Transcript of Dr. Matthew Goodwin presenting "Developing Innovative Technology to Enhance Research and Practice in Individuals on the Autism Spectrum: A Computational Behavioral Science Approach" at the 2019 Autism Summit on 11-02-2019.

****note: Dr. Goodwin's video presentations and his slides were not able to be shown due to a technical glitch. The slides corresponding to the transcript of his presentation are broadly outlined below.****

Slide 1:
Title Slide: Developing Innovative Technology to Enhance Research and Practice in Individuals on the Autism Spectrum: A Computational Behavioral Science Approach

Slide 2:
Disclosures:

- Scientific Advisory Board Member
- Affectiva, Inc.
- Empatica, Inc.
- Janssen Pharmaceuticals

Slides 3-6:
>> Dr. Matthew Goodwin: Okay I'm going to describe some things to you. Now it's not as exciting as showing you some of these videos, but I'll make these slides available and I'll embed those videos and if you have an archive of today's meeting you'll have full access.

So let's just think now. Open *gestures at head*

The purpose of this for me is... I have been working in the field of autism for about 20 years. I'm especially interested in those who are more severely impacted. So intellectual disability, limited verbal abilities, sensory sensitivity, high rates of challenging behaviors, stereotypical motor movements, aggression to other people, self-injury, property destruction, elopement, sort of darting and fleeing. I would argue that they require the most support from society, yet we have the least amount of information about them.

Slide 7:
If you look in the published literature in autism and you go to the participants and demographics section, overwhelming majority will show you normal intellectual, good verbal fluency. It's essentially a convenience sample for the researcher. Those individuals need support, we need to understand them better, I don't contest that at all. But you have to remember that most research is carried out in academic settings. These are typically in urban environments where you have universities. You've got very talented graduate and undergraduate students, who do high quality work and you don't have to pay them very much, and we can just make research discoveries and publish papers and get them out. That's academic currency. That's how we get promoted in the university setting. But remember, it's a strange place, with strange people, doing tasks you've never done before for ill-defined periods of time. Where you're supposed to be relaxed and acting naturally. If you have more of the severely affected profile, you are not often able to keep organized enough to travel to that setting, to have that wait time, to meet all these new people, have this total barrage of information, and for us to expect that they're going to be able to get through our assessments.

Slides 8-12:
So an early idea I had was what if we could take the lab to people instead of bring people into the lab. And so here's where an experimental psychologist with clinical interests starts working with electrical engineers and computer scientists. And about the time that I was pursuing this crazy idea there were two fields emerging out of computer science. One was ubiquitous computing, so those are cameras and microphones, radio frequency identification, motion sensors. Basically putting sensors into built spaces so that you could passively and unobtrusively collect information about inhabitant behavior.

And then there was another field called wearable computing. So these are engineers and computer scientists interested in weaving sensors into accessories or clothing or things that we can wear on the body and they're giving you information about peripheral physiology. Things like heart rate, heart rate variability, skin conductants or sweating, motor activity. This is what would be now your Apple watch, your Fitbit, your Nike Fuel, your Jawbone UP, there's a huge consumer wearable market now doing on body sensing.

The commonality between these two disciplines was they're able to record data, passively without conscious input from the person who's being observed, and that they all have a common time stamp, so any of your different sensors are referencing the same clock. And all that data is being stored in a centralized server that a researcher can access later and do offline analysis, or to the extent that you know what signals you want to see in the data you can do real time machine learning processing of that data, and create a kind of just in time triggers or adaptive feedback.

So if you've noticed if you're sitting, do people have an Apple watch? I don't work for Apple. Every now and again it'll tell you stand up. That's because it's been keeping track. There's an accelerometer in there. Accelerometer, when you have your phone and you go from portrait to landscape that's how it knows that you're turning it. That's what an accelerometer is. It knows where something is in space and time and three dimensions. So when you're sitting for a long period of time, it's basically doing some minor tracking over a period of time. How much motion? If not much, does he stand up?

