Key:

M: Dr. Montague

M: The study of the structure and conductivity the human brain has been going on now for almost 50 years. In recent years, a new method called functional MRI has been used and instead of taking a snapshot of the brain or like a snapshot of your stomach or your kidney or your knee where you get a static picture, functional MRI takes movies of microscopic blood flow changes in your brain. Microscopic blood flow changes in your brain are a good proxy stand in as it were, for changes in neural activity. So if I can make a bunch of microscopic blood flow movies with fMRI, then I have a way to read out increases and decreases in neural activity. And people have used this method to really revolutionize our understanding of at least one level of cognitive science and the cognitive traits that follow us all around and are underwritten by our brain activity and it's using functional MRI. Another way that functional MRI is used is in what's called connectivity analysis. And in this, people study all the areas of the brain they're active at any one moment and study that pattern of connectivity, activity and D activity the way they're correlated over time from moment to moment. And they look at which regions are talking to one another and different cognitive demands that are on board. And that becomes a new way of looking at disease states, perceptual states and in our particular case mental states associated with the legal code. One thing everyone wants to have is a kind of brain Rosetta Stone, a sort of pattern of activity or a spot or something like this it says, “You're capable of making this decision, you're incapable of making that decision, you're feeling guilt, you're feeling that way, you're not feeling this way.” And frankly, right now, the way brain imaging is done, we cannot do any of that. There's no bromide where you can say, “This person had a big chunk of their cerebral cortex cut out by a bullet that hit them and so they're unable to do X, Y, and Z.” Now we're starting to catalogue the way in which you can use brain activity to predict things that you already know about people. That is, you say, well, here's a category of people who are doing a visual perceptual experiment and they're seeing an ambiguity between one perception and another perception. And you can take brain activity across the entire brain and you can decode it in a way. And there's a movement afoot using those sorts of decoding techniques. Decoding is something that we have used in the MacArthur neuroscience in Law Project to look at what's called the Model Penal Code distinctions in about 1960 this is ensconced in basically English common law and of course the American system from that. All elements being equal, you can be differentially punished for a particular crime, depending on which mental state you're in, purposeful, knowing, reckless, or negligent. These mental states are noted in courts of law, they're thought over and they're decided upon. So in one case, you might get probation, in another case you might go to jail for quite a while, all the elements of the crime being equal, except for

this mental state. So, one of the questions we were after is, are these real? Can you decode brain states that look to be analogues to these mental states from the Model Penal Code? What we did was we set up a simulated scenario that was acceptable to the legal scholars on the committee, where the knowing and reckless boundary could be well defined. Did a series of behavioral experiments in people and then did the same behavioral experiments in a functional magnetic resonance imaging machine and ask the question, can we decode their brain activity such that we could distinguish the knowing state from the reckless state, under that scenario? Whether or not that scenario is representative of the real world, that's something that for people to debate. In any case, this was a project led by a PostDoc of mine, Iris Vilares, and she was quite able to decode knowing and reckless in that simulated scenario. Now lots of open questions here, one is, is that scenario representative of things in the real world? Would people accept that? The second is, what does this say about sort of proximate things going on in courts of law? And the answer to that latter question is probably nothing. In other words, you're not going to come into my laboratory and I'm having done something three years ago and I'm going to give you an experiment decode your brain activity and decide what your mental state was on August 27 2014. That's not going to happen. Maybe in some futuristic scenario things like that could happen and we'll worry about that then but for now, you can't do that. I think the way in which this will inform the law is people discussing the fact that there may be something like physical substrates the Model Penal Code states. And by learning that we may think about how we can mitigate the way that these things are used in courts of law, sort of outside the courtroom. So, I don't see it having any proximate use for day to day activities in a court of law. The best neural predictors of recidivism depend on so many factors that one would be shocked if you couldn't find neural correlates of somebody's propensity to commit a crime again or in some legal sense, backslide. However, they don't appear to be nor do they appear to be on the horizon any bromides for this. There's no single area that you can look at and there's no single response type that you can look at. These questions about the neural basis of legal distinctions or the neural basis of culpability aren't just science questions. There’re also questions about how the question is framed. Of course, your brain makes you do everything you do. The way you frame the question though may already trap somebody into a certain status such that you, of course, you can decode the brain in that situation but you're not really answering the question. In fact, you may be building in your own prejudices about what to expect to find. So, these are subtle questions. And as a lifelong brain scientist, I know that they're subtle questions they're hard to think about. And they're not just questions for the scientist to think about nor are they just questions for the lawyers to think about, the legal scholars, it’s a blend and it's a bit of a frontier. I think care and seriousness have to be taken. So, you can imagine glib, one size fits all solutions, where you say this area is completely responsible for this person's emotional status. It may be true that injuring that area of

