Episode 116: Mary Kite

# KL: Katie Linder

**ME:** Mary Ellen Dello Stritto

**MK:** Mary Kite

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

# [*intro music*]

# Segment 1:

# KL: Welcome to “Research in Action,” a weekly podcast where you can hear about topics and issues related to research in higher education from experts across a range of disciplines. I’m your host, Dr. Katie Linder, director of research at Oregon State University Ecampus. Along with every episode, we post show notes with links to resources mentioned in the episode, 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.

**ME:** 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 statics. On this episode, I am joined by Dr. Mary Kite, Mary received her bachelors, masters and PhD from Purdue University. A social psychologist, she is currently a professor of social psychology at Ball State University. Strongly committed to psychology and education at all levels. She is past president of the Society for the teaching of psychology, she has held a number of leadership roles also for the society for the teaching of psychology. She also shared the American Psychological Association presidential task force, on diversity in education resources. And is past president of the Mid-western psychological association. She is a fellow of the American psychological association divisions two, eight, nine, 35 and 44. And the association for psychological science and the mid-western psychological association. She maintains an active research program in the area of stereotyping and prejudice, including co-authoring the book *The Psychology of Prejudice and Discrimination* in its third edition with Bernard Whitley Jr. Kite and Whitley also co-authored the book *Principles of Research in Behavioral* science in its fourth edition. Her recognitions include the Charles Brewer award for distinguished teaching in psychology from the American psychological foundation in 2014 and a presidential citation from the society of the teaching psychology in 2011. She was selected as a G Stanley hall lecturer for the American psychological association in 2009. And was named a minority access role model in 2007.

Thank you for joining me today, Mary!

**MK**: Thank you, Mary Ellen, it’s a pleasure to be here.

**ME**: Great! So you have extensive experience as a professor and researcher in psychology. So, I wanted to begin today with discussing external validity and its importance in quantitative or experimental methodologies. So, can you begin by defining what we mean by external validity in the world of research?

**MK**: Sure. External validity addresses the question of whether results of a research study hold up under new conditions. And there are actually two ways to think about external validity. One way is to ask the question of whether the results of a particular setting or a population applies to my setting or population. So, for example an applied researcher or practitioner might what to know whether the results of research would work in my organization or my clinical population. If I was a human resource officer, I might want to know if a study on effective interviewing strategies would apply to hiring in my own organization. Or if I was a clinician working with a particular population, I would want to know whether the mindfulness techniques that have been shown to reduce stress and anxiety in clinical populations would work for the particular population that I have in my setting.

Another way to think about external validity is whether the results of a particular study would hold up under new conditions. And this is a question that a researcher is more likely to ask and it’s also known as generalizability. So, I might want to know, for example, whether the results of research conducted in a laboratory setting would generalize to an applied setting. So one example is if you look at the classic studies by Darley and Latané that were conducted in the laboratory. They showed that people are less likely to help when more people where present, than if they were alone in a particular situation. And it turns out that these results so replicate and cross lots of settings. For example Piliavin and Piliavin replicated these using a field study in a subway, and found the exact same pattern. So we can be pretty confident that, that phenomenon generalizes across settings.

Another common external validity question is whether the results apply to different people or populations. For example we often want to know whether results hold up cross culturally or with different ethnicities or whether women and men both respond the same way.

**ME**: That’s really interesting and it’s important to think about you know how the research that we do um may look different as you said in laboratories versus out in the real world. And that’s really what we are trying to get at with this issue of external validity.

**MK**: So, the questions are important because they let us known whether the results of research conducted in laboratories accurately represent the nature world. So in the helping studies that I described really briefly we can be pretty confident that, that relationship holds no matter where we look for it. So one question that is often asked of laboratory research is whether the behavior that we are studying looks like behavior in the real world. And sometimes it sounds like it really doesn’t. So let me give you an example. There is a whole lot of classic research where people are brought in and asked to memorize a list of non-sense syllables. And if you were not a psychologist and maybe where thinking about this you might ask yourself “Well when am I ever asked to memorize non-sense syllables. That’s not something I do”. So what does that research have to do with how people remember information? Well even though we don’t in our real lives sit down and memorize lists of non-sense syllables, except when we are trying to remember all the different kinds of passwords we have. It turns out that that laboratory research consistently shows us a lot about the psychological processes that underline memory. And gives us a lot of information about how quickly we forget about the things we learn. So even though at first glance it seems like those laboratory studies are really artificial and don’t apply to real life, they actual do apply to how people remember information.