Okay, so now translations. So stereotypical motor movements, hand flapping, body rocking, and finger flicking. It's one of the core features of autism. We see a higher occurrence in those who are more severely impacted. The consensus definition in the literature, I know we have behavior analytic people here, but if you sort of broaden to the bio psych community, they'll say that it's invariant in form, that means it always looks the same to our eye. It has no obvious soliciting stimulus. Meaning we don't know what may have occasioned this behavior. And it has no adaptive function. That is it's purposeless on the part of the individual. Now it's not tics, like Tourettes, it seems to be under regulatory control, but it seems to be just sort of aberrant behavior. And when the bio-psychologists look at this, they conclude that it's behavioral output of dysregulated neuronal systems, usually dopaminergic pathways and differences in the basal ganglia. But you have to remember that that, those conclusions are based on animal researchers who are looking at mice, cats, rats, pigs, and non-human primates who perform an anthropomorphically similar behaviors of repetitive grooming or shredding or self-injury, where they can then sacrifice those animals post-mortem and then they find those differences. We don't do that with humans, and we don't have a very big bio-bank of brains with good phenotypic behavioral data, so we're sort of making an assumption that that must be what's happening in autism.

But I want to just put a very simple thing out here for a second. Of all the things that we can do with our bodies, why flap, rock, and flinger flick? And you know if this is happening with the more severely impacted, they're not going to conferences, they're not on twitter, they're not snapchatting each other saying, "You know when you feel stressed, have you ever tried doing this?" Yet they're all engaging in very, very similar kinds of behaviors. I don't think we can explain that away.

And in fact there have been very interesting alternative hypotheses about these behaviors, that they regulate stress. In an overwhelming situation this may be calming. In an underwhelming situation this may be arousal producing. In the absence of language, when I engage in these behaviors, you may approach me and you're now an instrumental tool to carry out a goal that I can't do on my own. Or it may signal to you to stay away when I need a moment of reprieve. This may be how people feel their body better in space and time when they lack proprioceptive feedback. There are other clinically useful benefits of those behaviors. They may not start that way, but the environment may shape over time, and people's reactions may shape some time, that those become functional purposeful behaviors in the absence of achieving them in a more neurotypical way.

So for the nerd part of me, this becomes a measurement issue. So how do we establish whether a behavior is staying the same, increasing or decreasing? We have to do that through measurement. And if we want to know about a response to an intervention, or higher frequency, intensity, or duration of those behaviors is a function of your setting, the activity you're engaged in, the people that you're with, your affective state. I need to have very high fidelity measurement of those movements. We typically take these data in the field with a survey. You ask a parent to fill out 29 questions. What's the typography, what does it look like? About how often does it occur? And how disruptive is it?

So that's good. Presence/absence, kind of global. But you don't administer that survey moment by moment so you really don't get any information on intra-individual variation over time. People do live observation of these behaviors. So two of you would watch me flapping and rocking, and you would be independently jotting down was it a flap, rock, finger flick? When did it begin? When did it end? And then we can calculate frequency and duration. And you're not supposed to be conferring, you're doing that independently. And then afterwards we'll look at inter-rater reliability. What's the rate of agreement and disagreement between our periods? In the published literature, the highest inter-rater reliability I've seen is 33% for live observation. Think about it. If those two people turn their back and guess they have a 50% chance of agreement. You're going to do worse attending than you are just making it up.

You can videotape these behaviors, you can put them frame by frame offline, and we can get above 90% inter-rater. And I did this for my doctoral work. I would not wish this upon anybody. It's very boring, it's very time consuming, it's easy to get distracted. You have to keep drinking coffee. You have to keep reminding yourself, why am I doing this? Is this going to help people? So it's good but it's not scalable.

So what we sort of set out to do was say, “Can we come up with a method that is as simple as administering a survey?” So as efficient, but would be as accurate as though you were doing frame-by-frame ground truth coding. And so we used early versions of sensors that had three access accelerometers. We put them on 10 individuals with autism who are more severely impacted, we recorded those sessions, observations of them, when they're engaging in hand flapping and body rocking that we had observationally defined. The sensor is time-synchronized with the camera, so I have two independent people who are rating all those videos, that becomes our ground truth, of when those behaviors are present.

And then we give about 10% of that data, that's been labeled, to a pattern recognition algorithm. It's a decision tree classifier. And so it's training on what are the features that discriminate a flap and a rock and a finger flick from just general movement or no movement at all. That was one of the cool pictures I'm going to show you, is those streams coming out and you can very quickly see and discriminate with your own eye when he's engaging in these four different classes of behavior. And as your eye can see them and will in the future, we're mathematically now describing those. And basically 10% of the data that it's trained on, we then give those classifiers 90% of the data it hasn't seen yet, and we're seeing how well does it correctly identify a flap versus a rock versus a finger flick compared to the ground truth video, annotations of humans. Lot of cool technical stuff that would bore the heck out of you.

