. And these are a kind of technology that's, over the last couple of decades, transformed many areas of science and engineering. And if you haven't seen these before, just look at this one up here. And a network is a directed graph, circles and arrows. And just think of these as ways to represent causal structures in the world.

The key idea here is your mind is making probabilistic inferences, or just kind of good, plausible guesses with-- these intuitive theories are basically descriptions of the causal processes out there in the world, your mental models of them. So this directed graph up in the upper left should probably be basically familiar to us.

It's meant to capture very, very roughly what goes on inside of doctors' heads when they're making a diagnosis. You have diseases and symptoms or, just more generally, causes and effects. You observe a pattern of symptoms on the bottom. And you want to reason backwards to the pattern of diseases most likely to have caused those symptoms.

And of course, you can't be sure from any sparse observation of a couple of symptoms. But you can make a good guess if you have the right causal models and you put the right probabilities on here. And when we talk about doing Bayesian inference on a causal model, Bayesian inference is just taking a causal model, a probabilistic one, and kind of running it in reverse to make good guesses to the inputs, or the causes, that best explain a pattern of observed outputs of this causal process, or effects.

Now we do the same kind of thing in more sophisticated kinds of probabilistic models, which we call probabilistic programs, to capture common-sense physics and psychology. And they also kind of look like these directed graphs. But notice how I've put words on the arrows here. And that's because the arrows are not just sort of stand-ins for tables of numbers. But they're actually computer programs to capture the interesting causal processes that our mind-- the ways our minds capture the physical and psychological processes going on outside in the world of objects and agents.

We have found it most useful to build causal models on programs that other people in computer science have built for describing the same things in the world. So computer graphics and robotics are areas of computer science where basically what people do is write programs that describe how the physical world unfolds and how it looks in

images. And computer graphics people try to do this in ways that make very efficient approximations very fast, particularly the kind of computer graphics that's used in video games.

Sometimes we describe this as a view of like the video game engine in your head. Think about-- probably all of you are familiar with video games like the kinds that we have on our phones, these sort of physics-based video games. How many people have played some kind of physics game? OK, yeah, most of you. If you haven't, your kids or your grandkids, or maybe your grandparents have.

[00:35:06] Games where something like you have to stack up the blocks before they fall over, or fill this thing up with water. Some of them might even be non-Newtonian fluids. Cut the rope, capture the swinging.

Think about other kinds of games, like maybe it's sports video games or shooting video games. These are-- in order to create these kinds of immersive experiences and have them respond to human users in realistic ways in real time, you have to capture in computer programs things that go on in the world approximately. You don't actually have to capture all the things that are actually going on inside a soccer player's head, or all of the detailed physics of how a soccer ball bounces off the field, and how light is reflected and shadows, all the stuff needed to make just a freeze-frame of a movie of a soccer game look realistic enough to immerse you in the experience. But you have to capture it somehow enough.

And we think that in some sense, your brain is doing the same thing. Your brain has similar kinds of models of how objects move, how light bounces off things, sort of graphics, basically, and how agents have goals and pursue them. We can use the same kinds of programs that have been designed to build this. But by putting them in one of these Bayesian frameworks, we can run them backwards.

So you observe the outputs of these programs. And you make guesses about the inputs. In other words, you observe the experience that you're in.

It's not just observation, of course. It's interaction. But you get the data of your experience, which is like the video game of your life. And then your brain is making guesses about what were the inputs to these programs, things like the mental states of the other agents and the way the physics are unfolding. And that's, I guess, all I can do in this short talk to just give a sense of the kind of models.

I'll, in the time that I have left, just give you a sense of how we use these models and some of the implications of where you go for this. So we do experiments-- now it's sort of simpler and stripped down here- to test the intuitive physics engine in your head, if you like, where we show people a bunch of stimuli like these. Each one of these frames is one stimulus from an experiment.

And you might make a simple judgment, like, on a scale of one to seven, how likely do you think it is that the stack of blocks will fall under the influence of gravity? Probably

most of you agree that the ones on the upper left look pretty stable, whereas the ones in the lower right are likely to fall. And indeed, that's what most of our subjects think.

This is an example of the data from an experiment where we asked maybe 10 or 20 subjects-- we take their judgments, average them. And we're plotting along the vertical axis the mean judgment of a group of subjects. So each one of those crosses is one of these towers. And a cross which is high up, near one, means it's very likely to fall, the whole tower. A cross which is low down is one that is very stable.

And then along the x-axis, we're plotting the predictions of this model that I sort of sketched out to you. Basically what it's doing is it's trying to-- it looks at the image, makes a guess at the three-dimensional positions of these blocks, and then runs a simple sort of Newtonian physics simulator a few time steps forward in time under a couple of-- makes a couple of different guesses, because it's not exactly sure where the blocks are. And under that uncertainty, kind of sees what's likely to happen.

And what you get there in the x-- the horizontal position of each cross is our model's guess of how likely this stack is to fall. And you can see it does a pretty good job of predicting people's common-sense judgments here. There's a few stimuli shown in those colored frames. And they correspond to those dots there.

So the one that's in red is an example of one that people think is very unstable. And indeed, our model says it's very likely to fall. Interestingly though, it's actually not. It's actually-- this is sort of an illusion.

If you were to actually perfectly measure all those blocks in the red frame there, the very tall, precarious one, and perfectly localize them and perfectly simulate physics, it would be stable. It's much like these things.

You might have seen these other kind of physics illusions here if you've spent time out on the coast or in art galleries. This is a pretty artistic rendering of one of these rock mountains where-- how many people have seen things like this, either in the real world or in pictures? They're really cool.

Hopefully this doesn't look stable to you. It looks like the rocks should be falling. But they're actually stable.

They're just balanced exactly right. If they were just slightly in a different position, they would be falling. But the art here is to get them into a position that makes them look balanced.

Your brain thinks this is unstable the same way it thinks that thing is unstable, because probabilistically, it is unstable. It's not generic. There's just one precise alignment of these blocks in the bright red frame or these rocks here that works. But your intuitive physical intuitions-- your physical intuitions are only probabilistic.

Now we've actually been studying-- this is, by the way, work mostly done by Peter Battaglia, who was in my group until recently. Now he works for a particularly interesting part of Google called DeepMind. And he's actually extending these sort of ideas to look at a range of other kinds of physical systems, not just solid, rigid objects, including fluids, including even non-Newtonian fluids. But I'll talk about that later. I think this is a very interesting place where the cognitive science and art can interact.

