

Three-way Interaction Example

Say we are interested in the effects of publication pressure, and the motivational focus of researchers on the average quality of the papers they produce. With these three variables we will first build a model with one dependent variable, two predictor variables, and an interaction between the two predictor variables. After that, we will introduce another predictor variable to get a model with a three-way interaction. Plots are by far the easiest way to interpret interactions, so that is what we will mainly use here.

In this (completely made-up) example, average quality papers is our dependent variable. It is some continuous measure based on things like how reproducible the results in the paper are, the information value of the paper, the readability of the paper, etc. The higher the score from 1-100, the higher the average quality of the researcher's papers.

Publication pressure is the amount of pressure the researcher experiences to publish a lot of papers – this could be from demands to get a tenured position, their boss, the requirements to obtain grants, or the demands to get recognition in academia, etc. It is measured with a categorical variable with three levels, 0 low pressure, 1 medium pressure, 2 high pressure. Publication pressure will serve as a predictor variable. Given that it has three categories we will actually use two dummy predictor variables for publication pressure for our analyses, where low pressure is the reference category.

Motivational focus is another categorical predictor variable, with two categories (clearly oversimplified, but it will have to do for this example). Category 0 is 'truth-finding': A researcher in this category is mainly motivated by 'truth-finding', that is, wanting to figure out how the world truly works. Category 1 is 'status': A researcher in this category is mainly focused on obtaining a high 'status', that is, becoming successful in terms of earning a lot of money, getting recognition, and/or obtaining power.

Imagine that in reality, the following scores (depicted in the plot) would be the expected values for the average quality of the papers, for the different levels of publication pressure, and the different motivational foci of the researchers:



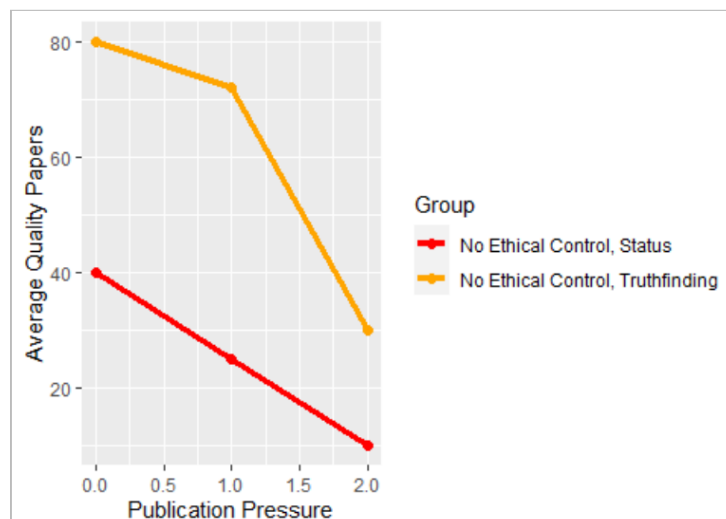
That is in reality, say we have three main effects here on Average quality papers, namely:

- An effect of each of the publication pressure dummies on quality, where the medium pressure group has a lower expected quality than the low pressure group (dummy_pubpresmed), and the higher pressure group is also expected to have a lower quality than the low pressure group (dummy_pubpreshigh).
- An effect of Motivational focus on quality, where we see that overall, the Truth finding group produces on average higher quality papers than the Status group.

There is also an interaction between Publication pressure and Motivational focus (that is, we have a two-way interaction for each publication pressure dummy and motivational focus). This means that the relationship between publication pressure and average paper quality differs depending on the motivational focus of the researcher, as can be seen from the figure above. (Note: We could also flip this around and say that the relationship between motivational focus and average paper quality differs depending on the level of pressure). What we see in the plot is that for both motivational focus groups, increasing publication pressure decreases the average paper quality. However, the difference is in the steepness of this relationship: We see that the truth-finding group seems to try to hold on to the quality, keeping it fairly high, until the pressure becomes really high, after which we see a steep drop in quality. For the status group we see a consistent, linear drop in quality as we increase pressure. In other words, for truth finding group we have a non-linear relationship across the three levels of publication pressure (which is not a problem because we are using two dummies), while for the status group we have a linear relationship across the three levels.

