

The following content is provided under a Creative Commons license. Your support will help MIT OpenCourseWare continue to offer high-quality educational resources for free. To make a donation or view additional materials from hundreds of MIT courses, visit MIT OpenCourseWare at [ocw.mit.edu](http://ocw.mit.edu).

**NANCY**

**KANWISHER:**

I'll just be brief today, but you can check out some of my stuff at the website up there. If you're confused by my appearance, if you've met me before, yes, I used to look like that. But at a deeper level, really, I look like this. This is me, and you look like that, too, inside. And these are parts of-- is that showing? Yeah. These are parts of my brain that we've mapped with functional MRI that were either discovered in my lab, or that my colleagues discovered, and that we then ran-- they're kinds of scans in my lab. These are all regions that do very specific things and that to me, are a big part of the story of how we are so smart.

So my interests, at a very general level, are to answer things like what is the architecture of the human mind? What are its fundamental components? And there are lots and lots of ways to find those fundamental components. Functional MRI, which is how we made this picture, is just one of a huge number. I loved Patrick's comment that you should find questions, not hammers. I kind of like my hammer, I have to confess. But questions are more important. And there are lots of ways to approach this question of the basic architecture of the human mind.

I also want to know how this structure, which is present in every normal person-- I could pop any of you in the scanner and make a picture like this of your brain, OK, it would take a little while, but wouldn't take that long. How does that structure arise over development? How does your genetic code and your experience work together to wire that up when you're an infant and a child? How did it evolve over human evolution?

This is sort of what's sometimes called a mesoscale, this really macroscopic picture of the major components of the human mind and brain. But of course, we also want to know how each of those bits work. What are the representations that live in each of those regions? And how are they computed? And what are the neural circuits that implement those computations?

And of course, cognition doesn't happen in just one little machine in there. It's a product of all of these bits working together. We want to understand how all of that works, too, and how all of that goes together to make us so smart. And that's related to a question that I'm deeply

interested in, which is what is so special about this machine that looks a lot like a rodent brain?

And it's smaller than a whale brain or a Neanderthal brain, so it's not just that we have more of it. What is so special about this thing that has put all of us here, interacting with each other and studying this thing, something that no other brain is doing, no other species brain? So are there special bits? Do those bits work differently? Are there special kinds of neurons? I don't think so, some people do. What is it about this that has brought us all right here?

OK, so that, at a top level, are some of the questions I would most like to answer. Not that I know how to approach any of them, but I think it's important to keep an eye on those goals, even when you don't quite see how you're going to get there. My particular focus in the CBMM Project is to look at social intelligence, which is one piece of that puzzle. And so, why social intelligence?

Well, just briefly, I think social cognition is in many ways the crux of human intelligence. OK, and it's a crux in a whole bunch of different senses. One is it's just the source of how we're so smart. Like, if you think about all the stuff you know, OK, do a quick mental inventory. OK, what's all the stuff you know? Like, make a little taxonomy. There's this kind of stuff, it's all lots of different kinds of stuff you know.

OK, now how much of that stuff that you know would you know if you had never interacted with another person? A lot of it, you wouldn't know, right? So a lot of the stuff we know and a lot of the ways that we're smart are things that we get from interacting with other people. That's social cognition.

OK. Another sense in which social cognition is the crux of human intelligence is many people think that the primary driver of the evolution of the human brain has been the requirement to interact with other people who are, after all, very complex entities, and to be able to understand how to work with them, and what they're doing, and what they'll do next is very cognitively demanding. And so that may be one of the major forces that has driven the evolution of our brain.

Another sense in which social intelligence is the crux of human intelligence is that it's just plain a large percent of human cognition. OK, so we do versions of social cognition much of every day. Right now, I'm having these thoughts in my head. God knows what that looks like neurally. I'm translating that into some noises that are coming out of my mouth, you're hearing those noises, and you're getting-- let's hope-- kind of similar thoughts in your head.

