Two-Way ANOVA: Analysis of Data from Studies with Two Non-Repeated Independent Variables

Dr. Tom Pierce
Department of Psychology
Radford University

In one type of study that I do, I measure people's blood pressure. It may not seem like a very glamorous variable, but if you think about it, the haemodynamic properties of… Okay, forget it. Blood pressure is not a very exciting variable for psychologists. But that's what I do. Anyway, I measure people's blood pressure. And if you're going to measure people's blood pressure, you have to do that while they're sitting in a nice big comfy chair.

Let's say that, one at a time, I bring a group of ten research participants into my lab and I tell them that I'm going to measure their pulse rate and blood pressure for ten minutes while they're sitting in the room quietly by themselves. At the end of this ten minute period, a little egg timer is going to go off and, when it does, I tell them I'd like them to turn a piece of paper over that's sitting on a table beside them and to fill out the questionnaire. The questionnaire is a measure of how anxious they feel at that moment.

Then I show them the device I'm going to use to get the readings on pulse rate and blood pressure. The device can take blood pressure automatically as often as I like. In fact, the machine is made for psychologists because to start it, all you have to do is press the GREEN BUTTON. To stop it, you press the RED BUTTON. Not bad, eh. It would have been nice to only have to worry about one button, but I think we can handle this. I get their informed consent and then I leave the room while the machine records their blood pressure every two minutes and they fill out the questionnaire ten minutes later.

For a second group of participants, I bring them in, tell them I'm going to have the machine measure their pulse rate and blood pressure every two minutes for ten minutes and that after ten minutes an egg timer is going to go off and I ask them to fill out the questionnaire. Everything is exactly the same as for the first group of participants, except for these ten people I have an undergraduate research assistant go into the room five minutes into the ten minute period. The assistant walks over to the other side of the room and sits down at a desk and is facing away from the participant. As they walk in the assistant tells the participant that they need to finish filling out some forms from the last participant and that the participant should just try and ignore them. If the participant says anything, the experimenter asks them not to talk because talking will just make their blood pressure go up.

In this second condition, the only thing that's different is that there is a stranger in the room with the participant. This is usually considered to be a more stressful situation than someone just being in the room by themselves. One variable that I'm interested in studying is stress. In this study, I want to know what will happen to people's self-reported scores for anxiety if I change the level of stress they're experiencing.

Let's say that's I bring in a third group of participants. For these people, the instructions are that I'm going to measure their blood pressure for ten minutes and that after ten minutes I'd like them to fill out the questionnaire. For this third group of participants I tell them that half-way through this ten minute period – after five minutes – a research assistant is going to come in the room and ask them to do some mental arithmetic problems for the rest of the ten minute period.

So, I've randomly assigned participants to one of three groups: (1) they're in the room by themselves for ten minutes, (2) they're in the room with a stranger, but they're not interacting with them, and (3) they're doing mental arithmetic. By changing the conditions in this way, I'm trying to provide an experimental manipulation of the amount of stress that people are experiencing. I'm assuming that these conditions represent increasing amounts of stress. The independent variable in this study is Level of Stress and the dependent variable is Self-Reported Anxiety. The logic of this experiment is exactly the same as in the other experiments that we've talked about: you ***change the conditions*** as far as the independent variable goes and then you ***see what happens*** in terms of the dependent variable.

I've got ten people in each of the three levels of stress. If I had to analyze the data right now, I could do a one-way ANOVA, like we've already learned about. But, let's say that there's one addition piece of information that I have about every participant in the study. I know whether each participant is an introvert or an extrovert. It turns out that half of the people in the study are introverts and half are extroverts. In fact, exactly half of the people assigned to each level of stress are introverts and half are extroverts. This mean that of the ten people in the low stress condition (alone for ten minutes), five are introverts and 5 are extroverts. And it's the same way for the moderate stress (stranger in the room) and high stress (mental arithmetic) conditions.

