Episode 133: M. Brooke Robertshaw

# MEDS: Mary Ellen Dello Stritto

# BR: M. Brooke Robertshaw

# KL: You’re listening to “Research in Action”: episode one hundred and thirty-three.

# [intro music]

# Segment 1:

# KL: Welcome to “Research in Action,” a weekly podcast about topics and issues related to research in higher education featuring experts across a range of disciplines. I’m your host, Dr. Katie Linder, research director at Oregon State University Ecampus, a national leader in online education. Along with every episode, we post show notes with links to resources mentioned in the episode, a full transcript, and an instructor guide for incorporating the episode into your courses. Visit our website at ecampus.oregonstate.edu/podcast to find all of these resources.

**MEDS:** I am your guest host, Dr. Mary Ellen Dello Stritto. I am pleased to bring you another episode in our periodic series focusing on quantitative methodology and statistics.

On this episode, I’m joined by Dr. Brooke Robertshaw, an Assistant Professor and the Assessment Librarian at Oregon State University. Her current research interest revolve around the ethics of learning analytics, with a particular interest in the contextual nature of quantitative methodologies. Brooke is a member of the data doubles team that is exploring student perspectives of learning analytics. She is passionate about quantitative literacy, social justice, and the intersection of the two. In her spare time she enjoys white water and flat water kayaking, discovering ways to give voice to the voiceless of the diaspora in the Middle East, and traveling to Jordan to spend time with her dear friends there.

Thank you for joining me today, Brooke.

**BR:** Oh thank you – yeah I’m thrilled to be a part of this podcast. I’ve been listening to it since it started.

**MEDS:** Great. So I’m really pleased to have Brooke in the studio today, and Brooke and I have had previous conversations about statistics, and I know one of her passions is talking about this topic of effect size. So let’s start our conversation about what an effect size is. Can we start there?

**BR:** Yeah. So an effect size – I have – I use a lot of different ways of explaining effect size. Um you’ll hear people talking about it’s a measure of magnitude. Um I had a student use the idea of a magnifying lens. So um – so when you run your statistical test you get a p-value, and a p-value can be seen as dichotomous, so is it statistically significant or not? The effect size then says how statistically significant it is, so it’s like holding up a magnifying lens and you see, and so you can see the noise. You can see how much is actually happening. And so yeah, so if you think about the p-value being yes or no, and then the effect size saying how, yeah, exactly how much is going on. So that’s a really simple definition of an effect size.

**MEDS:** So the effect size gives us kind of a magnitude of effect, is that correct?

**BR:** It uh, magnitude of effect or an effect size will actually tell us how much is actually going on, and so. So for me the effect size, you can have magnitude or you can have something going on, even if you don’t have statistical significance. Does that make sense? Did I say that right? Because you and I had talked about this before, and I always, like I want, I wish this were a video podcast, because I want to have my white board. Yeah so it’s, it’s a way of saying that something, it’s a way of, magnitude isn’t always the best term, because it’s not the easiest.

**MEDS:** Way to understand it?

**BR:** Yeah.

**MEDS:** So maybe it would help to talk about effect size in terms of how it’s calculated. Can we talk about that for a minute?

**BR:** So there are a bunch of different effect sizes**.** And so you have Cohen’s d, you have Hedges’ g, you have, uh I have a study out there with one of my mentors, Dr. Todd Campbell, and I use this effect size called Varga-Delaney’s A. Yeah.

**MEDS:** Right. So they all may be calculated differently, but…

**BR:** It’s all looking at means, and standard deviations. The most common is Cohen’s d.

**MEDS:** Is Cohen’s d. right. And in Cohen’s d, in general what they’re doing is they’re looking at is a difference between means, divided by some sort of standard error or standard deviation.

**BR:** Right. The whole standard deviation.

**MEDS:** Right. Exactly. And what that statistic can tell you is do you have a relatively small, medium, or large effect.