the brain makes it hard to control your emotions or having a frontal injury makes you frontal. There are classic symptoms of hypofrontality due to injury and disease. That doesn't mean that we have right now signatures to say the degree to which that's true and the way in which we might be willing to punish or not punish you because you've transgressed the law. There are subtle issues, I think they're important to think about. The question of what is a normal brain, basically, envelops the whole of brain science. We don't know exactly what a normal brain means. And normal also implies normative, that is the way we want you to behave, the way we think you should behave. And people bring a lot of preconceptions to the table for that, however being able to compare your brain to my brain is an important outcome of brain research. And so there are various kinds of ways that we can map one brain and another brain into common coordinate systems by slowly stretching and perturbing one brain's image until it fits this common coordinate system and taking mine and doing the same sort of thing. By aligning them this way we can compare regions that have homologous activity and regions that look the same and look different. This sort of average brain approach, if you will, has been going on for 25, 30 years. It's quite developed. And people have very developed mathematical schemes for representing that. What's happening of late to say, how is this brain different than that brain? Is people have begun to focus in detail on individual differences and their enormous individual differences and the structure of brains, the projection pattern in the brains and the connectivity pattern in the brains and that's a world that's just been opened up. One of the new ways in which individual differences can make a difference is in this study and analysis of people with mental disorders so one thing that we can do is we can say, this person is having perceptual events and behavioral issues, so that they can't hold a job. Maybe they're clinically depressed, for example or maybe they have autism spectrum disorder but just so they're just on the spectrum. Well, we know that's not functional in certain ways and we want to improve that. And so, anything that moves it from dysfunctional to slightly functional we're interested in. We're interested in all the biological variables that flutter in that transition, the brain being the centerpiece for where you go looking for those changes. And so, it's individual differences that account for these clinical differences that has begun to catch the notice of neuroscientists around the world. We use functional imaging or any kind of imaging for that matter, to decide what somebody intends to do or whether somebody has the capacity to not do something we'd rather they not do. That strikes at the heart of what it means to control your own behavior and to be assigned credit for the control of your behavior. We know that your prefrontal cortex and its interaction with other parts of your lower brain structures is central in that kind of control. We do not have good measures yet of what it means, neurally to be in control, so I can't look at your brain and decide, oh, well, this person is exerting a lot of cognitive control over this particular act. With that said, there's a very active region of research on cognitive control and the regions of your brain that sort of contribute to it.