That is a miracle. Nobody has the foggiest idea how that works at a neural level. Nobody can even make up a sketch of a hypothesis of a bunch of neural circuits that might be able to make that happen. Right? That's a fascinating puzzle, and it's also of the essence in human intelligence. And we do it all the time, not just speaking per se, but all the other ways that we share information with each other. So one, social cognition is just what we do all day long every day.

It's also a big part of the surface area of the cortex. So this cartoon here shows-- with some major poetic license-- brain regions that are involved in different aspects of social cognition. And it's just a big part of the cortical area as well. OK. Another sense in which social cognition is of the essence in human intelligence is that many of the greatest things that humanity has accomplished are products of people working together. So all of that is the big picture on why social cognition is cool, and important, and fundamental.

The part of it that we're focusing on in our thrust within this NSF grant is something I call social perception. OK, so by social perception, I mean this spectacularly impressive human ability to extract rich, multidimensional social information from a brief glimpse of a scene. From a brief glimpse at a person, you can tell not just who that person is, you can tell what they're trying to do. You can tell how they feel. You can tell what they're paying attention to. You can tell what they know and who they like. OK?

And that's just the beginning. OK? So the work in our thrust tries to approach all of these different kinds of questions that we are calling as part of our PR of this NSF grant. It's kind of an organizing principle. The Turing questions, these demanding, difficult computational problems of social perception. Who is that person? What are they paying attention to? What are they feeling? What are they like? Are they interacting with somebody? What is the nature of that interaction? And so on. OK?

So the general plan of action in how to approach this in our thrust is first to study these abilities in the computational system that's best at them, namely this one-- and those out there, yours, too-- the human brain. And so the roadmap here is to first do psychophysics, characterize simple behavioral measurements-- what can people do, what can't they do-- from simple stimuli, and quantify that in detail.

Ask, how good are we at it? Maybe some of these things that we think we can do, like size up somebody's personality in three seconds when we first meet them-- feels like you can do that,

or at least you get a read on them-- I mean, is that based on anything? Is that just garbage? Right? Are we actually tapping into real information there? What cues are we using when we make those high level social inferences? What is the input that we get, that we use as a basis for analyzing this particular percept or throughout life that we've used to train up our brains to be able to do this?

So the second approach is once we have some kind of sense of what are those abilities-- that's sometimes called Marr theory level, characterizing what can we do, right-- is we can then try to computationally model this. And so there's lots of different ways to do this, and many of the other thrusts that you'll hear about are really tackling that problem. Another thing we can do is, of course, characterize the brain basis of these abilities, and we can do that with all kinds of methods.

We're using, in our thrust, functional MRI, intracranial recordings, something called NIRS. This is the ability to make measurements of blood flow changes in very young infants. And so we can characterize these brain systems in adults and infants. And that gives you a leg up in understanding these other broader questions about how the whole system works in a number of different ways. Just seeing how the brain carves up the problem of social perception into pieces already gives you some clues about the kinds of computations that may go on in each of those pieces. OK?

OK. So that's the overview. There's many, many ways you can do this, and of course, people all over the place are doing this. There's nothing all that unique about it. This is just our framework here. Some of the specific projects that are going on include some work on face recognition, which of course, a really classic question that many people have been approaching.

My post-doc, Matt Peterson, here has done some very lovely work where he's shown that, actually, where you look on a face is very systematic. You don't just look anywhere, right? When you first make us saccade into a face, somebody appears in your visual periphery, right, of course, all the high-resolution visual abilities are all right near the center of gaze around the fovea, where you have a high density of photo receptors and a shitload of cortex-- to be technical about it-- devoted to allocating center of gaze.

Right back here, in primary visual cortex and with the first few retinotopic regions, you have 20 square centimeters-- that's like that-- of cortex allocated to just the central two degrees of

vision. Right? So you have a lot of computational machinery doing just that bit right there. Well, when a face appears in your periphery, you move that bit of your cortex, boom, right on top of it. So you have all that computational machinery to dig in on the face, right?

OK, so what Matt has shown is that the particular way that you allocate that computational machinery, namely by making an eye movement to put that stimulus right on your fovea, people do that slightly differently. Some people fixate on a face up here, some people fixate on a face down there, and most people fixate someplace in the middle. OK?