**BR:** And that’s debatable.

**MEDS:** and that’s debatable too.

**BR:** It’s all debatable, and it’s also - because then you also have your Ada squared, your partial ADA squared. So when we talk about effect sizes, so Mary Ellen asked how – you know, how do you calculate effect sizes? And so if you’re talking about Ada squared, Ada squared is, you can calculate it using your – so Ada squared is the effect sizes for ANOVA, and you can calculate it using your – Using your F-table from your ANOVA table. So yeah. So there are multiple effect sizes. Actually um your correlation – so in a correlation table, the r, that actually acts as an effect size as well. Um and so I’m actually working on a meta-analysis with a colleague at Indian University, and there’s so many different effect sizes. Um the nice thing about effect size, so that, Mary Ellen, previously you had a podcast on p-hacking. And so one place where effect size comes in is, people will often add, so they don’t have statistical significance, and they will go out and they will add, they will add to their sample size. They will increase their sample size so that then they can get statistical significance. Well you can have fifteen thousand people in your sample size, an n of 15,000, and you’ll get statistical significance. But if you run an effect size, that effect size will, sometimes it will show you that you have actually nothing going on. So that’s where the effect size kicks in.

**MEDS:** Good, and that was my next question. Is, you know, how do we use this effect size? So this is a statistic that we can use in a particular way, so explain.

**BR:** Um yeah so, so the effect size come in, because it’s yet another statistical test to look at exactly what’s going on. So you have you n of 15,000 or you have you n of seven, and you run that statistical test and it gives you that p-value, and like I said earlier, p-values aren’t as dichotomous as most people think they are. And I say most people, it’s – yeah. Most people think they see statistical significance and they say, “Yay! Woo-hoo!” and then people like Mary Ellen and I are like, “No. Wait a sec. What’s your effect size?” Because people, because we know that there’s stuff going on, and an effect size is a way of taking a deeper look into your stats. So it’s a way of saying, okay you have an n of 15,000, we know that – we know that you’re likely to get statistical significance, now let’s actually see how much is going on.

**MEDS:** Mhm. Or is there actually an effect here.

**BR:** Is there really an effect?

**MEDS:** Or if you’re doing a comparison between two groups, are they really different? Right?

**BR:** Yeah, and how different are they? And so that’s where, I was actually working with an engineering student who was working – who had come over to the college of to the college of ed, and was doing social science research for the first time, and I was trying to think of a way of explaining this, and she was like, “it’s like a magnifying lens.” So we’re taking a magnifying lens to the variance to see actually how much variance there is, because – because it. I was trying to like – the p, running the – getting the statistical significance will tell you that there’s noise, but the magnifying lens, the effect size, actually allows you to actually see how much noise there is going on.

**MEDS:** Hm interesting.

**BR:** Yeah. And then another engineering student was like, “Oh! So you just like – “she works – she’s a civil engineering student and she was like, “So you’re shaking out all the dust to see what’s actually going on.”

**MEDS:** That’s a good analogy.

**BR:** Yeah! So it’s like working with these engineering ED students is really cool. They’re helping me come up with different metaphors that – yeah. So that’s why I’m not a big fan of the word magnitude, because especially translating that to different cultures or just different ways of thinking, and so that’s why I’m always trying to come up with different metaphors. And so I use the word ‘noise’ – like looking at how much noise there is actually, and you can’t see me, but I’m actually raising my fingers back and forth. Um but yeah so – so when we’re running those studies. But then – so I actually have – I referred to it earlier, back during – one of the last things I did during my Ph.D. program was I did a study with one of my mentors Dr. Todd Campbell in science ED where we had this n of 7 repeated measures, um and this study is not generalizable, but with an n of 7 repeated measures, we didn’t have statistical; significance. And so then we ran an effect size, and we were able to see some effect. Again this isn’t generalizable, because it’s such a small n. But the effect size allowed me to pick up – to like pick up on the fact that there was something happening. Where as in some cases – on the other end of the spectrum, you’ll have an n of 15,000 – you’ll get statistical significance and you’ll run the effect size and there won’t be any effect. And then there’s all these cases in between. And so the effect size, and actually this is an APA thing, where APA – I think the citation is back in 1996 where APA actually said, “You need to be reporting not only statistical significance, but also effect size.”

**MEDS:** And that’s the American Psychological Association for folks who don’t know APA.

**BR:** Yeah. And so I come from education, and that’s where my Ph.D. is from – educational technology and learning sciences. And so we follow – that’s who we follow. Yeah. And so they also talk about reporting confidence intervals around, if possible, around your means, and then I’ve just recently discovered reporting confidence intervals around my effect sizes.

