Multiple Operationalism

at its core critical multiplism is a theoretical framework for thinking about how we do our research I know that some of you heard research in theory in the same sentence and you kind of tuned right out but try to bear with me here for a minute okay virtually all of you come from organizations that would like to use data to make predictions to better serve your clients or your stakeholders whether that be students customers program participants whomever when we are making predictions the biggest hurdle we need to overcome with overcome is the hurdle of generalizability okay that is will this predictive model accurately predict cases outside of the observations that I used to build it now in data science there are procedures that get used called cross-validation that try to get at this you if you train a model on a just a subset of the entire data set and then try to predict the rest how off is the model on average in essence that's how all variations of the procedure work um and it's a reasonable a reasonable approach to tackle one specific type of generalization but these procedures are going to be limited to the data that we are already collecting or have already collected there are steps that we can actually take in the data Gathering step in the formulation of the study of the specific design that we can also use to sort of increase our confidence in our model's ability to predict um that is it's generalizability I mean this is especially going to True going to be true when the things that we are interested are complex and not best represented by a single metric and that is most of what I'm going to be talking about in fact today

so many if not all of you have heard the phrase and operational definition that is to say we Define the thing that we are studying by how we measure it sometimes this will be excellent other times it's gonna miss the mark okay let's take the example of student engagement okay if we look at a theoretical definition of student engagement it usually ends up somewhere in the ballpark of a multi-dimensional construct comprised of cognitive behavioral and emotional components that describes the energy purpose and durability of students invest in their academic work and their participation in school now how do I go about measuring it

well in terms of the practice there are a few different ways that we see student engagement measure we see student self-reforms we see teacher ratings interview and focus groups behavioral observations and administrative data okay each of these methods has its individual pros and cons but I want to focus in here on what is by far the easiest way of gathering data and that's going to be this administrative data okay things like attendance graduation rates problem behaviors and course enrollment can all be considered indicators of this thing that we call student engagement but if I Define student engagement as the students attendance rate or the number of courses a student enrolls in have I really captured student engagement

well no right I've taken this multi-dimensional thing and I've reduced it to just one aspect we we have we have a metonymy right who's sort of confused the part for the whole attendance is not in fact

engagement but sort of a a representation or a manifestation of student engagement

so what's the solution let's say you're working for a district and you don't have the option for one of the other means of gathering student engagement you don't have time to do a student survey you don't have the resources necessary to do an interview or a focus group or behavioral observations you really just have the administrative data that you have on hand so are you stuck just picking one of these and running with it like are you just stuck with attendance no you can take some of these other things that you see happen and you can use those two use of in addition to you look to see how your predictors or maybe an intervention relates to attendance and problem behaviors and the number of courses enrolled and whatever else you have you use multiple operational definitions you look for consistency for parallel results and then you say I am going to Define student engagement as the thing or the construct that underlies all of these indicators okay I'm gonna I want to take it a step back and get a little bit abstract here a little bit more abstract here you have this construct that you're interested in in this case student engagement that construct manifests itself in different ways as attendance as graduation rates as problem behaviors as course enrollment okay what you are trying to do is you're trying to use these manifestations these things that you can actually measure these things that you can see to get a peek at sort of the shape of this construct that's behind them okay the thing that's driving these differences that you're seeing now you can do this in one of a few ways okay you can you can run either the same or very similar statistical models where all you do is sort of just switch out

that outcome or that dependent variable you can create a new metric if you want to sort of combine all of the items together you can make something like an index um sort of considering all the different variables together now there are some things that we have to think about don't just go making an index just to make an index we do have to think about how we are going to weight the individual variables how we're going to combine them mathematically things of that sort so there are a lot of problems uh sorry a lot of limitations and a lot of considerations we have to make if we are going to do that we may get into that next week we may not not sure yet um but you can also uh if you happen to know some fancy statistics you can do those as well there are procedures called latent variable modeling or measurement models and structural equations modeling that'll sort of all get at this idea of using these uh indicators or these manifest variables these manifestations and mathematically show what the relationship is among those variables and potentially other things as well okay and the way that they are ordered here is probably the most is probably the order of the most common way you'll see these represented in applied literatures um you tend to see folks using multiple similar dependent variables and sort of the the go-to creating a new index will be sort of the second and then using these higher level statistical models will be usually around the around the last option in applied work in theoretical work or basic research it's sort of flipped but that's the general idea okay

but this diversification of operations okay it's not just going to be unique to these types of data administrative data okay it's also going to apply to things like surveys and interviews and panel groups and all that

kind of thing but surveys sort of have a built-in Safeguard to some extent um they can sort of build in this idea of multiple operations from the test construction standpoint you can ask multiple questions pertaining to the same construct you can essentially getting multiple measures where each question is itself a measure of that construct you can measure multiple constructs within the same survey and things like that like if we take for example the big five openness to experience neuroticism conscientiousness agreeableness and extroversion you see here that each of these statements here is a specific manifestation or specific indicator of the underlying domain right we have eight here for extra Version 9 here for agreeableness nine here for conscientiousness and eight here for neuroticism and 10 here for openness right each of these items gets at the Domain in a slightly different way so we can end up with a single survey that asks a series of questions and yields us these sort of five sub scales or we can if we have some reason to combine these into one giant score we could have one giant score it really just depends on how these specific survey is constructed okay um this is particularly going to be true for the well-validated scales they'll sort of have this explicit structure built in or even implicit structure built in where they have sort of sampled across the different uh what they believe are theoretically relevant domains we want to approach the construct the thing that we are interested in from with slightly different metrics for the same reason that we said we want to use qualitative and quantitative methods together each metric is going to have its own unique biases and when we take them together collectively they should be able to make up for each other's weaknesses to some extent we want to be able to generalize from what we find from one set of metrics of a construct to another or to another set of metrics we want to say hey this pattern that I found here with these particular indicators I'm likely to see a similar pattern or a similar result into this other metric or these other metrics that are still related to this construct called student engagement and the reason we do this for a few reasons one of the major ones is that we don't want to overly rely on one single metric okay over Reliance on one metric to determine success can result in that metric becoming the end becoming the end itself if you think of something like the SAT it's supposed to measure aptitude student aptitude or student achievement but when it ends up being the chief criteria for things like scholarships or admissions to college schools begin to teach to the test they sacrifice other goals of Education in favor of teaching to a test if we think back to Campbell's law the more any quantitative indicator is used for social decision making the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it intended to monitor sat doesn't measure aptitude it measures your ability to take the SAT and we don't want that for our metrics for our outcomes particularly those that are essential for us understanding how our organization is functioning and how we might be able to better our organizations