Well, so why is it interesting? Here's why it's interesting. People do that in very systematic ways. And if you look up here, you pretty much always look up there. And if you look down there, you pretty much always look down there. And this has computational consequences. If we brought you guys into the lab and ran you on an eye tracker for 15 minutes, we'd find out which of you look up there and which of you look down there.

And if we took those of you who look up here, and we presented a face by flashing it briefly while you're fixating so that the face landed in your not-preferred looking position, your accuracy at recognizing that face would be much lower, and vice versa. If you're one of the people who looks down there, and we flash up a face so that it lands right there on your retina, you're much worse at recognizing it.

And what that means is that this fundamental problem that you'll hear about in the course, that Tommy has worked at in many people, it's one of the central problems in vision research of how we deal with the many different ways an object-- the many different kinds of images an object can make on our retina by where it lands on the retina, how close it is to you, the orientation, the lighting, all these things that create this central problem in vision of the variable ways an object can look.

A big part of how we solve that for face recognition is we just move our eyes to the same place. Position and variance problem solved, mostly. OK, it's kind of a low-tech solution. It's a good one. OK, anyway, so Matt has been working on that for a while, and so now, most of that is lab studies. Now what he's done is he's using mobile eye trackers, which look like this, and a GoPro attached to his head, because the mobile eye trackers don't have very good image resolution.

And so he's sending people around in the world, and he's finding that, first of all, yes, in fact, when you're walking around in the world-- not just when you're on a bike bar in a lab, you

know, with a tracker and a screen-- the people who look up here also look up there in the world, right? So that's just a reality check that shows that our technology is working. And now Matt is using this to ask all kinds of questions. For example, social interactions, where do people look in social interactions? Can you tell stuff about what they think about each other based on where they look on faces, right?

We want to run-- this is fruity. We haven't set it up yet, but we want to run speed dating experiments in the lab with people wearing eye trackers. I bet in the first few fixation positions, you can tell who's going to want to recontact who. I don't know. We haven't done that yet. OK, that's a little trashy, but it's kind of interesting. Some interesting scientific questions are a little bit trashy, you know. Some trashy questions are not scientifically interesting. I think that's one of those rare that's actually both. Anyway.

We also want to characterize-- a whole other part of this is this question that people have been considering for a few decades now of natural image statistics, right? So people have done all this stuff, collecting images, and at first, they did it really low-tech, and then the web appeared. And it's like, oh, now there's a lot of images out there, and we can just collect them easily. And let's characterize them. What are natural images like? So it's a whole set of math where people have looked at those natural images, and characterized them, and tried to ask how the statistical properties of natural images have-- how we have adjusted our visual systems to deal with the images that we confront. And that's a cool and important area of research.

But in all of that work, nobody's actually used real natural images, right? The images on the web, somebody stuck a camera and put it there, and then they threw away most of the pictures they took. The ones that land on the web are the ones that have good resolution, where people weren't moving in and out of frame, things weren't occluded. They're not at all like the actual images that land on your retina.

So we're collecting the actual images that land on your retina. And we're doing it with mobile eye trackers, sending people around in the world using these nice GoPro systems to give us high resolution. And importantly, not only are we collecting real natural image statistics from these real natural images, we know, for each frame, where the person was looking. And that's important for the reason I mentioned a while ago, that most of your high-resolution information is right at the center of gaze. And the information out in the periphery is pretty lousy. OK, so that's one project that I described too long, so I'll whip through the others more briefly.

We want to know how well people can read each other's direction of attention. OK, so when I'm lecturing now, if you guys get bored and look at the clock, I will see it, right? And that's just one of these things, you know? We're very attuned to where each other are looking, and that's very useful information.

You meet somebody at a conference, and you see them make a saccade down to your name tag, and it's like, damn it, doesn't this person remember who I am? You know? I'm very aware of this because I'm mildly prosopagnosic. So if I've met you before, and I'm slow to register, don't take it personally. I'm just lousy. It takes me a long time to encode a face. Anyway, we're very attuned at where each other are looking.