[00:40:05] We've also studied the same kinds of intuitive physics in babies by showing them much simpler displays of things bouncing around. And I won't go into any of the details because I only have a few minutes left. But this is an example of a stimulus you might show a baby.

It sees a few objects bouncing around. And then after a moment, it'll be occluded. And one of the objects will appear. And you can essentially ask a young baby who can't speak, how surprising is that? under their intuitive physics. And you can vary different factors, like whether it's the blue or yellow one, or whether it's the one that's close to the door or far from the door, or how long the delay is between when it's occluded and when they come out.

And it turns out you can quantitatively predict, as this graph here on the right shows, infants' looking time, which is the standard measure of surprise. Infants look longer, just like you, when they see something surprising. You can predict that with one of these probabilistic intuitive physics measures. And this is really one of the very first times people have built any kind of quantitative predictive model of anything that infants do. I think it's a nice example of how these probabilistic programs can be taken all the way down to the core of common sense even in very young babies.

In the center here, we're working on trying to reverse engineer how this kind of physical knowledge grows. But I won't go into that here. Sort of future work.

The same kind of model can be applied to the intuitive psychology case here. But now we add into these physics programs planning programs. What's a planning program? Well, again, I'll point to a phone.

You're all familiar with what Google Maps does, or Waze, or navigation things with GPS and maps, where you basically say, I want to go here. And it knows where you are now. And it has a map of the city. And it plans out a route for you. A self-driving car does a fancy version of the same thing.

That's basically what any-- at the heart of any robot acting in the world is this rational mapping from what you could call beliefs and desires to actions, where beliefs include your general model of the world, where you are, your state. Your desire is where you want to get to. And then planning is, again, some kind of finding an efficient, sort of minimal cost sequence of actions to achieve your desires given your beliefs. This is a way to formalize a very classic idea in intuitive psychology, the sort of belief, desire, action rational agent model.

And from the context of Bayesian inference and probabilistic programs and common sense, we think of what the common-sense human being is doing is observing the actions of others, seeing them as the result of a rational planning program, the effects, and working backwards to the causes, the hidden mental states, the beliefs and desires in the other person's head. And we can use this to build, again, quantitatively predictive models of this aspect of common sense.

So just to illustrate the intuition behind this, I'll show you one example from Chris Baker's work in our lab. This is the food-truck domain here. So again, as at MIT, if you're on campus or maybe in other places where you've been, imagine a grad student going to forage for lunch one day, or maybe what you might have done during the break. There's various food trucks that come to campus each day.

In this little world, there's three kinds of trucks, Korean, Lebanese, and Mexican. But there's only two parking spots. So some days, Korean gets there and Mexican. Some days, the Korean and Lebanese.

Here we see a day where Harold, our grad student, has come out of his office here. And he can see, by the way, just in his line of sight. So he can see on this slide of his building. Let me get my cursor here.

He can see the Korean truck has come and parked in this spot. He can't see what's on the other side, though he knows there's a parking spot there because there always is. So what does he do? He goes this way.

He goes past the Korean truck to the other side where he can see the Lebanese truck. Then, having seen it, he turns around and goes back to the Korean truck. So the question that we might ask people here is, what is Harold's favorite kind of food? Is it Korean, Lebanese, or Mexican?

What do you think? Yeah, Mexican. Now isn't that interesting, because he didn't go towards Mexican?

Standard computer vision algorithms will see an agent or a person reaching for something, try to analyze the visual motion, say like, OK, I'm reaching here for my phone. So you think that's my goal. Here, what we perceive him doing is going towards his mental representation of something that he wants and hopes is there but isn't actually there. And in fact, both our model, which can make that inference, and people not only think that Harold, his favorite truck is Mexican but also that he believes it was likely to have been there when he started off, because otherwise it wouldn't have made any sense.

We can make this more quantitative. But I won't go into it. We can also-- and I think this brings back to several of the other talks, basically all the talks in this session. We can turn this into what I think is a model of people's mental models of these organized activities. And I think the most basic one, which breastfeeding, as you put it, maybe is sort of the first example of what you could call helping or caretaking.

Developmental psychologists in cognitive science have been very interested in these kinds of activities. And again, I can't go into the details. But if we want, we can talk about this in the question period.

But infants, from a very early age, seem to be sensitive to watching other agents, including little animated shapes like circles and triangles, which ones are good or bad based on whether they seem to be helping or hindering other shapes. And we've been able to formalize, what does it mean to see an action as helping in terms of this kind of planning framework, where basically it's to take the desires-- one agent is helping another if the first agent takes his desires to be some kind of function of the other person's desires.

[00:45:24] The way these get formalized-- I guess I didn't really quite mention-- is in terms of a computational extension of classical economics, if you like. In the same way that our physics engines extend on Newton, this extends on the early ideas of expected utility theory. We can, again, talk about-- if we want to come back in the discussion to neuroeconomics and whether expected utilities are a good way to think about the actual ways that our brains work. But they seem to be a good way to think about how our brains think about how brains work. And that we can formalize the idea of, say, a desire as a kind of utility function. And then a helpful desire is sort of recursively having your utility function depend on another agent's utility function, a kind of a golden rule, if you like.

These kinds of common-sense physics and psychology can also be studied in the brain. So again, I'm out of time. And I will stop.

But I'll just refer to the work of my colleague Nancy Kanwisher and a bunch of people in her lab. And they've found, basically, with simple kinds of intuitive physics and intuitive psychology tests testing, again, people's intuitive sense of whether there's a joint activity going on or not, you can find dissociations between large brain systems involving multiple areas of the brain that seem to underlie these aspects of core common-sense physics and psychology. You can also find dissociations in various kinds of developmental disorders. But again, I'll refer you to Nancy's work on that.

And I would also refer you to the work of a number of people in robotics, including Rod Brooks, who many of you might be familiar with at MIT, or Pieter Abbeel's group at Berkeley. I'll just show you this nice video here.

If you're not familiar with robotics, this might not look like much, seeing a robot tying a knot. But it's incredibly impressive. And partly, what it's based on is the robot doing one of these sort of physics-based planning things, using some of the same kinds of ideas inside the robot's head.

The very last thing I'll end with, and then we'll go to the discussion, is, of course, looking beyond common sense. So I've emphasized how we can try to use ideas of computation to reverse engineer common sense. But I think many of us are interested in what goes beyond that.