**MEDS:** So yeah. So many statisticians, experts in the kind of quantitative realm, are really saying we should be doing this with all of our statistical tests. That we should be reporting effect sizes. It may not be happening kind of widespread, but that is the best practice. Is this idea that we could potentially have statistical significance, but it could be meaningless.

**BR:** It could be totally meaningless.

**MEDS:** especially if you have, like you were saying, these very large Ns. And so this is another index or another bit of statistics that will tell us more, and I like your magnifying lens idea. So I think that’s really good.

**BR:** And to me it falls into an ethical practice, because especially as we’re moving in to more and more decisions being based around statistical – around data, and I see more and more studies where people are saying, “we have statistical significance” and the first thing I look for is I look for an effect size.

**MEDS:** Can you provide an example of how using an effect size could really help with like a major decision?

**BR:** Yeah so – I’m not going to give a specific examples, because I don’t feel comfortable calling people out, but it’s an idea of – so coming from the field of – coming from the world of instructional technology where, you know, we go and we do research on implementing a technology and you say, “Hey. I’m going to implement this technology and I’m going to watch the impact.” And so you do your qausi-experimental design, and these people have access – you know, these students have access to the technology and these students don’t, and we say we implemented a cross – you know, like a whole series of English 101 courses, and we do our treatment control, you know, quasi-experimental, and then we show statistically, you know, we just want our statistical significance and these students perform better than these students, and so now we’re going to make this business decision to use this technology. Well, the n is high enough that we’re going to show a difference, but we haven’t ran the effect size.

**MEDS:** And is that a practical difference, right?

**BR:** Is it practical? We haven’t run our confidence intervals around our means, and our test statistic, or our effect size. We haven’t even ran our effect size and we haven’t even run the confidence interval around our effect size, so we haven’t followed the APA best practices, which have been -you know, which were written about 22 years ago? 1996 was 22 years ago? Yeah. And so we get into the ethics of when quantitative, when quantitative decisions start impacting business decisions in education, and what does that look like and what is our due diligence around that to make sure that our analyses are as rigorous as possible, that tour studies are as rigorous as possible? And for me, that effect size is just one piece – its. I love effect sizes. And one of my many mentors, so lucky to have so many mentors, but Andy Walker was like, “Effect sizes are so important,” and it took me awhile to learn why, but if I think about effect size plus the ethics of higher ED, and the millions and Billions of dollars that we spend in higher ED, it’s just so important.

**MEDS:** That’s great. I think that’s a good place to stop. So we’re going to take a brief break, and when we come back we’ll hear more from Brooke about best practices for using effect sizes.

The Research in Action Podcast is definitely a team effort, and I wanted to give kudos to our Oregon State Ecampus Multimedia Team, who ensures the podcast is the high quality production that it is. OSU Ecampus is home to national award winning multimedia developers and instructional designers, facilitating the highest level of student engagement in OSU’s online courses. See what else the team is up to by previewing what it’s like to learn online with Oregon State ecampus.oregonstate.edu/demo.

# Segment 2:

**MEDS:** So now that we’ve talked about the fundamentals of effect sizes, I’d like to go a little deeper to discuss some of the best practices for using them. So let’s begin there. What do we need to know about using effect sizes practically?

**BR:** You just always run them, you run them because they’re that important, because I’ve seen too many studies where people don’t have them, and my, I mean that’s actually the first thing I look at when I see a quantitative study, nowadays, is what’s the effect size? And if they don’t have it, I start looking for how can I run it, because I actually want to see it. And I know there are debates about – I mean there’s always debates in statistics about effect sizes, um and people always poking at things, but to me effect sizes are – it’s not the be all and end all statistic, but if you just have that statistical significance, I need to see that and I need to see confidence intervals, because I want to see how precise your study is as well.

**MEDS:** Yeah and I think that’s really great. I mean, and one of the things we try to do is to of course promote quantitative literacy and this is a piece that even if you are not actually doing this statistic, but if you are the person who is reading the journal articles, reading the research, you should be looking for these things, because it’s a piece of quantitative literacy.