Now, let's add a third categorical variable to understand a potential three-way interaction. A three-way interaction basically indicates that the two-way interactions we have change based on the levels of another variable. Here, we'll introduce the variable 'Ethical Control'. Ethical control is a measure of the level of checks, controls and experience pressure researchers experience to behave and perform research in an ethical way. The variable has two levels, high control (high pressure/many checks to behave ethically) and low control (low pressure/few checks to behave ethically).

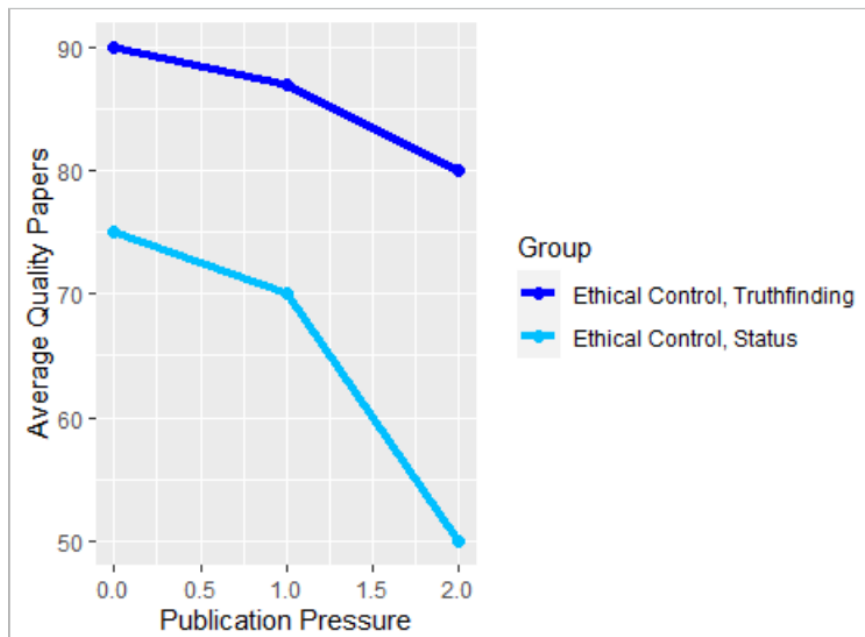
The results for the two-way interaction (plotted in orange/red) we saw before are actually the results for a group that experiences low Ethical Control:



If there is NO three-way interaction between publication pressure, motivational focus, and ethical control, then the plot of the interaction should look essentially the same for the group that

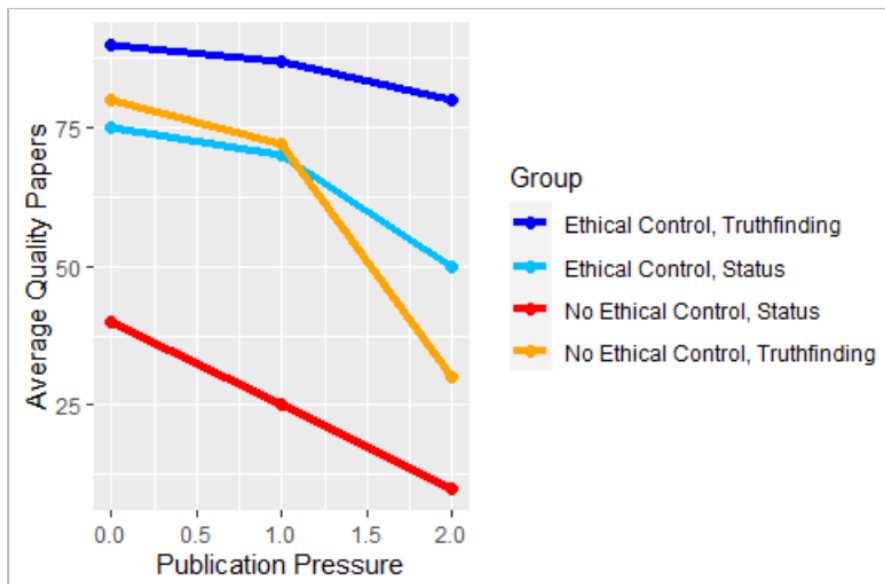
experiences high ethical control. That is, both groups would have a linear effect of pressure on quality of this size for the status group, with the specific non-linear effect for the truth-finding group.

