# Introduction to Mediation and Moderation

## Video 3 transcript

<https://www.ncrm.ac.uk/resources/online/all/?id=20839>

Oliver Perra: Hello. I’m Oliver Perra, and this is the third part of my introduction to mediation and moderation. So, in presentations one and two I’ve talked about mediation, but I had also introduced moderation, saying that moderator is a variable that qualifies the association between a predictor and an outcome. What this means is the strength of the association or even the sign of the association between the predictor and outcome changes depending on the values of the moderator.

Here I present a fictional example. Social capital refers to personal relationships, social networks, support, support, trust, etc., and social capital may be higher or lower. And here the relation between being exposed to traumatic events and the onset of major depressive disorder may depend on the availability of social capital so the people that are exposed to traumatic events may be less likely to develop major depression if they have a higher level of social capital, so the arrow that links exposure to traumatic events and major depression will be weaker for people who have higher social capital. This is an example of moderation, where the strength of the association between, say, an exposure to traumatic events and an outcome depends on a third variable, the moderator, in this case at the level of social capital.ere

[0:01:50]

People are familiar with factorial ANOVA. I have already dealt with tests of moderation. In fact, interactions between predictors in factorial ANOVA are testing moderation of the facts of a predictor on an outcome. In fact, mathematically, factorial ANOVA is identical to the regression procedures I will illustrate, but the advantage of using regression methods or procedures in test, in describing and testing moderation models is that these regression approaches are more general and flexible and allow to include, test and represent moderation by different types of variables. So, I’ll start with a standard regression model where the value of an outcome Y for individual I is the results of an intercept A which represents the predictive value of Y when the other variables are equal to zero, then we have a parameter b1 that represents the change in Y associated with a one-unit change in predictor X and a parameter b2 that represents the predictive change in Y associated with a one-unit increase in variable mode. Then we have an error term eY, which represents the variation in the Y scores that is not accounted by the predictors. This variation is supposed to be normally distributed around mean zero and the variance we estimate.

Here, since I consider the predictive values of Y in this example I’ve just put here, I’m not representing the error variance in the estimation. In this simple example, X takes two values, 1 and 5, and Mod takes two values, 1 and 2. And using the equation above we can calculate the predictive values of Y for each row. And if we plotted these predictive values, putting X on the horizontal axis and the outcome Y on the vertical axis and use different lines for the different values of Mod, we can see that these lines are always going to be parallel. This is because the effect of the predictors are independent. Any change in Y associated with one unit of X is not dependent on the values of Mod, no matter what the values of Mod and one-unit change in X will result in a change in Y that is equivalent to the coefficient b1, in this case 0.05 points. And on the other hand, the changes in Y associated with increasing values of Mod are not dependent on X.

If, however, we assume that the effect of X on Y depends on the value of the moderator Mod, we have to add a term that represents changes in Y for different combinations of X and Mod. Thus, we assume that there is a third coefficient, b3, that represents how much the values of Y will vary as a function of different combinations of X and Mod. The different combinations of X and Mod are obtained by multiplying the values of the two variables, and see the terms that I have highlighted in red in the equation here. So, now, when we estimate the predicted values of Y for different individual values of X and Mod in the rows, we will see that these vary depending on the combinations of these different predictors. And, in fact, if we plotted those values, those predicted values, so again we have on the horizontal axis the values of X, on the vertical axis the value of Y, and we have two lines representing the different values of the moderating variable Mod, we can see that those lines are no longer parallel, but the way in which the two lines diverge or converge depends on the sign and value of the coefficient b3 that represents the combination of the different values of the two predictors.

Here I have used a dataset that was used by McElreath in Statical Rethinking, a book that I highly recommend, and you can follow this example in the material provided with this module. The data here represents information on countries’ log of GDP, gross domestic product per capita, in the year 2000. These are from different continents, whereas the variable ruggedness is an index of the topographic diversity of the landscape. In most countries, higher level of ruggedness is associated with lower GDP, most likely because transport is more challenging or expensive in rugged terrain, and this can hinder access to markets. However, the association between ruggedness and GDP is positive in African countries. And, as I said, you can follow the example using the R script in the online material. Here the log of the GDP is standardised so that the variable in the dataset represents the ratio of the log GDP, so a country with value 1 has GDP equal to the world average. Ruggedness is also standardised so that it varies from zero, a completely flat country, to 1, the most rugged. Having loaded the dataset, I created a dummy variable to indicate if a country is in Africa or not, and I also create an interaction term that represents the different combination of values of the predictor ruggedness and the other covariate (inaudible 0:08:22) for countries being in Africa. I then ran the regression, as you can see here in the grey shaded area. The outcome standardised log of GDP is regressed on ruggedness, the dummy for African country, and the interaction. And you can see the results where the coefficient for the interaction is significantly different from zero. So, we assert that supposed moderation effect is not negligible.