**ME**: Ah that’s a great example. Um and many times you hear criticism of folks who say “Oh what does that have to do with anything that happens in my life?” and I like how you kind of brought that up and brought those things together. So that’s great.

So we are going to take a brief break and when we come back we will hear more from Mary about sampling in quantitative methods.

[*Music plays in background*]

**KL:** The research in action podcast is brought to you by Oregon State University Ecampus, a national ranked leader in online education. Oregon State Ecampus is renowned for its innovation and it recently won three national rewards for the development of a 3D virtual microscope that gives students everywhere increased assess to learning. Learn more at ecampus.oregonstate.edu/microscope

# Segment 2:

**ME:** So next, I would like to talk about the issue of sampling in quantitative methodologies. So what should we know about selecting samples?

**MK:** So what might seem to be the gold standard is that we select a random sample of participants who represent the population that’s of interest to us. So for example, a lot of political polls are based on reasonably random samples of the population and those give us a pretty good index of what people in general think about that an issue. But when you think about who can come to the laboratory or who is available for psychological research, those samples are pretty hard to find. So if you think about drawing a random sample of everybody in your city and then calling them or contacting them by email and getting them to come to the lab chances are they are not going to be willing to do that. So, you are not going to be able to really represent the population in the way we think about with national samples.

So psychology, for example, tends to use restricted samples. And most often those restricted samples are convenient samples, and most often those are college students. So there is a change now to what’s called crowd sourcing, which I will get back to in a minute. But let’s just start with this convenient sample of college students and what those samples look like. So one criticism that is often leveled in psychology is its reliance on what’s called college sophomores. The science of college sophomores. And that is in fact true that in psychological research college sophomores are much too over represented. So for example P.J. Henry did a survey of research published in social psychology journals between 1990 and 2005. He focused specifically on the topic of prejudice. And he found even in journals with a very applied focus where you might expect to see less reliance on college students about 70% of the studies where based on a college student samples. [**ME:** Wow that’s a really high percentage.] It is a very high percentage.

There is another well-known study by Arnett who looked at country of origin in psychological research and he found even though people from the United States only make up five percent of the world’s population. People from the United States are about 68 percent of samples used in psychological research. The next most common source are samples from Europe which is another 13 percent, and there are only tiny representations from Asia, Latin American and Africa and the Middle East. So there is a huge biased towards representing people from the U.S. and Western Europe in psychological research.

There is also another really well known study by Henrich and colleagues who identified what are known as weird samples. And what they mean by that is that in psychology we over-represent people who come from western, educated, industrialized, rich and democratic countries. So we are pretty clear about the source of the biased in our samples.

I also want to just mention that even within those samples there is another bias that’s underneath that and that bias affects who volunteers for psychological research and who doesn’t. So Rosenthal, for example, found that people who volunteered for psychological research tend to be more educated, tend to come from higher socioeconomic statuses, tend to be a little less authoritarian female and younger. And so we have another kinds of bias that aren’t just related to demographic variables. That are a part of the constrictive band of people we have in psychological research.

**ME**: So that relates to what we were originally talking about is generalizing. Right?

**MK**: Exactly. So that asks the question of whether the results that are based on those samples really represent the world in a much broader way. Or even represent for example everybody from the state of Indiana, they don’t really. And we are pretty clear about that. And I will get to a little bit later why that matters and what the question comes down to when we think about that generalizability.

**ME**: Okay; great.

**MK**: So let me say also a little bit about um what’s called crowd sourcing. So one thing that is much more common now than in the past is that we can collect data using the internet. And one way we can get data on the internet are from crowd sourcing sites. And the most well-known site is Amazon’s Mechanical Turk. So mechanical turk is a site where researchers can go and ask people who are mechanical turk workers to complete their researcher. And so if you have an online study you can put it out there on Amazon’s turk, and people will sign up for your study in exchange for payment. So this is very much changed how psychologists collect data and it’s a very nice change because if you want people who are not from the United States you can get them through mturk samples. And we can also get people who are not college students and who represent the population more broadly.