So for example, the origins of knowledge in theories which are not just intuitive theories but scientific theories. Like, we ask, how did Darwin come to his theory of natural selection by looking at finches in Galapagos? Or Mendel come to his theory of genetics from studying pea pods? Or Newton come to his theory of gravitation from looking at orbits of planets and dropping apples and so on?

Well, in some sense, it doesn't fit into the model that I talked about. But in other ways, it does build on it. And this is a frontier area that, again, we're just working on.

I think it's also interesting to come back to one of the themes from last night, to think about some of the challenges that face us on a societal and global scale, like for example, our society's inability to grapple with issues like climate change, or our whole globe's inabilities to do this. And to think about both ways in which understanding what these common-sense intuitive theories are like and understanding a computational framework that can both say, what is the core of common sense, the kind of theories that are of the physical world, for example, that are present even in young infant? And also understand how that's similar to and different from, say, more sophisticated scientific theory. That can help us understand both why, in some ways, people fail to grapple with what's really at stake in climate change, but also maybe give some hints on what we can do about that.

OK, so I'll just conclude. I've talked about what I called the common-sense core as this very early developing understanding of the world in terms of intuitive physics and intuitive psychology. And I've tried to at least sketch out how, in our research, we're reverse engineering this common-sense core and something about where it comes from. And I think this will be not only of interest for cognitive neuroscience but for many other people.

The technical idea behind this is something that I could just gesture at. But if you're interested in learning more about this, I would suggest reading up on the topic of probabilistic programs. And I just wanted to end with that Heider and Simmel video, or just bring the picture back there, because I think it's very exciting. I think we're actually starting to understand something of what's going on there.

But also to bring things back to what art does that science necessarily doesn't. I think even this video here, which was constructed by hand by two very artful psychologists, has a lot of art in it that we aren't able to handle. Even something as simple as this, I think our models are just scratching the surface of what we experience there. So again, something we could come back to in the discussion. Thank you.

[00:50:05] **Question & Answer**

Natasha Schüll: So I'll ask Alva to come back up. And Carrie Lambert-Beatty, whom I introduced earlier, an art historian coming to us from Harvard, will be moderating a discussion with these two panelists. And then we're going to slide, in about 25, 20 minutes, into a Q&A with the audience.

Carrie Lambert-Beatty:

So I thought that this panel was totally fascinating. And was going back to thinking about the structure of this incredible symposium as a whole and how it has moved from individual senses to thinking about sensing in general in the session today. And from perception, which was our focus in the panel yesterday and the one this morning, to questions of action, if I understand our sort of brief here. And therefore, what do we do with perception? And how do artists, scientists, engineers, historians help us get at some of those problems, both of how we do it and of what we should do, what kind of actions?

It also was a session that seemed to me to be perhaps problematically but interestingly related to an idea that I know Alva's very strongly associated with, which is this idea that we are not our brain. And that the limitations on thinking of consciousness as bounded by, embedded in, and sort of a product of the brain machine-- as incredibly useful as it is to find out which parts of the brain do what-- that we want never to lose sight of the way that the action of neurons, the action of cells with the perceptual system is-- nothing is really happening until those things are interacting with the world and with the body. And when we talk about consciousness, it's something both infinitely complex but also infinitely relational.

And this was something that I was thinking about a lot today, in part because of something that happened at one of the previous panels, the one about color, where a kind of classic question about color came up a couple of different ways. And that is, how do I know that what you say is-- that when I point to this and say, green, and you point at it and say, green, that we're actually seeing the same things. And like Tauba Auerbach, I remember this as like my first big, abstract, mind-blowing thought as a kid. And it's of course one that a few decades in college and graduate school taught me is just not a problem.

I've been taught that it's not a problem because we can think pragmatically. Green is what we all say it is. And let's get on with things.

Or we can say that that whole model, the sort of anxiety that drives that question of, how can I ever know what's going on in your brain? That is kind of a false problem. And I think that Alva's way of thinking about consciousness is one among other sort of philosophical models, like phenomenology and so on, that help us try to think otherwise.

[00:54:34] And yet that question and problem persists. It's maybe not quite possible to get out of. And I thought that part of what we were talking about here today, as we moved into thinking about consciousness in action, consciousness of others, and trying to produce organized behaviors, trying to improvise together, trying to create kind of new systems of experience, raised some of those issues for me again. And maybe it's something that we'd like to come back to. Does anyone want to sort of take off from there? I'm looking at you, Mr. brain is not self.

Alva Noë: Oh, gosh, you went for the hardest and most truly unresolvable question. It's so interesting that two things seem to be true. One is that we don't actually suffer from very powerful existential anxieties about whether, for example, our kids don't see the same colors that we see. At the same time, we find it almost impossible to talk ourselves out of that puzzlement. Very similar, in general, if you think that all that I know of you is what I see, what you say and do, the exterior features, how do I ever get the confidence that you are a locus of consciousness yourself or that you have a mind yourself? And yet--

Carrie Lambert-Beatty:

And yet.

Alva Noë: --it would be cause for clinical concern if I took that possibility seriously. And I think that shows that in some sense, the puzzle presupposes a theoretical framing of the issues, which isn't the right one. It isn't the real one. It isn't the live one.

Carrie Lambert-Beatty:

And it's a pretty historically specific one, culturally and historically specific, at least-- does every culture worry about whether we see the same green?

Alva Noë: Well, one thing I'm sure we can all agree, as you've already sort of indicated, is this particular one about color, it starts early.

Carrie Lambert-Beatty:

Yeah.

Alva Noë: I think I've heard children raise it spontaneously at a very, very young age as sort of an obvious possibility. And yet it's not a source for concern. There are real sources of concern about what's going on in the minds of others.

If you think of gaslight, or is somebody trying to manipulate you or control you? Does somebody love you? Is somebody actually trying to cheat you? But not, does red look to them the same way it looks to me?

That question doesn't arrest-- actually the person who's been most effective at labeling and describing this problem is a former MIT philosophy professor, Ned Block. He had a very good way of putting it, because it's very hard to dissociate the linguistic from the substantive in the language question. So for him, the way to formulate it is, how can we tell that the things we both call green look to me the same way the things we both call red look to you?

So where you have to sort of factor out-- take for granted that there's no behavioral dissimilarity between us. What's left? What residue is left for puzzlement? If you're waiting for my one-line solution to the problem, I don't have it. But you might.

Josh Tenenbaum: I don't have a one-line solution. But I think what we do offers some ways of thinking about some of these aspects. I agree with most of what you said.