**BR:** I’m the assessment librarian here at Oregon State, and I even, so if you’re doing assessment, there’s also debates in the assessment community about how much literacy do you need to have in order to be doing assessment, and be doing evaluation. And a lot of those assessment reports make a difference in how money is spent and those decisions, and if you’re not running the appropriate statistics. And so – and it needs to be written so another best practices, and this is why I come up with all different ways of explaining what an effect size is, it’s not just so I can teach my students – but the initial driver for me to try to explain what an effect size was, was so my boss, so my first job out of my Ph.D. was at Purdue, and so my different bosses could understand, or their bosses. So people high up in admin could understand what an effect size was. Because they needed to be able to make decisions about what to do next about how to spend money based on my calculations, and so if they didn’t understand what the effect size was – if they didn’t understand and so I needed to be as diligent as possible. So in assessment you need, so one of the best practices, you need to run your effect sizes. You’ve got to run your effect sizes, and then you’ve got to explain in in a language that people who don’t have the statistical literacy – so you if you – I mean, I’m in higher ED. I’m an assessment librarian in higher ED and my boss is freakishly smart, but she doesn’t necessarily speak stats and she’s trusting me to explain to her what this means. Make sure your running the right effect size, because there are multiple effect sizes. Reach out to me – I’m really easy to find. Um I can’t always help you, but I can point you to resources. If you’re part of the OSU community, oh by the way, I can help you because it’s part of my job.

**MEDS:** So let me go back to something you mentioned in the first segment, and that’s this concept of confidence intervals, so let’s circle back to that. Can you talk a little bit more about how you use confidence intervals?

**BR:** Yeah so confidence intervals are, so overall, the idea of confidence intervals are a way of looking at the preciseness of the statistic that you run and how precise this statistic – the statistic you run is an estimate of the population. So I’ve seen it described as a precision, um a precision estimate. So you can run confidence intervals on – so you have confidence intervals on test statistics including your effect size if it’s normally distributed.

**MEDS:** Okay.

**BR:** So yeah. So you run your effect size, and then you can – and then if it’s normally distributed you can run precision round your effect size. So recently I was looking at a student’s work, and they had, they had a 95% confidence interval that was just huge, which made sense because their n was like 50. Um and so to me that says, you know, this study is – so this particular um piece that they had run is not a very good estimate of what, you know, of what the population would look like. Then you get to see how precise it is. So that’s even like, so that’s even more information that you’re providing yourself. That’s more information than you’re getting about your quantitative analysis that you’re getting about your study. I mean the more information we could get about – the more information we could get about our study, the better, right? To me stats is a language, and effect sizes are just one piece of this language of stats, and it gives us more information. And the more information we can get and the more literate we can become, I can’t – I can’t tell you about my study. I mean, I can try to give you some descriptions and I can try to translate that, but if you understand it and the more information you can give numerically, the better.

**MEDS:** Right. And this concept of confidence intervals I think is well named, and a lot of statistics aren’t well named. But in the case of confidence intervals, it is how confident are we that the sample the we are looking at and this finding that we are trying to find with in our sample, you know, comes from this population that that exist out there? And I love how well that confidence interval concept really encapsulates this idea of we’re confident that what we are showing in our study is actually reflective of our population. Or not, right? So let’s go on to some resources. So can you suggest some resources for our listeners to learn more about effect sizes?

**BR:** Um. Yeah. Uh to learn more about effect sizes, there’s – there’s a really great article about – it’s called, *It’s the Effect Size, Stupid*. And that’s just a great title. And so if you – the author’s last name is Coe. Um so that’s a resource that you can learn about effect sizes. Um and then –

**MEDS:** And we will have links to these sources in the show notes.

**BR:** Um yeah. And then I have resources about where you can run effect sizes on my website. So I have a personal website, and believe me I didn’t come on the podcast to promote myself. I came of the podcast nerd about effect sizes. Um but if you go to stats.brookerobertshaw.com, I have a bunch of just stats resources for the behavioral sciences for education, and the behavioral sciences for education, and I have a section on effect sizes. So um yeah. So there’s a place where you can actually go run different effect sizes, because they’re used in meta-analyses, and so there are a lot of different ways you can run effect sizes. Um yeah so there’s campbellcorportation.org has got a whole set of effect sizes, and I vetted that – I’ve looked at the person who developed that. And then there’s freees.usu.edu that my professor Andy Walker developed, and you can run effect sizes and actually store stuff there.