**ME**: Now Mary that’s one place and I believe there are other places like Qualtrics also allows you to get paid participates as well. Is that right? Are you aware of that?

**MK**: That’s correct. There are other ones also. When we think about those crowd-sourcing samples it’s very important to keep in mind that they are also convenient samples. So we don’t want to give the impression that mtruk samples or crowd sources samples are random samples of the population. There is still the convenience sample and they have bias built into them. For example we know from research by Kraut and colleagues that internet samples over represent whites, young people and parents. And Chris Fraley also notes that they underrepresent the poor and working class.

**ME:** Interesting. So these samples are at face value. Not that much different even though people many think they are.

**MK**: That’s right. [**ME:** Interesting] There is another interesting bias build into crowd sourcing data that’s kind of hard to recognize, but also important. And that’s the fact that the people who are workers for example Amazon mturk are people who stay in the pool for a long time and probably have participated in other psychological research. So it’s pretty possible that any workers that you get in your study have already completed a study that is at least come what similar to yours. And they might have been debriefed about that study which means that they have learned about the reason for this study and maybe have even been told about manipulations that won’t create a different source of bias.

**ME**: So interesting in that way they are not that much different than the college sophomores.

**MK**: In some ways they are not. And also again I want to emphasis that you do get more breath, but it’s just in some ways you are substituting one kind of bias for another kind of bias. And it’s always important to keep in mind what that bias is. Another problem with internet research is that people might represent themselves inaccurately. So, keep in mind people who are mturk workers get paid. And those workers are motivated to participate in a lot of studies because they earn more money. So, if you are trying to screen people for a particular characteristic. For example some sexual orientation you can’t be 100 percent sure people don’t say that they are gay or lesbian or transgender so they can complete the study. And there is no way you can find out for sure whether they have misrepresented themselves. And there are ways around this, but it is a little bit complicated and you have to be mindful of those techniques. So that you can look for that kind of bias in your own sample.

**ME**: Those are really interesting comparisons between you know the traditional samples that we think about, especially in psychology, and then this new opportunity to read a broader set of samples. So thanks for sharing the research and what you understand from those. So that was great.

So we are going to take a brief break and when we come back we will hear more from Mary about her own work.

[*Music plays*]

# Segment 3:

**ME:** So finally, Mary, you have experience with a specific methodology called meta-analysis that I would like to talk about. So can you briefly describe what meta-analysis entails?

**MK:** Yes, so a meta-analysis is often described as a study of studies. So in a typical research study an individual respondent provides a data point. In a meta-analysis a study results provides a data point. So you can think of a meta-analysis as a study of studies. And its useful because we can then take a data point from a broad range of studies and look at whether a pattern we are interested in holds across all of those studies. Or we can also break it down and look at what our moderated variables or the variables that show when an effect holds and when it doesn’t hold.

So meta-analysis are useful for studies that can provide us with a numerical comparison and can give us a data point that looks at a broad question. So when you do a meta-analysis you try to locate all the studies you can find that examine a particular hypothesis. And then you set criteria for whether the study does need to be include or does not need to be included. And once you identify those studies, you compute what is called an effective size. That allows you to make a comparison between the groups that you are interested in and look at whether the effective that you think is there holds across the series of studies. So meta-analyses are an objective ways to look across a set of findings and draw conclusions about whether an effect is there or not there. One qualification is that the interpretation is still limited by the samples that are in your study. So this doesn’t result in a random sample of studies. It’s still representing for example probably likely to over represent college students, but it also allows you to look at whether for example the results differ between college students and non-college students. So you can ask that as what’s called a moderator question.

**ME**: One of the powers I guess of meta-analysis is that we get to take a group of studies on a topic that may have differing results. Right? And find out what they mean together. And I think that that’s an interesting way to think about meta-analysis. So, you recently completed a twenty-year replication of a meta-analysis that you had previously conducted. So can you share some of those findings?