Certainly, just the way you put it, like the problem of green and red and shared subjective experience. I think it's a huge puzzle. And I think it's not a worry. And it's interesting that it's a puzzle but not a worry.

I think anybody who starts to think about how minds work will be puzzled by it. And I do think it's probably culturally universal to think about how minds work, because even young infants think about how minds work. Different people in different cultures might think about that question differently. But we all think about that.

And I think this sort of common-sense core perspective I was talking about might have something to do with why we don't seriously question whether somebody else is an inanimate object or a zombie or whatever it is. From the very beginning, we really have a few basic ways to understand the world. We can understand the world in terms of inanimate objects, like objects subject to forces. And then we can understand the world in terms of some other kinds of objects, which have internal mental lives much like our own and exert forces on the world in order to achieve goals subject to beliefs.

[00:59:20] We have basically a very small number of models like that. And there are systems in the brain which are there in part because of evolution putting them there, and in part from our early experience, which is responsible for that. It's not that we can't understand other things in the world, like the complex causal feedback loops involved in global warming. But those take a lot of other resources.

They build on language. They involve a lot of conscious thinking. And they're hard won and often not very won.

So I think that perspective is consistent. And I also think that the reverse engineering perspective that I talked about-- maybe some of the parts of it that I didn't talk about. But I think parts that are shared between me and a lot of other people in brain and cognitive science are consistent with some version of the idea that we are not our brains, or you can't understand thought and consciousness by just looking inside. It's very much not biological reductionism.

Think about engineering in engineering. Like, you can't understand why your computer works the way it does, or why the software works the way it does, or why any engineered system works the way it does without understanding the social context in which it's designed to be used. And I would say, again, I don't think-- I hope this won't sound reductionist. But I would say evolution has designed our brains to be used in certain ways which only make sense in terms of these organized activities that our species participates in.

So it's a view that says, if you want to understand-- I don't think it's that different. I would say that some of my neuroscience colleagues often look-- like, they take a different perspective. Like that if you just focus on spikes or molecules inside synapses

that that's the heart of things. But I think, in that sense, I have a lot more sympathy with what you're talking about than what some of the same people in my own department do, at least as far as how it's going to bear on questions of intelligence. What they do might bear much more directly on how you might develop drugs to cure certain kinds of diseases. But I think these different ways that cognitive neurosciences look out to the world-- I think the kind of reverse engineering perspective that we have is very broadly consistent with the kind of perspective that you're talking about.

Carrie Lambert-Beatty:

It's interesting. During your talk, I was thinking a lot about anthropomorphism, which your model-- at least the way you demonstrate the core common-sense seems to imply it is actually kind of fundamental to how we look at things. Is that fair to say?

Josh Tenenbaum: Yeah, you mean a sense that we can look at circles and triangles and right sense?

Carrie Lambert-Beatty:

And we impute belief to them.

Josh Tenenbaum: Yeah, that's right, because if they move in ways that are consistent with forces being generated to maximize expected utility, basically, and those are all things we could formalize, or they move in the way a self-driving car seems to move, then, yes, it makes sense to anthropomorphize them.

Carrie Lambert-Beatty:

And anthropomorphism, of course, can be a very dangerous thing to do. But there's also people now who would argue that a kind of tactical use of anthropomorphism is necessary exactly if we're going to be able to rethink and become sensitive to Gaia, as Professor Latour was suggesting. In other words, yeah, these questions about climate and political and incredibly complicated questions have, maybe, at their root, how do we understand our relationship to things? And it sounded like you're saying that's actually one of our very core--

Josh Tenenbaum: Yeah. Yeah, I don't know--

Carrie Lambert-Beatty:

--brain powers.

Josh Tenenbaum: I don't think I understand the Gaia idea enough to know whether that's a valuable one or not for solving global problems. It might be. I think that the picture I put up there as a little slide from some people's climate change presentationss-- to me, from the little I know about the problem, the thing that we should most worry about are these feedback loops where global warming will lead to changes, say, in the tundra that will

release much, much more methane than any human-caused action does. So things that can just run away and get out of control.

And I think a lot of people have an intuitive theory that says, well, if humans are causing the problem, then, well, OK, when it gets too bad and we can't breathe, we'll just back off. And it'll be OK. Without realizing that you could cross a point where you can't back off, where even if you stop doing anything, the problem would just be out of your control. So that idea of a complex feedback dynamic is one that exists in some places in the physical world. But it's not part of our core intuitive physics.

Carrie Lambert-Beatty:

That's very interesting.

Josh Tenenbaum: It might be more part of our core intuitive psychology. Like, these kinds of bad dynamics are things we're more familiar with in relationships, including, sometimes, when breastfeeding is less pleasurable and more anxiety producing. But so it might be that anthropomorphizing in that sense would help us understand some of these complex feedback loops in a way that just thinking about the physical system as a purely physical system wouldn't. But it's also possible that we just need to help people develop better ways of understanding that causality in physics isn't quite as simple as just one ball hitting another. And coming up with ways to understand more complex but not kind of common-sense core kinds of physical causality would be useful, too.

Carrie Lambert-Beatty:

And does that maybe sound like a way of describing some of the things you're up to, projects like the one you showed today?

Tomás Saraceno: Yeah, uh, yes.

[1:04:47] No, I mean, I don't know how to put it. But very simple, what I talk about is butterfly effect. I'm always somehow plucked by these -- last year, I broke one of my legs. And I was walking all the time with crutches on the street. And then suddenly, you see in the street that everybody's walking with crutches. Or half of the population is with crutches.

And then I was very preoccupied because of this. And I say, oh, my god. Then you start to see, oh, the road is not well built. And it's a lot of things that you don't see. I never saw that there were other people walking.

And then somehow, I got better. And I don't see -- never again, I see people now with crutches.

Carrie Lambert-Beatty:

They're all gone.

Tomás Saraceno: And they're all gone. And I'm all the time busy. And I like what Bruno noticed. It's like, how we can become more sensitive? Or how I can stop forgetting something?

And only when it happened to me, then I start to see somebody else that is another problem. And then minorities, like weakness, they heal. And then it's, how I can activate this in my senses, not to kind of forget when my problem kind of disappear?

Alva Noë: It's very interesting that you can't help but see the figures in the videos as animate. We seem to be able not to see the globe as animate. And the philosopher Wittgenstein once said, we only attribute mind to that which has a face or looks and acts like us. But of course, what it is to look and act like us is really an abstract notion, because in what sense do the triangles look and act like us?