**MEDS:** Great.

**BR:** Yeah. I’m pretty sure that it’s okay that I’ve – so these are resources where you can run effect sizes, because that’s often really difficult – is to, because SPS has got it built in in one place that I know of . Yeah.

**MEDS:** That’s a great point that maybe our listeners don’t know, is that sometimes this is over and above the basic statistical packages that you may be using.

**BR:** Yeah. It really is, which is bothersome for me, because it should be. Like, when you run a t-test in SPSS, you should be able to – it should just give you the effect size. Uh and I know there are a lot of people who dis on SPSS, but it’s what I use, it’s what I use *R* to sometimes, but it’s what a lot of people in education in the behavioral sciences use, because we’re not statisticians. I consider myself a data analyst. I consider myself and educational researcher. And yeah, so effect sizes aren’t necessarily built into some of our stats packages. Other places, um, yeah. I’m just trying to think. There’s an APA – there’s that APA article, but actually it’s really difficult… I’m trying to think of places – like to learn about them. Like I think

**MEDS:** that’s a good point, they’re not always easy.

**BR:** when we were talking about it, because it’s – even though there’s this article, even though APA put out this recommendation there’s not - I mean, all kinds book talked about, the new stats book talk, like I’m seeing them being talked about. My stats classes actually didn’t talk about them, I learned about affect sizes because I was in a research group for five years during my Ph.D. program, um and Andy just, Andy talked about them. Um – but where to learn about them, that’s a really good MEDS: Question, yeah.

**BR:** That is a – that. You learn about them from people like me and Mary Ellen who.

**MEDS:** And this is why we have an episode on effect sizes) – who are like really passionate. And people like Andy walker who are really passionate about them, who are really passionate about this really tiny statistic that makes a huge difference

**MEDS:** and the power of it.

**BR:** And the power. I mean the power – Because let’s face it, statistics are changing people’s lives right now. Every day your life is changed. If you go online, your life is being changed by statistics, and effect size is one that people are not using. They’re just not using it enough. Go out and pull any research. If you go out and you pull ten research articles, I would bet that in social and behavioral sciences – I can’t speak for out of the social and behavioral sciences, but how many of them have effect sizes? I don’t know

**MEDS:** Not so many from my experience.

**BR:** Not so many. And yeah there’s a whole other discussion about - but yeah. Look at my website. And my website has a section of effect sizes and how you can run them, and this is a great point, that I need to developed something about why you should run them.

**MEDS:** For your website. There you go.

**BR:** For my website, and actually just because in my role of assessment librarian, like his is good practice. Yeah. Thank you.

**MEDS:** So thanks for talking with me Brooke, and thanks to all our listeners for joining us for this week’s episode of Research in Action. I’m Mary Ellen Dello Stritto, join us next week for another episode.

# Show notes with links to resources mentioned in the episode, a full transcript, and an instructor guide for incorporating the episode into your courses, can be found at the show’s website at [ecampus.oregonstate.edu/podcast](http://www.ecampus.oregonstate.edu/podcast).

# There are several ways to connect with the “Research in Action” podcast. Visit the website to post a comment about a specific episode, suggest a future guest, or ask a question that could be featured in a future episode. Email us at riapodcast@oregonstate.edu. You can also offer feedback about “Research in Action,” episodes or share research-related resources, by contacting the Research in Action podcast via Twitter @RIA\_podcast. Finally, you can call the Research in Action voicemail line at 541-737-1111 to ask a question or leave a comment. If you listen to the podcast via iTunes, please consider leaving us a review.

# The “Research in Action” podcast is a resource funded by Oregon State University Ecampus, ranked one of the nation’s best providers of online education with more than fifty degree programs and over one thousand classes online. Learn more about Ecampus by visiting ecampus.oregonstate.edu. This podcast is produced by the phenomenal Ecampus Multimedia team.

# “Research in Action” transcripts are sometimes created on a rush deadline and accuracy may vary. Please be aware that the authoritative record of the “Research in Action” podcast is the audio.