**MK**: I can. So in 1996, Bernie Whitely and I published a study that looked at whether there is a sex difference in attitudes towards gays and lesbians. And back in 1996, we did find that there is a sex difference. And so we thought that it would be worth replicating our meta-analysis in current times and see whether that sex difference has gone away. And one reason we thought the sex difference might disappear is that attitudes towards homosexuality in the United States are defiantly changing toward greater acceptance. And if that’s turn you might except when women and men would look at things more similarly now that they did twenty years ago. So back in 1996 we had one-hundred and sixty effect sizes representing about a hundred thousand people. And now we have six hundred and five effect sizes representing about six hundred thousand people. So a much bigger sample [**ME:** Wow! That’s a lot of data]. It is a lot of data. And what we found and we are just fascinated by. In 1996 we found a D statistic of .38 and I realize a lot of people won’t know what a D statistic is. But it basically means a pretty reasonable size difference, always headed in the direction of men being more negative than women. And what we found now 2017, it exactly the same D. It hasn’t changed at all over time. And when we tired to look at year data collection as a moderator of that effect to see if those two things. If that sex difference might be shrinking over time we found no evidence that it is shrinking over time. So well social attitudes are becoming more accepting. Men are still less accepting than women at about the same level of change or level of effect as we saw for back in 1996.

**ME**: So you are saying that that D value is saying that men are more um likely to have negative attitudes towards homosexuality. [**MK**: Correct] Correct okay. And that’s what has not changed in that twenty year period [**MK:** It has not changed.] That is remarkable. [**MK:** You think it would have but it has not.] Yeah. Fascinating, with more studies and going back to those other issues. Probably pretty good samples in those studies.

**MK**: Yes! And what’s nice about our research set is that a lot of the studies that we included asked looked at whether women and men have different attitudes. But that’s not the main part of their study. So that gets us out of one problem with meta-analysis and that’s the issue of whether only those studies that find a difference actually get published. That’s less likely in our sample because most of the research was looking at a different set of questions and so it the sex difference wouldn’t have been a factor of whether it got published or didn’t get published.

**ME**: Ah that’s very interesting. Okay now a meta-analysis is an enormous amount of work. Can you talk a little bit about your process?

**MK**: So, we have been working on this for five years. [**ME:** Wow] And so for each of the studies that we find you have to go in and compute the D statistics that compares women and men’s attitudes. And so that has to be done basically by hand and then someone has to check it. And so that took a long time for us to locate all the studies and then compute the D statistics and then have them checked. And then you have to code every single study for the moderators that you are interested in and looking at. Like we had to code for year the data where collected. And then also you have to compute the statistics um and that. A lot of statistical computation are more of less automated now you can put them into a statistical program like SBSS and kind of get data out pretty quickly. But that is not really true for meta-analyses. So it does take a very long time to get your data and we are just now getting to the stage where we are ready to write up our results.

**ME**: So can you talk about maybe your speculation to why you think this hasn’t changed?

**MK**: Um I don’t really have an answer to that question, honestly.

**ME**: Okay; that’s fine.

**MK**: But let me ask let me answer a different question. [**ME:** Okay] Which is an important qualifier to the general set of findings that I described. So one moderator that is really important is whether we are looking at attitudes towards gay men or whether we are looking at attitudes towards lesbian women. And what happens in our data both in 1996 and today is that men are especially negative toward gay men compared to women. And that when you look at attitudes towards lesbians women and men have very similar attitudes. So one thing we are pretty sure about is that it has to do with gay men and the belief that gay men are seen as feminine. And there is an idea that we have that may other people share that men are taught to reject femininity and gay men are seen as feminine. So, one reason that gay men are rejected by heterosexual men is because of that belief, that gay men are feminine. Women in contrast are not taught to reject femininity after all though most of them do identify as feminine at least um to some extent. And so for women there is not that need to reject gay men because they are seen as feminine.

So we found again in 1996, and today a much, much larger effect size when you look at gay men alone than when you look at lesbian women.