And kind of a question I wanted to raise-- and it's not a challenge. It's just simply an interesting question. The Bayesianism, the reverse engineering, the belief, desire, psychology, and inference, all of that suggests a kind of a theoretical attitude to the world, as if there's the world. It's this uninterpreted something. Let's figure it out. And then the idea is that each of us has a brain and a body, has this burden to figure it out for him or herself.

But the amazing thing is that that's not what we experience. We don't experience-- we don't live life as detectives. In some domains we do. My colleague Alison Gopnik wrote a book, *The Scientist in the Crib*, about these developmental accomplishments.

But what's so interesting is the non-theoretical quality of our relationship to each other. I look at you. I see your animacy. And I don't infer that you're animacy on the basis of what I see.

Josh Tenenbaum: Yeah, no, but I think words like "infer" can mean many different things. But I think that the view that-- so Alison Gopnik is very influential in my thinking here, also Susan Carey, who has been married to Ned Block for many years. And some of his ideas and her ideas, also very influential.

And I completely agree with that. And I think we, as a field, are starting to understand actually how to resolve that tension. It's that there are some kinds of common-sense theories, you could call them, which are, yeah, mostly unconscious. They're present in very young infants. And they're fundamentally different than sort of later-developing scientific theories or the kind of intuitive science that Alison is talking about.

They're automatic. They're not subject to the same kind of belief, revision, and sort of rational checking against all possible evidence. They're very modular, to cite, maybe, the opposite philosophical tradition. They're sort of like these Fodorian theories, Fodorian models. Jerry Fodor is the person we're referring to. So that seems to be, to me at least, a way to describe this very early developing common sense that we can't help but apply everywhere.

But then there's also later-developing intuitive theories, which are more explicit, more verbalizable, very much caught up with learning language. They only really come on when you're looking at three- and four-year-olds and kids. You could say one of the things that happens when you start to learn language is you re-describe, among other things. And part of learning about the world is you re-describe these core domains of intuitive physics and intuitive psychology in words.

You start to talk to yourself and explain to yourself, what makes something balance or be stable? Or why are people saying the things they're saying? And you see this really interesting phenomena that-- across these and other domains in physics and psychology, you see what developmental psychologists often call U-shaped curves, where kids suddenly sort of get worse at things that, when they were young infants, they were good at.

[1:09:26] So classic examples of this are in the so-called false belief task. Three-year-olds traditionally fail to understand ways in which somebody else could have a belief that's out of register with their understanding of reality. And it used to thought that that's just only something that's learned when you go from three to four. But then it was shown by Renee Baillargeon and others, about 10 years ago, that actually even young infants can do that. But it seems like, when you're telling--

Alva Noë: Depending on how you measure it.

Josh Tenenbaum: Depending on how you measure it, if you measure it nonverbally and sort of unconsciously and automatically. But if you tell a story in words and ask them for a response in words, then of course, a 12-month-old can't do that. A three-year-old can do it. They just get it wrong. And a four-year-old gets it right.

And you have the same thing with these kind of intuitive physics like stability. Is this stably supported by that? There's inferences that 12-month-old makes understanding aspects of center of mass, which a six-year-old might get it wrong. But then a seven-year-old might get it right.

So there's something about language and the way language both mediates our experience of the world and our cognition that is incredibly valuable and essential. We wouldn't have the culture we have without language. We wouldn't be able to have the interaction we're having now without language. It does enrich and extend our early developing intuitive theories. But maybe there's also sort of--

Carrie Lambert-Beatty:

A loss.

Josh Tenenbaum: -- an initial cost. But I don't know. I think it might be-- if we're going to talk about how to get beyond certain limitations or resist instinctual responses to things, I think we're going to have to engage with that part of our brains, the sort of language-based, more explicit part. At least that seems to be what cognitive science is telling us is that's the

part that's malleable that education acts on. The other part is no more really intervenable on than our sense of green and red.

Alva Noë: She's flashing you the 3-minute sign.

Carrie Lambert-Beatty:

OK, so does that mean we're about ready to open it up?

I think we might as well open up to the audience then at this point. So let's do it.

Natasha Schüll: Microphone over here.

Audience: In the interest of time, I can speak loud. When you spoke about core common sense, you did not mention any of the artistic urges we have. You alluded it might be too difficult. But do you think that is maybe for our beyond common-sense urges, the art playing that role? Did you do any research on that? Or are you working on it?

Josh Tenenbaum: Well, yeah, I think here, I would defer to the other panelists on that. But what I was referring to about that when I pointed to the Heider and Simmel thing at the end is I think what you said about the artistic sort of orientation that's-- it's not about, how do we get so much from so little? But how do we see so little in so much?

I would say that what you see in those Heider and Simmel videos with the triangles and circles, that is getting at the core of this automatic common sense. When we look at a much richer experience, like just the world around us, our everyday life, actual people interacting in a street or in our life, any of our real experiences, our brains are boiling it down to two triangles and a circle. That's the core of our common sense. And all the rest we're missing, like whether somebody's got crutches on or an infinite number of other things.

So I think maybe what your colleague might have been getting at, and maybe what you were getting at, is there's all these other aspects to experience which are not just what's shown in the triangles and circle. Because they're not part of our core common sense, because it's so important to us, to our most basic organized activities that we've pruned those out and focused on certain essences that actually we need other ways. And art can give us other ways to be sensitive to those.

I think that there's a key-- there's a nice difference between the Heider and Simmel, the triangles and circle, and the chasing. I think the chasing scene that I showed, we can understand that perfectly, basically. The models we have basically explain what's going on there. There's nothing very artistic about it.

And it's important. That was designed to study 13-month-old intuitions. The Heider and Simmel video with the triangles and the circle and this whole story, that was designed by social psychologists to study almost everything.

So even that scene, I think, has many aspects of thought and consciousness that we put onto it that I don't yet know how to understand, that looking forward, we'd like to understand. And somehow the artistic rendering of it was able to capture these aspects. And then there's all the rest of real-world experience, which doesn't fit into even the two triangles and the circle, which I think most of us are just not even seeing because they don't fit into this common-sense core.

Alva Noë: Do you mind if I just--

Josh Tenenbaum: Yeah.

Alva Noë: If I just add-- I don't want to contradict you in any way.

Josh Tenenbaum: Please do, though.

Alva Noë: No, no, because I think there's not a disagreement. It's maybe merely verbal. But I think it's worth maybe mentioning.