And so there's been a lot of work on how precisely we can tell whether somebody is looking right at you versus off to the side. Try this at lunch. When you're in the middle of a conversation with somebody, fixate on just the side of their face, not way off to the side, just like here, and just do that for a few seconds. It's deeply weird. The person you're talking to will detect it immediately, will feel uncomfortable, until they realize what you're doing, and then you guys will have a good laugh. And that will show you how exquisitely precise your ability to read another person's gaze is. It's really very precisely tuned.

OK. So there's a lot of work on that, but there's less work on how well I can tell what exactly you're looking at if it's not me. That is, I can tell if you're looking at me or off to the side, or this side, or that side. But what we're looking at is how well can I tell what object you're looking at? And that's an important question because many people have pointed out that a central little microcosm, kind of a unit of social interaction, is something called joint attention. And joint attention is when you're looking at this thing, and I'm looking at it, and I know you're looking at it, and you know I'm looking at it. That's a cosmic little thing. Like, we can have this little moment, right? Joint attention, OK?

And people have argued that that's of the essence in children learning language. It's of the essence in all kinds of social interactions. And by most accounts, no other species has it, not even chimps. OK? I mean, there's still some debate about this, and people niggle and stuff, but basically, they don't have it in anything like the way we have it. So we want to know, what is the acuity of joint attention? OK, so I was supposed to do that briefly. I can't seem to be brief. OK. So that's a whole project that's going on with Danny Harari and Tao Gao.

We're also asking how well people can predict the target of another person's action, right? So

if I go out to reach this, at one point-- well, there's only one thing there-- but if we had a whole array of things, at one point when I'm reaching for an object, can you extrapolate my trajectory, look at my eye gaze, and use all of those cues to figure out what is the goal of my action?

Here's a cool way to look at how well people can predict each other's actions. This is work by Maryam Vaziri-Pashkam, shown here, who's a post-doc at Harvard working with Ken Nakayama, who will give a lecture later in the course. And what they're trying to do is get an online read of how well people can predict each other's actions. And so obviously, this happens in all kinds of situations, especially in sports, right? If you're playing basketball or ultimate frisbee, it's all about predicting who's going to go where when and trying to take that into account with your actions.

So they've set this up in the lab. And they have a piece of glass here, and there's two Post-its on this piece of glass. And one person's task is to reach out and touch one of those targets quickly. And the other person who's the goalie watches them through the glass and tries to touch that target as soon as possible after the first one does. OK? And so it's just a basic little game. And so they have little sensors on each person's finger so they can track the exact trajectories and get reaction times. They're just behavioral measurements, but they're very cool.

So what they find first of all is that the goalie, the person who's trying to reach to respond to the other person, can do that extremely fast, right? They launch their hand to the correct target within 150 milliseconds. Well, you should immediately realize that something's fishy. You can't do that. It takes about 100 milliseconds just to get to V1. It takes, I forget how long, but a few tens of milliseconds to send the signal out from your brain out your arm to initiate the movement. So how could you possibly do all of that in that time?

Well, you can't. And what that means is that people are actually launching the hand action, the goalie's launching the action before the other person has actually started moving their finger. They've started processing it before. And the way they've done that is before this person starts, before their hand moves at all, they've subtly changed their body configuration in ways that the other person can read. OK? Now, on the one hand, OK, duh. You're playing this game. You learn to exploit cues. We're really great at figuring out cues quickly, and using them, and learning to use them.

But here's the-- one second-- here's the cool thing about this task is that this immediate, ultrafast reaction time happens on the very first few trials. So the ability that this task is tapping into is not that the goalie can learn what cues are predictive given enough trials and feedback. No, they do it right off the bat. This task is tapping into an ability that we all have already, right now, to read each other's actions and predict each other's behavior. And so people with no instruction and no experience whatsoever in this novel task know that this subtle little cue of the way the body is moving a little bit before the person's finger even starts to move, they can tell what it's predictive of.

So that's just another way to characterize people's abilities in social perceptions, so one of some of the many different things that we just see really well in other people's actions. OK, that's what I just said, all right. All right. I'm going to skip over some stuff.