**ME**: That’s really interesting. In that and important to look at that level of complexity.

**MK**: And so I think getting back to your level of what might account for the effect. The belief that femininity is less acceptable than masculinity probably hasn’t changed across time. That might be one reason why men are still negative toward homosexuality than women are, but we don’t a hundred percent know that for sure. We are hoping to look at general attitudes as a moderator we haven’t don’t those analyses yet.

**ME**: Wow that is really, really interesting. So that must have been a big surprise to you [**MK**: It was a surprise] to think you would see significant change and that it did not.

**MK**: I defiantly expected the sex difference to shrink at least a little bit over time. The other thing that I mentioned which we are still working on these analyses. But I can tell you tentatively that it looks like to have enough question when you ask. So, for example if you ask about civil rights should gay men and lesbian women be allowed to marry, for example. Or should they be allowed to adopt children or be in the military. The sex difference is much smaller than if you ask about the morality of homosexual behavior. And that sex difference is much larger. So what probably depends on the question you are asking and where we see in the national survey data the biggest change is usually on questions like whether gay marriage is acceptable. Those questions those some survey data don’t tend to ask about homosexual behavior. Which is something that people still have reservations about in our culture and in many cultures.

**ME**: So thank you for talking with me today, Mary. And thanks to your listeners for joining us for this week’s episode of Research in Action. I am Mary Ellen Dello Stritto. Join us next week for another episode.

[*Music plays in background*]

# Show notes with links to resources mentioned in the episode, a full transcript, and an instructor’s 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 50 degree programs and over 1,000 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.

# Bonus Clip

**KL**: In this bonus clip for episode 116 for the “Research in Action” podcast, Dr. Mary Kite discusses the importance of having a representative sample. Take a listen!

**ME**: So, Mary, can you tell us about the importance of having a representative sample.

**MK**: So, it’s very easy to kind of pick that issue of convenient samples, and draw the conclusion that psychology isn’t telling us very much about the real world or people actual behavior. I just want to talk a little bit about the positives that we get from laboratory research that serve as a counter weight to those criticisms.

So we know from laboratory research that we have a lot of control over what happens in the laboratory. So if we think about a really simple experiment for example for the helping studies that I mentioned at the beginning. Darley and Latané were able to control how many people where present. Whether it was one or five for example. And so they could clearly test the hypothesis that when several people where present helping decreased. And so one strength of laboratory research is that it allows researchers to really focus in on a specific hypothesis. And to test that hypothesis in a way that allows us to say variable one caused variable two. And those goals are important to many psychologists and they are really important if we are looking at something that is very complicated and hard to study in the real world. Or if we are looking at something that has never been studied before. And so it’s really useful to remember it that laboratory research has a lot of strengths that are not really that are really again um you kind of have to set them side by side to question whether the results match the population in a random sample kind of way.

It’s also important to keep in mind that there is a lot of evidence that lab research does generalize to the real world. So I gave you an example earlier on the work of non-syllables that really does tell us a lot of how people remember in everyday life. So we shouldn’t draw the conclusion because laboratory research has some limitations that it doesn’t tell us anything that is important. In fact we know a lot of very important things based on laboratory research.

And another important thing to keep in mind is that if I can conduct research in a natural setting. Let’s say I go out to our organization and I particularly things that are happening there. That research isn’t necessarily generalized to another organization its self. So you can’t draw the conclusion that things that are conducted out in a natural setting would automatically generalize to another setting.

So the most important piece in this discussion is to remember that we have to replicate research so that we can demonstrate generalizability. That every single study has strengths and limitations and what psychologists do and what all good research does is replicate as many different times and ways we think are important. So we can show whether the pattern does in fact apply across settings. Or that we can identify what are called the boundary conditions. Or where the research applies and where is doesn’t. So again it’s just it’s important to keep in mind that we need a really good reason to think that results won’t generalize before we come to the conclusion that laboratory research doesn’t have anything to do with real life.

**ME**: Great. And I like how you brought up all those things together the sampling um and the generalizability and the validity. That’s a nice way to kind of end that topic.

[*Music plays in background*]

**KL**: You have just heard a bonus clip from episode 116 of the “Research in Action” podcast with Dr. Mary Kite discussing the importance of having a representative sample. Thanks for listening.