# Multilevel models to study intersectionality

## Transcript MAIHDA - Limitations and Extensions (video 3)

Part of the resource: <https://www.ncrm.ac.uk/resources/online/all/?id=20849>

So, we’ve seen in the previous video some of the advantages of the MAIHDA approach in comparison to more conventional approaches involving interactions or dummy variables. But MAIHDA is not a panacea. It has limitations and we need to be wary of those limitations and some of the conceptual challenges that we face when we’re setting up these models. So, we’re going to talk about some of those now, as well as talking about some extensions, potential extensions of the MAIHDA approach.

So, the first thing to say about MAIHDA is that it’s an explicitly exploratory descriptive approach. Really what we’re doing here is finding a really nice, robust way of estimating the means of strata in a particular outcome and finding the differences in those means between different groups. So, we’re generally not testing a particular hypothesis in a kind of confirmatory way, rather we’re looking at finding estimates of these different strata means and exploring those differences and understanding those differences. Now, that’s not a problem. Exploratory research can be robust. Exploratory research is an important part of the research canon, but it’s important to understand this is generally the approach that this method takes.

Now, probably the most common question we get asked by people that are interested in using MAIHDA is around sample size, so how big a sample do I need in my dataset, how many strata should I have, how many individuals is it okay to have within each strata? And the answer to these questions, I’m afraid, is it depends a bit. But it’s worth kind of talking through how we might get into these kind of approaches, because broadly we are dividing our population into subgroups and as a result of those divisions some of those subgroups are going to be relatively small, in particular some of the subgroups that we’re particularly interested in may be relatively small, often unusual combinations of identities that have a relatively small sample size might be groups that face particular disadvantage in society. But fundamentally, if we’re dividing society up in this way we’re going to end up with relatively small numbers in some groups, and we’re going to have some strata where there’s no-one in them, some combinations of identity where no-one in our sample matches that. And that requires some careful thought, okay? It requires us to think about how many variables we should include to define our strata because we could think of lots, right? We could think of sex, sexuality, age, ethnicity, disability status, social class, education, I could go on and on, but if I include all those variables in defining my strata, I’m going to end up with a lot of strata and therefore fewer people within each one.

Similarly in terms of how do I group, how many groups for each variable do I need? So, taking an example of ethnicity, I could divide ethnicity into two groups, right? Majority ethnic group and minoritised ethnic group. But that’s potentially problematic, that’s potentially assuming that all of those ethnic groups that we’re categorising as minoritised experience that disadvantage in the same way. That may well not be true. At the other end of the scale we could have potentially 20-something different categories, thinking about the different ethnicity categories that, for instance, the Office for National Statistics in the UK produce. That’s probably going to be too many to be able to find anything useful because taking some of those smaller categories when we’re dividing those into other variables as well, we’re going to have very small numbers for a lot of those.

So, we need to come up with an answer that balances these things, and it’s going to depend on… the answer that we go with is going to depend on a few different things. It’ll depend on how big our dataset is, so if our dataset is really large we can probably divide things up a bit more, if it’s relatively small that might be a bit more tricky. But it will also depend on the nature of the question that we’re being asked. So, it might be that for some outcomes we’re particularly interested in particular intersectional groups or particular categories of variables. If we’re interested in a particular ethnicity, for instance, it doesn’t make sense to combine that ethnicity in with all other minoritised groups, if we’re interested in a particular age group it doesn’t make sense to combine that in with other age groups. So, it really depends to some extent on the question that’s being asked and the nature of the inequalities that we want to find out about. We might want to include a particular variable in our modelling for one outcome, but for another we’re less interested in that one and might be interested in another.

Fundamentally, this all comes down to theory and theorising what we’re trying to find out and the theory that underlies what we’re trying to find out. Taking an approach that is theory-based would give us insight into what’s important and therefore we need to include and what complexity we really want to look at, and also looking at thinking about what actually for this particular study we’re perhaps less interested in. And it’s worth saying that if we choose not to include a particular inequality or not to include a particular variable or a particular categorical division, that doesn’t mean we’re saying that that division is unimportant, it’s just that it’s not something that we’re looking at in this particular analysis.

More broadly there are some conceptual challenges that we need to work through when we’re doing a MAIHDA analysis or anything that we’re kind of thinking of as intersectional. Firstly, as we mentioned in the first video, intersectionality is all about societal power to structures, the idea that societal oppression, discrimination, etc., produce a lot of the inequalities that exist in society. That won’t always be true. There are inequalities that exist in society for other reasons, and that’s okay as well. There’s nothing wrong with us doing that and thinking in that way. And it’s worth being aware that in any analysis there will be differences in what’s being produced here. The MAIHDA method doesn’t tell us anything about the process that produces the inequalities. That’s all down to the theory that’s underlying this and the interpretations that we might make afterwards.