[0:08:58]

How to interpret these parameters? Remember the question for this model, so the coefficient for the interaction is approximately 0.29, which represents the change in the outcome associated with different combination of the predictor X, ruggedness, and the moderator being in Africa. So, this represents the change in the slope of the line that links predictor and outcome for values of the moderator. So, since in this case we only have two values of the moderator, this means that the line linking ruggedness and GDP will tilt 0.29 units higher with increasing values of ruggedness for African countries. However, the interpretation of this parameter cannot be made separately from the other parameters. We are talking about contextual effects here, and indeed saying that the slope of the association between ruggedness and GDP is higher for African countries than countries in other continents does not mean much if we do not consider this with the other effects. The other effects, in fact, tell us that ruggedness is negatively associated with GDP, so more rugged countries tend to have lower GDP, and African countries tend to have lower GDP. So, in the context of the other effects, the coefficient b3 indicates that across African countries the reduction of 0.15 units in GDP associated with increasing ruggedness is countered by as much as 0.29, so it’s even reversed to an extent for African countries.

The best way to explain and represent these interaction effects, the effects of moderation, are by plotting the effects in graphs like this one. And in this figure represented as a scatter plot, where each point represents (inaudible 0:11:52), so you can see in the graph that on the horizontal axis there is the standardised ruggedness from 0 to 1, and in the vertical axis is the outcome, the standardised log of GDP. So, the points here represent different countries and the blue points are the countries that are not in Africa, so European, Asian and so on, and the red points are the African countries. In this graph I also use lines that represent the regressions of log of GDP on ruggedness for African and non-African countries. The shaded areas around the lines also represent 90% confidence intervals of the predicted values. So, in using a graph like that illustrates effectively the effects that I was trying to explain by words in the previous slide, and you can see that for non-European countries, increasing ruggedness is associated with lower GDP, but the opposite relation is observed across African countries, where increased ruggedness is actually associated with increasing GDP. And we can interpret this result invoking historical reasons and the legacy of colonialism, more rugged African countries were more difficult to reach and this probably protected them against the exploitation of the resources by colonial powers and also from the slave trade that had severe impact on the wealth of these countries.

In the examples and the exercises provided with this module, you can also see how to create similar graphs, so I refer you to those. Since moderation describes situations in which associations between two variables are different across the values of a third variable, you may wonder would it be not sensible to split the participants in different groups and check how those associations differ across these groups. In the example I’ve made before, wouldn’t it be okay to investigate the association between ruggedness and GDP in African countries first and then in non-African countries? Well, it would be a bad idea to split samples in this way, firstly because other parameters in the models, like variances, are not dependent on the moderators, and by splitting the samples we would have more inaccuracy and error in estimation of these other parameters. And also if we split the sample between African and non-African countries, we would effectively assume that the variance of these outcomes differ among these two groups, but there is nothing to justify this assumption. Another reason why it wouldn’t be a good idea to split samples is that we may want to compare models that include moderation effects against models that treat all covariates as having independent effects, and to run this comparison we need to estimate the models on the same data. And here I put an example of how to run likely ratio test between the two models I’ve described before. So, by using the same sample we can also run these comparisons between models.

[0:15:54]

Here I’m providing an example of moderation where the moderator is a continuous variable, unlike the previous example. So, here I’m using data from a previous example used in previous presentations, where I’m looking at the math scores of adolescents in grade 12, and here I am supposing that the association between math scores in grade 8 and math scores in grade 8, it’s moderated by reading scores in grade 8. Again, this example is provided, the syntax to run this example is provided and examples attached with this module. And here the graph shows a scatter plot where the horizontal axis represents grade 8 math scores and the vertical one grade 12 math scores. Since the moderator here, reading scores, is continuous, the easiest way to represent the interaction is to select some meaningful values in the distribution of the moderator and plot the predictor scores for these values. And in this graph I’ve used a package R called interactions, which automatically selected the values of the moderator that corresponded to one standard deviation below the mean, the mean of the moderator and one standard deviation above the mean, so these lines basically represent association between math scores for the adolescents that are average or below average in reading scores. The shaded areas again represent the confidence intervals of these predictions. And this shows that the association between math scores tends to be stronger for adolescents with worse reading scores.