Introversion/Extroversion represent a second independent variable. And I've got every combination of a level of stress that a participant could have been in and whether they were an introvert or an extrovert. So, in effect, what I really have are six different groups because I have six combinations of the three levels of stress and the two levels of Introversion/Extroversion. Below is a table showing the means for the six groups in the study. Possible scores for the measure for anxiety range from one to twelve, with higher scores reflecting higher levels of anxiety.

Introversion/Extroversion (A)

n=5

		Introverts a1	Extroverts a2
	Low Stress b1	*M* = 2	*M* = 2
Level of Stress (B)	Moderate Stress b2	*M* = 5	*M* = 2
	High Stress b3	*M* = 9	*M* = 3

So what questions could you answer using these data? Well, the question we started out with was whether stress has an effect on scores for anxiety. What treatment means could you look at to see if changing the level of stress results in changes in anxiety? To answer this question you'd have to ask what the mean level of anxiety is like for people in the low stress condition, for the people in the moderate stress condition, and for people in the high stress condition. What's the average score for people in the low stress condition? Well, the average for the five introverts under low stress is 2.0 and the average for extroverts under low stress is 2.0. The average of these two means is 2.0. So the average for all ten people in the low stress condition is 2.0. Taking the same approach, the mean score for the ten people in the moderate stress condition is 3.5 and the average score for the ten people in the high stress condition is 6.0. The means are displayed in the margin on the right-hand side of the graph.

Introversion/Extroversion (A)

		Introverts a1	Extroverts a2	
	Low Stress b1	*M* = 2	*M* = 2	*M* = 2.0
Level of Stress (B)	Moderate Stress b2	*M* = 5	*M* = 2	*M* = 3.5
	High Stress b3	*M* = 9	*M* = 3	*M* = 6.0

These mean of 2.0, 3.5, and 6.0 are referred to as the **marginal means** for the effect of stress. These are the means that you get when you average over the effect of introversion/extroversion. In other words, just for moment, we're pretending like we didn't care whether people were introverted or extroverted. We just had ten people giving us scores in each of three levels of stress. From looking at these means, does it seem like there's an effect of stress on anxiety? It does. A little later we'll do the actual statistical test to see if there are statistically significant differences among these means.

In this design, it's possible examine the effect of stress all by itself. When we do this, a statistician would say that we're examining the main effect of stress on anxiety. As a definition, a ***main effect represents the effect of an independent variable on the dependent variable when you average over the effects of a second independent variable***.

In this design, not only can we look at the effect of stress all by itself, we can also look at the effect of introversion/extroversion all by itself. In other words, do introverts look different from extroverts, in terms of anxiety, when you average over the three levels of stress. Now take a look at the means at the bottom of each column. You'll notice that the independent variable Introversion/Extroversion has been assigned the letter A and that the independent variable Stress Level has been assigned the letter B. Just like the situation when we had one independent variable the upper case letter represents the independent variable. A small case letter with a number subscripted beneath it represents a particular level of the independent variable. For example, "a1" represented a group of people. Here, to identify a particular group of people you have to specify the level of both the independent variables A and B.

		Introversion/Extroversion (A)	
		Introverts a1	Extroverts a2
	Low Stress b1	*M* = 2	*M* = 2
Level of Stress (B)	Moderate Stress b2	*M* = 5	*M* = 2
	High Stress b3	*M* = 9	*M* = 3
		M = 5.33	*M* = 2.33

The mean of 5.33 represents the average score for the 15 introverts. The mean of 2.33 represents the mean score for the 15 extroverts in the study. Because these means show up in the bottom margin of the table they're referred to as the ***marginal means for introversion/extroversion***. If we can show that there's a significant difference between these two means, we'll be able to say that there is a significant ***main effect for introversion/extroversion***.

So, one way of looking at the study is that you get two sets of results for the price of one study. When you look for differences among the means in the right-hand margin you're testing the effect of stress, all by itself, on anxiety. When you look for differences among the means in the bottom margin you're testing the effect of introversion/extroversion, all by itself, on anxiety. Two for the price of one sounds like a pretty good deal.