So, we want to be a bit careful here. Both are claiming that we’re talking about intersectional approach when actually we’re interpreting things in a way that’s not particularly intersectional, but also we want to be careful of not assuming an intersectional interpretation when maybe it might not be always be justified.

We want to think about what we mean when we say intersectionality because I think a lot of people think intersectionality is kind of a synonym of multiplicative or a synonym of interaction, that is if you found that there were no multiplicative effects, that disproves intersectionality, and I don’t think that’s the case. I think that intersectionality is the approach that one takes, so a MAIHDA model might take an intersectional approach, and it’s the approach that is intersectional rather than us trying to prove or disprove intersectionality one way or another. So, I think we can find intersectional inequalities within additive inequalities and I don’t think, from my perspective, that results have to be multiplicative in order to be called intersectional.

The third thing to think about in this area is how are we going to tell a story about the results that we find, and this is going to be particularly tricky if we have a lot of strata. So, if we have, say, 200 strata, it’s going to be very difficult for us to present those inequalities all together, and it’s also going to be very difficult to kind of tell a story that sums up all of the inequalities that are present across those 200 strata. So, we need to think carefully about when we have that many results, how do we focus our results such that a reader can understand the story of the paper that we’re writing? And we need to be open about the way in which we’re doing that because, as we’ll see in the next video, the way in which we plot our results can potentially affect the story that those plots end up telling us. So, we want to be open about that. We want to understand the processes and the biases that we’re putting into what we’re doing.

And fundamentally, MAIHDA isn’t in itself an intersectional method. MAIHDA is an approach that allows us to robustly estimate means of particular subgroups and identify the proportion of variants that is additive or multiplicative. It doesn’t tell us anything about the kind of social power side of intersectionality, and it doesn’t necessarily have to be used in an intersectional way, so we could use a MAIHDA method to find differences between subgroups based on all sorts of different variables that are nothing particularly to do with identity or social justice or anything else, and we could use that method to understand those in a really useful, meaningful way. That’s a MAIHDA method but I would say it’s not an intersectional method and so we want to be careful about what we are and aren’t doing with any study that we run.

So, just a few things about where we could potentially extend this method. So, multilevel models are highly flexible and MAIHDA is just a fairly multilevel model, so we can extend MAIHDA in all sorts of ways that we can extend multilevel models, so we can add control variables, for instance. These might be variables that we’re not including in our strata but that we do want to control for so that they’re not accounted for in the inequalities. So, for instance, we might have a MAIHDA model where our outcome is school attainment, but we control for past attainment so that we can kind of identify inequalities in progress between two timepoints in terms of school attainment rather than in absolute attainment at a particular timepoint.

It might be that we want to include some two-way interactions. So, it might be that we run our model 2 of MAIHDA and we find that there is some multiplicative variants, and that could be produced in two ways. It could be produced by particular intersectional groups behaving in particular ways, or it could be produced by something like a two-way interaction where a particular group of those intersections are just different from some of the others. So, we could explore that multiplicative difference in a bit more detail if we wanted to. We could potentially add some random slopes, so as well as looking at how inequalities exist in the mean of an outcome, we could also look at how the inequalities exist in the effect of a particular variable on an outcome. So, this has the potential to be particularly useful for things like policy evaluation, so we might have a policy which, on average, seems to work but a MAIHDA random slopes approach could identify whether that policy works for all groups in the same way or whether it works particularly for some groups and less well for others.

We could also add some levels, so we could do a kind of geographical analysis that incorporates a geographical level as well as our strata level in a cross-classified model, and we could have a longitudinal analysis, where we have observations nested in individuals and individuals nested in strata, and that would allow us to not just see how individuals’ means change and are different for different strata, but also how the kind of life course trajectories change. And there are different… and we’ve got some examples of all of those different sorts of model and you can have a look at those further if you’re interested.

National Centre for Research Methods (NCRM)  
Social Sciences  
Murray Building (Bldg 58)  
University of Southampton  
Southampton SO17 1BJ  
United Kingdom

**Web** www.ncrm.ac.uk   
**Email** info@ncrm.ac.uk  
**Tel** +44 23 8059 4539  
**Twitter** @NCRMUK