But this example also provides a chance to illustrate a property of moderation models, which is the symmetry. We may say that reading scores are moderating the strength of the association between math scores, so that association depends on the reading scores. However, in the same way, we may say that grade 8 math scores are moderating the association between grade 8 reading and grade 12 math results. From a statistical point of view and mathematical point of view, the two possibilities are equivalent. There is nothing in the equation that I represent here that says that grade 8 reading scores is the moderator. The interpretation of one construct as a mediator does not come from the model or the analysis, in fact, but it comes from designing studies that give support to this causal interpretation, as well as our substantive knowledge about the issues that can give support to our interpretation of a variable as being a moderator. So, the point is, once again, that statistical models cannot tell us everything we need to know, and the correct interpretation of the results necessitates further reflection and collection of information from different sources.

When running a moderation model, the test of moderation can tell us that interaction term is significant so the strength or the sign of the association between a predictor and the outcome depends on the values of a third variable. However, it’s not always evident at which values of the moderator does the association between predictor X and outcome Y become significantly stronger. And here, for example, I use data from the 1977 census in the USA, and this data will be built into R. This data reports information on income, illiteracy, murder rate, and percentages of people with high school degrees by state. And the model here, the level of income is a function, it’s supposed to be a function of high school grade, illiteracy, murder rate and an interaction between illiteracy and murder rate. And the interaction is significant, and I plotted predictions of this interaction using the package interactions. This shows that the association between literacy and income goes from positive to negative with increased murder rate, and we can also see that when murder rate is average, the association between literacy and income seems very weak. And so we may want to probe the interaction. By probing, we can test if these slopes that are represented here are significantly different from zero, that is if the association between predictor and (inaudible 0:21:13) is different from zero at these different values of the moderating variable.

And here I report the syntax and outputs that are provided by this option in the interactions package. And you can see that none of these slopes is significantly different from zero, although the slope appears to become larger and negative with increasing values of murder. We could change these values of the moderator to include points that correspond to two standard deviation units higher in murder rate or two standard deviation rates lower than the mean murder rate and check if the values of the slope are significantly different from zero. But another option is to use the Johnson Neyman model, and this provides an estimation of the points in which the slope that links predictor and outcome are significantly different across the distribution of the moderator, that is where in the distribution of the moderator does the association between income and illiteracy, in this example, is significantly different from zero. And the output of these methods and the graph shows that when murder rate is higher than 11.74, the association between the predictor and outcome is significantly different from zero and, in this case, is negative. So, the interpretation of the interaction will be that the association between state income and literacy is not significant except in cases where murder rate is higher than 11.74, whereby illiteracy predicts lower income.

[0:23:30

But I also wanted to highlight that you need care when using this type of analysis, and particularly the graphs, because you can see, for example, the graph is showing that the model is suggesting that if murder rate was negative, the slope will become positive, significantly positive. But these values are never observed and indeed they are not even possible, there isn’t a negative murder rate. So, the model applies this computation, but we need to think and interpret them and make sense of them and the model is just slavishly(?) using the map to present something, but we need to be careful of when we are interpreting it.

To conclude, I also wanted to mention that it’s possible to combine mediation and moderation into more complex models. For example, models of moderating mediation are models whereby mediation of an effect, of an independent or variable or a predictor, is supposed to be moderated by a fourth variable, so the mediation of the effect of an independent variable on an outcome may be stronger in some contexts, or depending on some variables than in other contexts, so the mediation effect may be moderated by another variable.

And another approach that is becoming increasingly popular and can also be implemented with packages in R is the causal mediation analysis. I refer particularly to work by Imai and others, but it’s a general approach for the definition, identification and estimation of causal mediation effects. That doesn’t refer to specific statistical models, and it’s therefore quite flexible in accommodating different types of relationships between variables, both parametric and non-parametric models, and different types of mediators and outcomes. So, it’s becoming more popular and it’s based on a counterfactual model of causal inference, so it provides a nice way also to define the causal links between different variables in a model. And I refer to some references if you want to learn more about this approach.

Well, thank you for your attention, and I refer to the material provided with this module again, where you will find the examples I have used in the presentations, as well as exercises that will guide you in applying some simple mediation and moderation models to data and also will show you and allow you to practice creating some of the figures that I have presented (inaudible 0:26:58) this presentation. So, thank you very much. Bye.

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