We're looking at perception of emotional expressions. Almost the entire literature is based on staged emotional expressions on faces, huge literature with neuroimaging and behavior, and it goes back forever. But my colleague Elinor McKone has pointed out that actually, it would be important to look at real emotional expressions on faces. Maybe that's different behaviorally. It turns out it's very different behaviorally. One, you can tell if somebody's faking an emotional expression or if it's a real one. Like, OK, which of these is real fear, and which of these is staged fear? Duh!

OK, so one, we're really attuned to that. I think that's really interesting. Just as a social perceptual ability, we spend a lot of time trying to figure out who's sincere, who's genuine, who's faking something, what's for real, right? You know? There's all kinds of shades of that. And here's one little piece of it, right? So I think that's very interesting. And they've shown that behaviorally, these phenomenon are very different.

Just one example. A prior literature had shown that people with schizophrenia are particularly bad at reading facial expressions, using the standard measures, a standard stimuli, the Ekman six facial expressions. These guys replicated that finding and then showed that when you run the same experiment, but using not staged but real emotional expressions, schizophrenics are better than everyone else. OK, so it matters behaviorally, and it's interesting. OK.

All right. Other things that we're doing-- right. Leyla, your TA here, who's done beautiful work on her thesis work with Tommy using MEG and other methods, is now working with me and Gabriel, using some of this magnificent data that Gabriel has collected over a bunch of years,

where he's got intracranial recordings from human brains while people watch movies.

This is so precious. These data are like a dream to me, as somebody who's been using functional MRI as my main hammer for the last 15 years. Functional MRI is magnificent, it's wonderful, it's fun, but it has fundamental limits. One, it has no time information worth a damn. And the computations that make up perception, including social perception, and language processing, and most of the interesting aspects of cognition, happen on the order of tens of milliseconds. We can't see any of that. It's all just squashed together like a pancake, right, with functional MRI. With intracranial recordings, you have exquisite time information, and you can see computations unfold over time. That's very precious.

Second of all, in principle, with intracranial electrodes, you can test causality, something you can't do with functional MRI. You can stimulate and ask what tasks are disrupted. All right? So there's a huge number of cool things you can do with intracranial recordings. Leyla is looking at some of the data that Gabriel has been collecting, with intracranial recordings of people watching movies. And because these are rich, complex social stimuli, she's going to look at all kinds of things that we can try to extract from those data.

Like, can you tell the identity of the person who's on the screen right now? Can you tell from their face, their voice, their body? Can you tell what action they're carrying out? Can you tell if the person on the screen right now is a good guy or a bad guy? Right? Can you tell what kind of social interactions are going on? So we know all of this stuff, all this information is extracted in the brain, because people are good at it. But to get a handle on the actual neural basis of how we carry out those perceptual processes, this will be a really cool tool. So that project is just starting now.

And in other projects going on, Lindsey Powell, shown here, who's working with Rebecca Saxe, and Liz Spelke, and others, is using this NIRS method to look at blood flow changes in response to neural activity in infant brains. She's looking at some of those specializations that I showed you in my brain at the beginning and asking, which of those are present in infancy, a totally cool question.

And Ben Deen, and Rebecca Saxe, and me, and a bunch of others are looking at a big chunk of the human brain that was one of my colored patches before. This whole dark gray region here is called the superior temporal sulcus. This is an inflated picture of the brain. That means-- usually, the cortexes are all folded up inside the head. You have to do that to fit it in

there. But if you want to see the whole thing, you can mathematically inflate it. So that's what's happened here. And the dark bits are the bits that were inside of folds before it was inflated. So they're inside a sulcus, but now shown blown out to the surface.

So this superior temporal sulcus running down here is one of the longest sulci in the human brain and one of the coolest. And an awful lot of social perception goes on right there. Ben Deen has a paper in press and some ongoing work where he shows that lots of different kinds of social, cognitive, and perceptual abilities actually inhabit somewhat distinct regions along the superior temporal sulcus. They're not perfectly discrete. Nothing is a neat little oval in the brain. Actually, they somewhat overlap, but there's a lot of organization in there. And that's cool because it gives us a lever to try to understand this whole big space of cognition.