[1:14:42] And that is that I think there's a very important sense in which it would be a mistake to say that when we walk down the street, we're just experiencing circles and triangles. Even if it is true that what we pay attention to or what we notice-- or that what we see is a function of what we're paying attention to, what we're noticing, what we're concerned about, what our task is, it's very important that people show up as people, and cars show up as cars, and the sunshine shows up as the sunshine.

Josh Tenenbaum: Yeah.

Alva Noë: Even if, in some sense, it wouldn't be in an individual economic nationalistic model of how to predict the behavior. It should help everyone understand human experience. It's you are not just a moving dot on a two-dimensional screen for me. And

I think it's so interesting to point that we don't-- you see, if you think of what's represented in our heads, then you are left with saying it's triangles and circles. The point is we don't need to represent the world for the world to be there for us in our heads. We have access to it thanks the fact that it's there and that we have all sorts of skills when we needed the access.

Josh Tenenbaum: I'm not sure I understand all those words. But I think I would agree with most of them, maybe even all of them.

Alva Noë: Fair. That's a fair response.

Audience: Is this on? In Tomas's artwork, he basically created a social-psychological model, I think. It was like a social-psychological experiment in which people had to think as a group to survive. And I'm wondering if you could project common-sense theory onto-- if it's possible-- and I think we're trying here-- to project it onto a social-psychological lens.

And how might it relate to a larger population? Because you talk about intuition as something that can be generalized statistically. So we're talking about large numbers anyway. That make sense? Do you understand the question?

Josh Tenenbaum: I'm not sure. Maybe. Are you asking me to see if I can understand what Tomas was doing?

Audience: Why don't you talk to Tomas?

Josh Tenenbaum: Yeah.

Audience: Yeah, both of you, maybe.

Josh Tenenbaum: I love your stuff. I wish I was there in the hangar.

Audience: It looked like a social-psychological experiment.

Josh Tenenbaum: Yeah, is that how you thought of it?

Audience: No.

Josh Tenenbaum: No, I'm asking her. Or maybe to people who were-- so you were actually one of the people in there, right? Did you feel like you were working together with the other people to survive? Or do you feel like you were more in your own body?

Leila Kinney: Both. It wasn't a question of survival, no. But I do think there was some of this intuitive physics involved in it, because, again, it was that losing certain aspects of your sensory apparatus.

So you couldn't really communicate verbally. You couldn't really walk. You could see, depending on which layer you were in, better or worse, that kind of thing. So yes, you had to reassemble your ways of navigating a space.

Alva Noë: How come you couldn't communicate verbally? Was it too loud?

Leila Kinney: Yes, it was loud. There was that roar. I mean, you could have tried. But it wasn't-- suddenly you realize, OK, that's not the most relevant thing to do.

Josh Tenenbaum: It's also hard to establish the usual communication channels, probably. People are floating all around you in different ways. And you can barely see them sometimes.

Leila Kinney: Right, and it is moving as you're moving through it.

Tomás Saraceno: And also, your body and the weight of your body talk much more quickly than what you could express verbally, because it influenced your presence just as a weight, because the medium is so unstable. It go much quicker.

Josh Tenenbaum: Does it feel like you're communicating with the other people there in how you move? Like, do you feel a kind of height interaction, like a dance, for example?

Leila Kinney: I wouldn't say it's like a dance, no. But in order to accomplish something-- and there does become a sort of utility moment. In order to accomplish something, yes, there has to be a kind of duet. I suppose one could use that analogy.

Alva Noë: Maybe more like driving, less like dancing and more like driving in the sense that you're relying on another person's perceiving what you need and acting in such a way as not to run you over.

Leila Kinney: Well, yes.

Josh Tenenbaum: It is interesting to compare the experience, maybe, of being there with like the experience of being a red and a blue ball chasing each other around in its case. Because there, what made that activity make sense was these physical constraints that you couldn't pass through, the walls, basically. Whereas here, what you've done is you've put some physical constraints or circumstances to motion.

Carrie Lambert-Beatty:

Opportunities.

Josh Tenenbaum: Well, some of them are concerns. Like, there's good ones, like you're not falling down to the ground at least.

Leila Kinney: Right.

Josh Tenenbaum: You could see someone on one of the upper layers if you're in the middle layer. And you can't actually touch them or get to them. So there's some constraints.

[1:19:35] But they're very different from our normal experience in a lot of ways. You can see through these layers, unlike most obstacles. But you're always moving around.

So it seems like a really interesting way of sort of taking people out of the normal ways in which, say, our interactions with other people, our intuitive psychology depends on intuitive physics. That's a different intuitive physics. So to the extent that there's any kind of interaction with other agents going on there, it is sort of changing the rules of that game in some way.

Leila Kinney: Yes.

Josh Tenenbaum: That would be very interesting.

Leila Kinney: And it is playful, too. It is fun. So you do call on that experience.

Josh Tenenbaum: It looked awesome.

Natasha Schüll: OK, back over here, this woman in the blue shawl.

Audience: OK, right, I think I have a question for every single panelist. So there's one single question. And I want to bring a concept which I haven't heard from any one of you. And I wonder why that is and if it's useful in any way.

And it may overlap with the intuitive physics or perhaps the naive physics, as I have sometimes read. But also the kind of projections of somebody else's behavior and how you can anticipate and react to it. And this is a term which comes from art history but also dance spectatorship. And it's kinesthetic empathy.

And it looks at the way we kind of cross a number of realms in which you would say, as a human being, I anthropomorphize geometric objects. But also I would put myself in their position. And I would move around that kind of environment in the same way. And that would be a kind of a generative bond or a kind of anticipation of how these different agents would move about in space, and how would I react to them?

So kinesthetic empathy, I haven't heard that from anyone. And I wonder why and if it's any way helpful to you.

Leila Kinney: I can just say that I think Caroline and I, in talking about this session and what we were trying to accomplish, she was bringing into play the idea of joint action, so this reciprocity. And Josh can talk much more effectively about this than I. But the mirroring effect that we are sort of tuned to do when we do anticipate someone else's actions and then react to it, or mirror it, or slightly alter our own in response.

So there is-- well, I've heard Tomas use this phrase. I'm trying to get people to tune to one another. So he sometimes-- I think that's why you're moving towards sound and vibration. And so that's becoming a more important part of your work because he's understanding that perhaps in the way that musicians understand each other when they play in an ensemble or something-- I don't know. You can comment on that.