We get about 95% average accuracy with less than .05% false positive rates, with that classifier identifying those behaviors, as though you were coding on a video. That means that a person, if an individual is wearing an accelerometer on each wrist and on the torso, if you use a cellphone and you opportunistically said, "He's doing it now and then he just stopped." "She's doing it now and she just stopped." And we get 50 of those examples, that classifier can now run on that phone and detect correctly 9 out of every 10 onsets and offsets of those three different classes of behavior, as though you were doing ground truth rating by video.

That frees humans up to now get very high accuracy, but very efficient measures of these behaviors, that we can now start to test hypotheses about what function does it serve. Because I personally think that because they are socially stigmatizing, and because people tend to make inferences about your intelligence when you're engaging in those behaviors and they're underestimating you, that if these are about communication, if these are about stress regulation, if they are about feeling your body in space and time or communication, can we teach a more socially acceptable way to achieve those functions? That would do enormous for quality of life.

What do you guys colloquially call this class of behavior?

>>Audience: Stimming.

>> Dr. Goodwin: Alright, what is stimming shorthand for?

>>Audience: Self-stimulation.

>> Dr. Goodwin: Self-stimulation. There's not one published study demonstrating that these behaviors have a self-stimulatory function. Now I'm not saying that they don't because I've seen in some individuals that I think that's the case, but we have all automatically decided we know what's going on... Based on what information? Across hundreds of thousands of people doing the same behavior with high measurement, reliability, and validity.

Slides 13-16 :
So this is going be an example where I would like us all, the technology piece for me is trying to get behind or inside why this is happening, and not to assume I know what it is and try to shape it away. So the next example that I was going to show you is using biosensors now, that are not just about physical activity but also have, physiological recordings and looking at aggression to other people.

So I'm fortunate to be a part of a network of seven different psychiatric inpatient units around the country, that are all using the same infrastructure for data collection. This is initiative from the Simons

Foundation and the Nancy Lurie Marks Family Foundation called the Autism Inpatient Collection. And it's Colorado Children's Hospital, it's Sheppard Pratt, and Kennedy Krieger in Baltimore. It's Bradley Hospital in Providence, Rhode Island. It's Western Psychiatric in Pittsburg, and Maine Medical up in Portland, Maine. They have the highest rates of referrals of individuals with autism who are aggressing to themselves or other people. It's the number one reason why families ask them to do an inpatient stay.

So they're there typically for 2-6 weeks at a time and so we did a study using the E4 by Empatica, the company I do some work with. So this records motion activity, cardiovascular, electrodermal, and skin surface temperature. And then we had a cellphone application where we asked research assistants to annotate opportunistically when 20 individuals in the inpatient unit were engaging in aggression to other people. Kicking, biting, hitting. And they're clinically being responded to. These are naturally emitted behaviors, we're not eliciting them. And so the clinicians are responding, as they typically would, but the RA is now helping us get some of that ground truth data. Videos would've been the best way to do it, but we can't record in this environment.

Slide 17:
So in those 20 individuals we had 87 hours of data collection. Four hours per participant. We saw 598 aggressions while these individuals were wearing the biosensor. So now we can go and separate that data into periods of time when there's no aggression, and when there's aggression. And then we can look at what features in the physiology discriminate aggressive periods versus non-aggressive periods.

And then we can do an interesting experiment where we say how much data from the past do we need to make a prediction how far into the future at what level of accuracy? And where we are, kind of what we've just published in autism research a couple months ago, was the last three minutes of data enabled us to predict an aggression one minute into the future with 84% average accuracy, if you train the classifier per participant. And 71% accurate if you use just one model for everybody.

Slide 18:
So that is to say, if someone's wearing the sensor for 3 minutes, I can predict correctly 8 out of 10 times that they're going to aggress in the next one minute before they have. And we've just recently done support vector machine on these which is the fancier way of doing classification, and we can use two minutes from the past to predict 3 minutes into the future with 98% accuracy. So we have funding now to scale that up to 270 inpatients across three different inpatient sites, so we can start looking at the generalizability in the population in the setting for carrying this out. And then we'd like to try to take this into real time push notifications. So I could, when the algorithm running on the phone or in the cloud detects the high likelihood, it can push a notification to a phone, to a caregiver, to now attend, triage, keep people safe, try to preemptively intervene.