One is the anterior cingulate cortex and the degree to which you are planning to exert and label control on future actions can be decoded in that region of the brain. But these are still very well defined scenarios where you put people into machines, you prescribe the scenario and you have a good chance of guessing whether or not they're in control or not in control, or maybe parametrically to a degree, which they're in control. What we can't do now is put you in a brain scanner and let your mind wander and decide, oh, now they're thinking about a cheese sandwich for lunch, now they're thinking about controlling that impulse, now they're deciding to move their right finger. That's not quite on the horizon yet. That's not to say that sort of thing isn't coming. The work of Jack [INAUDIBLE] at Berkeley in decoding, I think has stunned a lot of people in its capacity to use relatively crude functional magnetic resonance imaging measures to decode very sophisticated and abstract features of your cognition. Lots of groups are engaged in this decoding enterprise and time will tell how it develops. If you're looking for some pill to dig out of someone's head for that, I think that's not the direction that it's going to go. On the other hand, I wouldn't ignore it either. These decoding methods are getting more sophisticated and they're growing just about as fast as artificial intelligence methods are growing. And we've seen what they've done recently with board games and automatic picture recognition and car driving, etc. There have been really a fever and a revolution going on there. And so, it's something that the legal system, I think, has to keep an eye on. The big issue is acclaim in neuroscience and maybe it is that sets of detectable neural processes can be associated with psychological states that you care about. Some of the psychological states that you care about we can name some of them in this particular little snippet I'm talking about, legally relevant mental states, those can be distinguished, they can be defined, they can be put to a group of legal scholars and otherwise normal everyday people and they can distinguish these things in different settings, they're more or less good at them, depending on the kind of information they're given. The question about whether those distinctions in the psychological sense have direct and readable and distinguishable neural correlates, is the question that we're asking. In a restricted scenario, the answer is yes. Whether it does sort of in the wild remains to be seen and I think it's waiting for the next generation of brain technologies that can, in a sense, go into the wild and monitor people while they do things more naturally, rather than in a setup in a laboratory where the number of variables is very small. And it's not completely shocking that we can decode variables into psychological space, it can be read out narrowly. A better question is; can you decode variables in a person as they're functioning? Your brain is doing enormous amounts of things when is doing normal functioning, it's indexing things, it's accessing memory, it's reading things, it's writing things, it's controlling your balance and your posture and your digestion and on and on. All those things are in your head, we can detect those things, but they will also interfere with your ability to decode psychological states that you think are important. So, the question is in the wild, could

you actually do it and make cogent remarks about it in the sense of informing a case? Well, there is some large-scale guidance that these findings can provide. The first is we're very early on, I think a lot more work needs to be done about this. The second is these legal distinctions, these mental state distinctions it has been claimed that they were in a sense of illusory people had used them in a way to keep the disenfranchised and we know that goes on. We would like to know that the way they're used in a court of law is reasonable with the respect to what's possible in terms of the brain. Imagine that you invented a category, a legal boundary, which 99% of all adolescence was incapable of staying on the right side of, okay, that's not a very useful legal boundary. And maybe there's a neural correlate of that, that we could track before you even worry about the legal boundary. And so, I think knowledge in this domain is going to inform the way people think about these legal distinctions. But it's not going to inform cases I think, I think what it will inform are legal scholars outside the context of court rooms. The main question to ask to what degree are you uncertain about the answer you're giving me? None of these technologies is going to tell you X is true and X prime is not true. It's not going to be like that. In fact, that's not even the way you perceive the world, you perceive the world in a statistical way, your perceptions of the world are in a very direct sense, statistical inferences over data that comes into your nervous system. Likewise, anybody eavesdropping on my nervous system while I perceive the world they make decisions, is going to be able to make probabilistic statements about what goes on in my brain and that's it. The degree to which you use that in a court to decide something, it's just another piece of data and a big quiver of data in the context of a whole case. It would be completely out of bounds to think of these things as being in any sense deterministic, certainly at this stage. And even providing overwhelming evidence neurally, that X was true, and Y was not true. We will at some point reach a stage I suspect where we are going to decode useful mental states, that is mental states that we care about, depictions of human behavior that we care about and that seem to make a difference in the way we treat people, the methodologies we use, the medications we use to move them from some area of dysfunction to some area of function. So, you have ADHD, you can't pay attention in school and we give you drugs, we do therapies, and we move you over here a little bit. We're going to understand that a lot better in the coming 20 years. Our understanding of it will still be statistical. One way to think about it, just think about the book, Grey's Anatomy, not the television show, the book Grey's Anatomy, it's a book of human anatomy, elbows, long bones, fingers, muscles, viscera, etc. Well, whose anatomy is that? It's no particular person's anatomy, its average anatomy. Your elbow has a certain kind of structure on average, sometimes people have an extra bump here, sometimes people have it hooked in differently. This goes for your nervous system too. There's never going to be one neuroanatomy and neuro function for that fits all. We're going to have to understand

individual differences and the range of variations and what these mean for the way we wanted to decide on how it builds into our perceptions.