Alva Noë: I wasn't there. So I don't know what it's like to be there. But just giving you an observation from the outside watching the video, I didn't see a lot of kinesthetic empathy. What I kind of actually saw was individuals exploring their own sensorium.

That was the impression I had. I don't mean that that's the case. I'd be curious to know if that's not-- I mean, how much it really isn't the case.

But it struck me that everybody was sort of thrown back on themselves. And while they needed each other to solve certain problems, it looked like a slightly self-involved state that they were in.

Josh Tenenbaum: But I bet you could imagine, in a similar setup in which-- maybe with a smaller number of people, it would be more dance-like or more socially interactive.

Alva Noë: Have you ever slept on a waterbed before?

Josh Tenenbaum: Yeah, I was going to say waterbed.

Alva Noë: And every time they roll over, it wakes you up?

Josh Tenenabaum: Yeah.

Alva Noë: That's a kind of kinesthetic empathy, or at least the breakdown in it.

Josh Tenenbaum: Right. I was going to give the same example.

Natasha Schüll: So when we're talking about cultivating that kind of empathy but across time and looking toward the future-- and I'm bringing it back here to the angel of Gia's story. And what is she to do to grasp that future that is so sort of horrifying? And we've been talking a lot here about phenomenologically rooted exercises and experiences being in time, or seeing triangles and circles reacting to each other in time. But so I'm just struggling and wondering how to translate that insight, whether we reverse engineer it or sideways engineer it-- which is kind of how I think of Tomas's work. How to translate into that some sense of the climate change, the sort of temporal, complex layers of causality that it seems you're suggesting we don't have in our common core, right? So is this a question of transduction or translation?

Josh Tenenbaum: Again, I think our common-sense physics is basically like, build your ball of causation. And we understand some things about friction and mass, very simple Newtonian mechanics. And there's a lot of things in Newtonian mechanics that we don't understand.

[1:24:46] Our intuitive physics famously does a bad job with angular momentum and other concepts which really you only see when you study systems that are rotating at high velocity, which, again, are not ones that were very common in our evolutionary history or in the experience of young children. Certainly, I'd say, more our evolutionary history. So I think climate change-- we can barely understand sling shots is what I'm trying to say. Climate change is a much more complex sort of thing.

And when we try to think about-- we try to apply intuition to the physical, the inanimate world. To put it pretty bluntly, I think most people in first-world societies like ours are not Gaias. They think that the world is a physical system not an animate one. So they apply very simple mental models that are not that much more complicated than billiard ball causality.

One of the things that very young infants show-- one of the slides I showed from Renee Baillargeon's work. By five months, she showed, but not by two and a half months, infants start to understand the basic principle of conservation of momentum in collisions, that if a medium-size ball rolls down a hill and knocks an object a certain distance, then if you show them a much larger ball, they expect it might knock that object a farther distance, but not if you show them a smaller ball. So that's a very rough kind of understanding of transfer of momentum. But it's a very simple, linear one.

Bigger ball, heavier maybe, knocks it further. That is not the way climate change works. But if we're thinking-- if our core common-sense physics is that kind of a system, then it's no wonder we're not going to understand the feedback loops of things like, well, what happens if the Arctic tundra starts melting, and then you get much more methane than any humans are producing?

Natasha Schüll: And I see Tomas's work and Bruno Latour and the global Gaia circuit as trying to sensitize in some way to-- in a range of different ways.

Josh Tenenbaum: They have different possibilities for sensitizing. So one is think about it like an agent or a system of agents. That might be a version of Gaia. Another one-- like, I liked your non-Newtonian fluids example. Another one is, hey, the physical world is more complex than just the first few pages of a Newtonian mechanics textbook.

And maybe if art could give people a sense of the ways-- everybody agrees that cornstarch and water or whatever is inanimate. And yet, look, you amp up the Kenwood speaker to 60 Hertz and above, and it starts behaving in very non-intuitive and even somewhat animate ways. Well, that is one very simple way in which this kind of art could--

Alva Noë: But there's-- excuse me.

Carrie Lambert-Beatty:

I just wanted to-- because I think the way in which talking about climate change and ecological emergency is filtered through the whole symposium has been interesting and unexpected. But I think when we talk about it, we also have to remember-- I mean, I'm not sure that the problem is every individual's need to understand the science of climate change as much as it is concerted political efforts to debunk science so that what's missing is an understanding of--

Josh Tenenbaum: Well, it's a complex problem. It's a complex problem with a bunch of different aspects, I think.

Carrie Lambert-Beatty:

Yeah, of course.

Josh Tenenbaum: I guess I was just thinking of individuals I've met who've undergone kind of conversion experiences where when they understood what the causal thing was, they totally changed not only their own actions but their whole careers.

Alva Noë: I think it's important to separate the question of what it is to understand and take seriously the reality of climate change from somehow having an aesthetic intuitive insight into the world as a unit. Because one of the really interesting things about the world is it's not an object for us to contemplate. It's our world. We're inside it.

And I think while there is-- before, when we were talking about anthropomorphism, in a way that was the wrong idea, because I think the Gaia hypothesis, as I understand it, is not the idea that the world has a mind but that the world is alive. It's a living entity that's one system that sort of has the kinds of properties of self-organization that you find in living things. But to say that is not to say that we should, therefore, attribute intentions or desires and needs to it in a truly anthropomorphic sense.

[1:29:08] And I don't think we can take up the relationship to the world as the world is my baby, or the world is my friend, or the world is my mother. I don't think that's an intuitive way to think about it. But we can terrorize ourself about what's happening to our world.

And I live in California. People wash their hands differently in California than they do here. They turn the water on. They wet their hands. They turn the water off.

They lather their hands. They turn the water on. They rinse their hands.

They turn the water off. They don't go flip off-- it's really different. My kids have been raised with terror about the world, which is not an interpersonal thing. That's a metaphysical thing.

Natasha Schüll: Pick one. We have 10.

Audience: Hello, thank you very much for the very interesting speech. My question is open, I guess. I was wondering if we say that an embodiment experience, cognitive, like Tomas's artwork is somehow changing belief through an intuitive experience, these beliefs that Josh was talking about are altered. And so this models and reorganizes our awareness.

In a way, I was wondering how-- and you also mentioned, Josh, that as infants, these intuitive experiences very soon became language. And I was wondering, in terms of lab experimentation or phenomenology or artistic, what has been done in terms of trying to do the reverse, so try to experiment with language in that it is used in a different way so that it will create an empirical embodied experience that will reorganize and create this awareness? So what has been done in terms of even virtual mental images? And what do we know in terms of mental images, language, in terms of how it is applied in all these fields of remodeling and reorganizing our awareness? Did I make sense at all?