However, in reality, there is a three-way interaction, and hence the plot for the high ethical control group looks different:



In fact, the pattern has essentially reversed for the truth-finding group and the status group: The truth-finding group has an (almost) linear effect of publication pressure on quality. One may say that due to the value that is placed on ethical behavior, the truth finding group manages to not have a steep decrease in quality even when the publication pressure is high. The status group now first has a relatively small decrease in quality when pressure increase to medium, but has a steep drop in quality when the publication pressure becomes high.

Here are the expected values for all the variables in one plot, with different lines for each combination of motivational focus and ethical control:



Here is the regression equation for this model:

$$\text{Av_quality papers} = \text{int} + B1 \cdot \text{pubpressm} + B2 \cdot \text{pubpressl} + B3 \cdot \text{motiv} + B4 \cdot \text{ethictrl} + B5 \cdot \text{pubpressm} \cdot \text{motiv} + B6 \cdot \text{pubpressl} \cdot \text{motiv} + B7 \cdot \text{ethictrl} \cdot \text{motiv} + B8 \cdot \text{ethictrl} \cdot \text{pubpressm} + B9 \cdot \text{ethictrl} \cdot \text{pubpressl} + B10 \cdot \text{pubpressm} \cdot \text{motiv} \cdot \text{ethictrl} + B11 \cdot \text{pubpressl} \cdot \text{ethictrl} \cdot \text{motiv} + \text{error}$$

These are the true (population) values of the coefficients that I made up to make the plots:

int=80

B1=-8

B2=-50

B3=-40

B4=10

B5=-7

B6=20

B7=25

B8=5

B9=40

B10=5

B11=-35

By writing out the regression equation for each of the possible groups, and filling in the scores for the predictor variables and the regression coefficients, you get the expected values for each group that you see in the plots I made. This is a lot of work, but can be helpful to see what particular interaction terms do:

Group: Low pressure (0 both dummies), truth finding 0, no control 0.

$$\begin{aligned} \text{Expected value } Av_qual_papers &= \text{int} + B1*0 + B2*0 + B3*0 + B4*0 \\ &+ B5*0 + B6*0 + B7*0 + B8*0 + B9*0 \\ &+ B10*0 + B11*0 = \text{int} \end{aligned}$$

Group: Medium pressure (1,0), truthfinding 0, no control 0.

$$\begin{aligned} \text{Expected value } Av_qual_papers &= \text{int} + B1*1 + B2*0 + B3*0 + B4*0 \\ &+ B5*1*0 + B6*0*0 + B7*0*0 + B8*0*1 + B9*0*0 \\ &+ B10*1*0*0 + B11*0*0*0 = \text{int} + B1 \end{aligned}$$

Group: Large pressure (0, 1), truthfinding 0, no control 0.

$$\begin{aligned} \text{Expected value } Av_qual_papers &= \text{int} + B1*0 + B2*1 + B3*0 + B4*0 \\ &+ B5*0*0 + B6*1*0 + B7*0*0 + B8*0*0 + B9*0*1 \\ &+ B10*0*0*0 + B11*1*0*0 = \text{int} + B2 \end{aligned}$$

Group: No pressure 00, status 1, no control 0.

$$\begin{aligned} \text{Expected value } Av_qual_papers &= \text{int} + B1*0 + B2*0 + B3*1 + B4*0 \\ &+ B5*0*1 + B6*0*1 + B7*0*1 + B8*0*0 + B9*0*0 \\ &+ B10*0*1*0 + B11*0*0*1 = \text{int} + B3 \end{aligned}$$

Group: Medium pressure 10, status 1, no control 0.

$$\begin{aligned} \text{Expected value } Av_qual_papers &= \text{int} + B1*1 + B2*0 + B3*1 + B4*0 \\ &+ B5*1*1 + B6*0*1 + B7*0*1 + B8*0*1 + B9*0*0 \\ &+ B10*1*1*0 + B11*0*0*1 = \text{int} + B1 + B3 + B5 \end{aligned}$$

Group: large pressure 01, status 1, no control 0.

$$\begin{aligned} \text{Expected value } Av_qual_papers &= \text{int} + B1*0 + B2*1 + B3*1 + B4*0 \\ &+ B5*0*1 + B6*1*1 + B7*0*1 + B8*0*0 + B9*0*1 \\ &+ B10*0*1*0 + B11*1*0*1 = \text{int} + B2 + B3 + B6 \end{aligned}$$

No pressure 00, truthfinding 0, control 1.