Slide 19:
These behaviors are typically considered, escape, avoidant, oppositional, defiant, or forensic. I would like us to reconsider and think of this as maladaptive stressing coping responses in the absence of language. And all the comorbidities we've been learning about this morning and what we know about the core characteristics of autism, the world is not a very predictable place. It's scary, people are in pain, they can't communicate it. They're not overtly, we're making inferences on people's internal states based on their overt presentation, and often they're flat. That's another video I can show you. I would like to think of these as fight or flight responses. Any mammal, any reptile, if it is in a position where it feels like it's personal safety is threatened, will attack or will flee. This is self-preservation.

Slide 20:
It changes our conceptualization of this though, because it's not about dangerous kids doing dangerous things that we need to seclude and restrain and medicate. It's that we could play a role in trying to identify- thank you. Trying to identify what is causing that distress that is leading to these behaviors? And think about this more as an emotion dysregulation issue that we could be teaching better coping strategies that make those behaviors non-necessary. Right?

Slides 21-23:
And then the last thing that I was going to show you, which gets Ron excited, and it gets me excited. And it gets Ron excited because I need help with this dataset. I did a study with the Simons Foundation where we took off-the-shelf commercially available cameras, microphones, daytime physiological sensors, like the one I keep pointing, to that doesn't mean anything to you. And that's how it does the cardiovascular, that's called photoplethysmography. So every time your heart beats it's pushing blood to your extremities, and there's oxygen in blood, and so these two different color arrays can basically calculate pulse transit time, how quickly your heart is beating and then you can get heart rate variability derived from that. So these are good measures of sympathetic nervous system arousal.

>> Audience member: Did you have to desensitize the individual?

>> Dr. Goodwin: So good question. So we got usable data from all 20 participants in that first study. We tend to have a pretty high acceptance rate. Some folks will just put it on and they could care less. Others we have to model it, we have to come up with social stories and scripts, we have to shape, you know sort of start, five seconds and get up to 40 minutes. We reinforce the heck out of them. You know we put stickers on them, we can personalize it. This can look like a wonder woman band, or this can look like a Viking band. You know to try to- most cases, once it's on they tend to forget about it. It's also shock proof and water proof. This is partly what took so long to do these studies, is we had to get these all built in a way that I know that in an unsupervised way they could be worn and I don't have to worry about it being damaged, and swallowing electronics, etcetera.

So we had commercially available cameras, microphones, physiological sensors. We had to write a code base so that all of those sensors could write to a same common data set with the same time stamp. And then we had 20 families with autism who deployed this in their homes for up for four months with zero instructions. And we were just going to see what would we get? And what is the quality of that data? If memory serves, we have 4,000 hours worth of video. We have 1,300 hours of biosensor data. We have 116 nights of sleep through sensor-based actigraphy. And then I've asked my graduate students and undergrads, basically anybody who's willing to help, to watch these videos after they have CITI certification in human subjects, for identifying clinically relevant behaviors in those data sets.

Slide 24:
And now I would like to share that with people who do work in machine learning and deep learning, and see if a human and machine in the loop. So if a human can identify, these are periods of time when someone's doing something that we think is interesting. Can an algorithm do computer vision, audio and speech analysis, bio-signal processing, multimodal data fusion, and basically say this is how me from a machine, would identify that label. Or the machine says here are times when we see recurrence across these different sensor streams, and it's showing, it's a human saying, "Is this interesting and what is it?" And in this way we will do what's called transfer learning where the machine is doing automated detection, humans are doing detection, and together we can try to… So here's the issue, it's

tremendously valuable data about function and about response intervention. But it would take 3-4,000 hours just to watch the video. Let alone to splice it up and code it and do any analytics of it. So how in this sort of data rich world where we're doing deep learning. I mean, think about Siri. I mean, think about all the interactions you have that you've never had to do any training for and it works great. We can be doing that with health data and have much, much, much better measurement to think about functional analysis and response to intervention.

Slide 25:
So I'm really sorry. This is a very boring way to present this work. I promise you it's much cooler with videos and that will happen later. Thanks.

[audience clapping]