Josh Tenenbaum: You made a few different senses. I'm not sure which one I have anything to say about. But I guess, again, I'll try to say very briefly, I'm not a language researcher primarily. But from what I've learned from colleagues who do work on language, yeah, these same kinds of common-sense core theories that are present in early infancy about intuitive physics and intuitive psychology are at the heart of early language.

The first words we learn and the syntax of even our first sentences that we construct, they're building on these kinds of concepts. So language builds very much on this. At

the same time then, it completely transforms the way we're able to think. We can construct models in words, whether we're Isaac Newton or any one of us here, that just don't fit at all into our basic intuitive theory.

So am I getting at that at all? I don't know.

Audience: You are. But I would like to know not just in terms of language, but also language as virtual mental image. So not just in its interaction and its own manifestation. What has been done in terms of experimentation to create a cognitive transformation through language, either speech or mental? What has been done? Or what is being done in these terms, if you know of anything?

Josh Tenenbaum: Probably, in this panel, there's not enough time to go into it. But there's a huge number of interesting experiments out there. People looking at, say, for example, eye tracking, what you're looking at as you're processing language. There's people who've worked on how language transforms what you see.

We talked about color. So going back to that theme, one of my former colleagues here, Lera Boroditsky, was well-known for studying how the experience of learning a language like Russian, where you have two different words for two different shades of blue that, in English, we just call blue, roughly sort of a dark blue and a light blue-- how does that change your percept in some fundamental way? I guess the interface between language and perception is a very exciting area. And it does seem like in some cases it might. But it's pretty hard.

[1:34:19] If there's one lesson you draw from that, it's that, again, like perception and a lot of the core common-sense stuff we're talking about here that's present early, it resists. Language can give you other ways to think. But it doesn't redraw the basic cognitive contours of that infant mind.

Audience: But language-- sorry. Just language in that it--

Natasha Schüll: I think we need to ask one more before we--

Audience: --in that it creates experience.

Josh Tenenbaum: I'm happy to talk more about this after. But thank you.

Alva Noë: But it's just worth very quickly mentioning that although-- Tomas's work affords us an opportunity to talk at great length about it. And that's really interesting that there's so much to be said about it even if it's not a primarily linguistic vehicle. And moreover, it does so much to us which is not linguistic, which we can talk about.

I was thinking there was a show in London a couple of years ago, which I wonder if you saw. Actually, it was in Germany, as well. It was called Move-- Choreographing You. And the idea was works of art that make you-- they choreograph you. So there was a piece in it by Forsythe that consisted of rings hanging.

And what you were supposed to do is climb through the piece. So you put one foot in one ring, hold onto another ring, raise your foot up onto that ring. It's like a kids' jungle gym. But you ended up being choreographed in this way that you could look-- you could just stand there in the room and watch people moving in ways that was absolutely dictated, as it were, by the affordances of the objects, but completely unpredictable in terms of normal human movement.

And I had that experience also watching you. This strange-- if you sort of abstracted away from the fabric that was holding them, you found people in these alternative poses that were striking.

Carrie Lambert-Beatty:

I mean, a chair choreographs us.

Alva Noë: Yes, exactly.

Carrie Lambert-Beatty:

And so it's choreographing something else, something that we don't ordinarily do that work like this.

Natasha Schüll: OK, for the last question.

Audience: This is a question, I guess, directed towards Alva. But I guess we can open it up. I was interested in aesthetics.

We've talked a little bit about aesthetics. And one of the notions of aesthetics that came from Professor Latour's talk was the idea of rendering us sensitive to dot, dot, dot. But it seems to me-- correct me if I'm wrong-- that the notion of aesthetics that you're talking about, Alva, is actually not exactly about rendering us sensitive to, in the sense that it's not about, say, the revelation of phenomena, the exposure of phenomena. But it's also about the way in which aesthetics helps us to fashion ourselves as subjects in our interaction with the world.

So when you're talking about the painting and fashion, there's a kind of important dialectic that's happening here. We become subjects insofar as we are engaged with the work of art. But it's not necessarily an active, say, rendering sensitive to. So I was wondering if you could comment, elaborate a little bit more about the kind of aesthetic theory you're talking about.

Alva Noë: That's very helpful, actually, because-- is Bruno Latour still here? I don't know. But he made a comment. He said, I'm using the word "aesthetic" in its original meaning. Essentially, he meant it as a perceptual faculty.

And we can cultivate our perceptual faculties. If you read Arabic, you can look at a page of Arabic writing and see the individual words. You can cultivate that sensitivity. We're

all experts at reading each other's facial expressions. And if we're all English speakers, we're capable of understanding what we're saying. So we cultivate, in that sense, our aesthetic sensitivities all the time.

And I don't think that has anything especially to do or particularly to do with art. Art is something to which we are blind, however, if we don't cultivate aesthetic sensitivities to what it's playing with, what it's doing with. An example of an artist that I think beautifully illustrates this is-- oh, no, all of a sudden I've forgotten his name. He's a German choreographer who's not really a choreographer.

Audience: Tino.

Alva Noë: Tino Sehgal. Oh, in fact, we had this conversation about Tino. So when you saw Tino Sehgal's work at the Biennale in Venice last year, you didn't see it. You walked right through it.

It was people on the floor. They were invisible. They were invisible until they weren't.

So one of the beautiful things about that work and about performance in general is very often performance is invisible except that it gives you resources. It's like an online, quick, real-time tutorial to see it. So all of a sudden, it comes into view.

[1:39:01] So,you would have people on the cellphone standing in the middle of the dancers on the floor. And then they'd notice that all these people are looking at them. And all of a sudden, they'd realize they're in the piece. And then all of a sudden, there's a piece.

And so there is a way in which art is concerned with the cultivation of sensitivity of a perceptual kind in something like the way Bruno Latour said. But then there's this further question of, in what does the achievement of a sensitivity to the meaning or significance or importance of the work consist? And in a way, that's what I was trying to get at, this idea that works of art do this other thing, which doesn't have to do with just seeing them. It has to do with defamiliarizing, destabilizing. So if you think of Tomas's work, it brings so much into the foreground which is otherwise in the background, our frailty, our vulnerability, our dependence on gravity. And in doing that, it gives us the opportunity to be different.

Natasha Schüll: And our responsibility, right?

Alva Noë: And our responsibility, yeah.