Expected value Av_qual_papers = int + B1*0 + B2*0 + B3*0 + B4*1
 + B5*0*0 + B6*0*0 + B7*1*0 + B8*1*0 + B9*1*0
 + B10*0*0*1 + B11*0*1*0 = **int+B4**

Group: Medium pressure 10, truthfinding 0, control 1.

Expected value Av_qual_papers = int + B1*1 + B2*0 + B3*0 + B4*1
 + B5*1*0 + B6*0*0 + B7*1*0 + B8*1*1 + B9*1*0
 + B10*1*0*1 + B11*0*1*0 = **int+B1+B4+B8**

Group: large pressure 01, truthfinding 0, control 1.

Expected value Av_qual_papers = int + B1*0 + B2*1 + B3*0 + B4*1
 + B5*0*0 + B6*1*0 + B7*1*0 + B8*1*0 + B9*1*1
 + B10*0*0*1 + B11*1*1*0 = **int+B2+B4+B9**

Group: Low pressure 00, status 1, control 1.

Expected value Av_qual_papers = int + B1*0 + B2*0 + B3*1 + B4*1
 + B5*0*1 + B6*0*1 + B7*1*1 + B8*1*0 + B9*1*0
 + B10*0*1*1 + B11*0*1*1 = **int+B3+B4+B7**

Group: Medium pressure 10, status 1, control 1.

Expected value Av_qual_papers = int + B1*1 + B2*0 + B3*1 + B4*1
 + B5*1*1 + B6*0*1 + B7*1*1 + B8*1*1 + B9*1*0
 + B10*1*1*1 + B11*0*1*1 = **int+B1+B3+B4+B5+B7+B8+B10**

Group: large pressure 01, status 1, control 1.

Expected value Av_qual_papers = int + B1*0 + B2*1 + B3*1 + B4*1
 + B5*0*1 + B6*1*1 + B7*1*1 + B8*1*0 + B9*1*1
 + B10*0*1*1 + B11*1*1*1 = **int+B2+B3+B4+B6+B7+B9+B11**

True (population) parameters versus estimated regression coefficients

Note, in this conceptual example of a 3-way interaction, we knew we had a 3-way interaction in reality, and had access to the population model parameters. This makes things simple, because we didn't have to do any estimation, or worry about significance. Of course, in practice we do not have the population parameters (otherwise we wouldn't need to do research), instead we would collect data and analyze the data to estimate the population parameters. Based on the estimated parameters and the data, we would make plots similar to those above. Such plots would still be the main way to interpret (significant) interaction effects, but there would be some uncertainty about the expected values.

For example, say I collected data from about 500 participants, divided equally among the different groups. Here are the results, the estimated coefficients, standard errors, t-values, and p-values, for a random sample of N=500 (estimated in R, not SPSS, but the idea is identical):

Coefficients:

	Estimate	Std. Error	t value	p-value (two-tailed)
(Intercept)	79.7982	0.8427	94.696	< 2e-16 ***
pubpressm	-7.8187	1.1014	-7.099	4.46e-12 ***
pubpressl	-48.9863	1.1165	-43.874	< 2e-16 ***
motiv	-39.5119	1.1220	-35.216	< 2e-16 ***
ethictrl	10.6415	1.1165	9.531	< 2e-16 ***
motiv:ethictrl	24.4966	1.5352	15.957	< 2e-16 ***
pubpressm:motiv	-6.9532	1.5722	-4.422	1.20e-05 ***
pubpressm:ethictrl	5.0737	1.5202	3.337	0.00091 ***
pubpressl:motiv	17.4329	1.5538	11.219	< 2e-16 ***
pubpressl:ethictrl	38.2477	1.5665	24.417	< 2e-16 ***
pubpressm:motiv:ethictrl	3.6682	2.1694	1.691	0.09149 .
pubpressl:motiv:ethictrl	-31.9749	2.1672	-14.754	< 2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Model results:

Residual standard error: 4.914 on 488 degrees of freedom

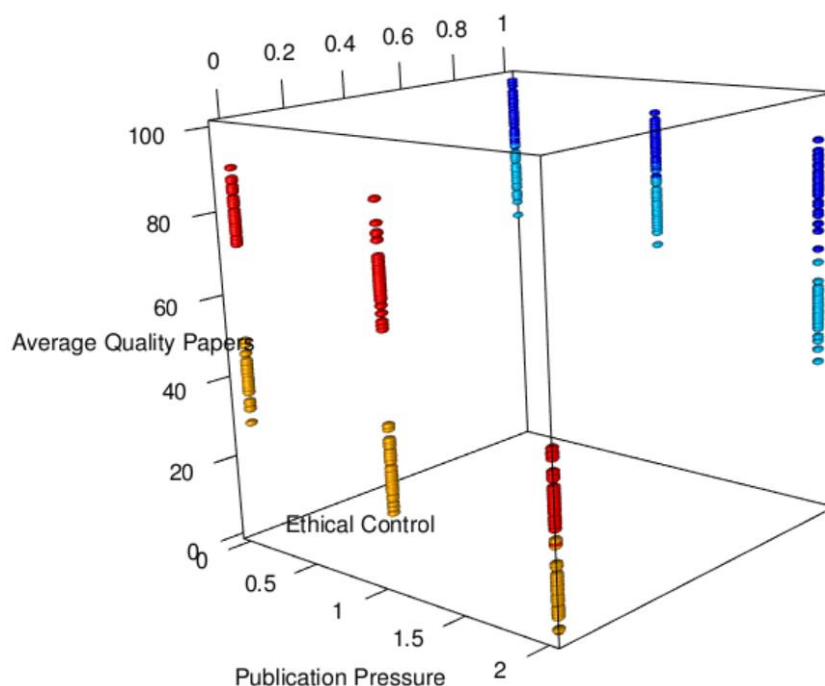
Multiple R-squared: 0.9651, Adjusted R-squared: 0.9643

F-statistic: 1226 on 11 and 488 DF, p-value: < 2.2e-16

Note that the p-values are almost all very small, most of them smaller than 0.001. This is because it is a large dataset and the effects I made up for the population are quite large. The population parameters are estimated back quite well because of this too (compare the estimated coefficients to the true coefficients I presented earlier). We have one type 2 error, because the 3way interaction for the low vs medium publication pressure dummy is non-significant, while there is an effect in reality. The effect is however not very big, so it has not reached significance with this sample. You can also see this from the plots; the largest differences in the interaction effects occur for the high pressure groups (low vs high pressure dummy), the differences among the groups for the difference between low and medium pressure are much smaller.

To help interpret the results you could make 2d plots of the predicted values for each group, as I did earlier for the population values. We practiced this during the course by saving predicted scores in SPSS based on our regression analysis, and then making plots out of those. You could of course also make plots of the observed data instead of the predicted values.

Below I have plotted the simulated data (n=500) for this example in 3d. The axis at the top of the box is for ethical control, the label is a bit misplaced.



In this example the sample size, the effects and explained variance is very large to make everything very clear, of course in practice the uncertainty about all the effects is typically larger.

What about interactions with continuous variables

If you have continuous variables, the upside is that you do not have to mess around with dummy variables. The downside is that it is a bit harder to make plots like the above, where we made nice separate lines for different groups to highlight the effects of the moderating variables. Typically, however, people do make exactly those kind of plots for continuous variables. This is sometimes called 'simple slope analysis'. Basically, what you do is for your continuous predictors, you pick a bunch of values, and for those specific values you make plots like the 2d plots we made before.

Imagine for example that ethical control was a continuous variable ranging from 0 to 10. What people do is then pick two or more values out of that range, and make plots for those values. So, we could pick a score of 2 and of 8 to represent 'little control' and 'lots of control' and make exactly the kinds of 2d plots you saw before. This will help you see the effect of this moderating variable on the 2way interaction effects.