But it turns out that the deal is even better than that. Not only do you get to look at the effects of each independent variable separately; you get an additional piece of information that you couldn't possibly get from a study that had just a single independent variable. Think about it for a second. When you look at the means for the main effect for stress it seems like higher levels of stress result in higher levels of self-reported anxiety. *But is that always true?* Is it true for everybody included in the study? NO! Take a look at the column for the introverts. The means go from 2.0 to 5.0 to 9.0 as the level of stress increases. It certainly looks like there's an effect of stress on anxiety when you only look at the introverts. However, when you look at the means in the column for the extroverts, you get a very different story. Here the means go from 2.0 to 2.0 to 3.0 as the level of stress increases. When you only look at the extroverts – half the participants in the study – the researcher increased the level of stress, but next to nothing happened to the scores! There's <u>no</u> effect of stress when you only look at the extroverts. There <u>is</u> an effect of stress when you only look at the introverts. That's the piece of information you can only get when you look at the independent variables in combination with each other. When you get this type of pattern in the results – ***when the effect of one independent variable on the dependent variable doesn't look the same at every level of a second independent variable – we're able to say that there is an <u>interaction</u> between the two independent variables***.

So there are three major questions that you can answers using this type of design. You can see whether there's an effect of stress, all by itself, on anxiety. That's the main effect of Stress. You can look at the effect of I/E, all by itself, on anxiety. That's the main effect of I/E. And, you can see whether the two independent variables interact with each other.

Okay, so how do you test each of these three questions? By using the same kind of strategy as in the One-Way ANOVA example. In this study there is a certain amount of variability that needs to be accounted for. In other words, if you took the 30 scores for anxiety and had SPSS calculate the sum of squares for those scores, you'd get 126.167. That's all the variability there is that needs to be explained. In the One-Way ANOVA example, we had SPSS calculate the amount of variability that our one possible explanation – our independent variable -- could explain. And we could see how much variability it couldn't explain. The SS Total equaled the SS squares accounted-for plus the SS not-accounted-for. When we had one independent variable, there was only one source of variability accounted-for. It's the same thing here, except that now instead of having one source of variability accounted-for, we've got three: the main effect of A, the main effect of B, and the interaction between A and B (AXB). These sources of variability are orthogonal to each other. In other words, they provide completely independent pieces of information. Having a main effect for A has nothing to do with having a main effect for B or with having an interaction between A and B.

If we (a) calculate the amount of variability in sums of squares units that we can attribute to each of these three sources of variability and then (b) add these three sums of squares up, we'll have all of the variability that's accounted-for in this study. The way a statistician would say it is that we can **partition** the variability that's accounted for (between-groups variability) into three pieces.

$$SS_{Total} = SS_{Between\text{-}Groups} + SS_{Within\text{-}Groups}$$

$$SS_{Total} = (SS_A + SS_B + SS_{AXB}) + SS_{S/AB}$$

This means that there are now three overall effects. When you use the GLM (general linear model) module in SPSS, the program will give you an ANOVA table that provides separate F-tests for each of these three overall effects.

As we talked about, A, B, A X B, and S/AB each have a sum of squared deviations associated them. When you add these sums of squares up you get the sum of squares total. To calculate an observed value for F for each of the three overall effects you have to take a Mean Square for each effect and divide it by the Mean Square Within-Groups. To get these Mean Squares you've got to take each sum of squares and divide it by the appropriate number of degrees of freedom. The ANOVA table below shows the sums of squares for the effects and how these degrees of freedom are calculated.

$df_{AXB} = (a-1)(b-1) = (1)(2) = 2$
$df_A = a-1 = 2 - 1 = 1$
$df_B = b - 1 = 3 - 1 = 2$
$df_{S/AB} = (a)(b)(n-1) = (2)(3)(4) = 24$

Source	SS	df	MS	F	$F_{Critical}$
Stress * I/E	45.00	2			
I/E	67.50	1			
Stress	81.57	2			
Error	32.00	24			
Total	226.17	29			

df_{Total} = (number of participants) - 1

Once you take each sum of squares and divide it by the appropriate number of degrees of freedom you get the Mean Square for each effect. Then, to calculate the F-ratio for each effect, you divide the Mean Square for each effect by the Mean Square for the error term MS S/AB. A researcher could then look up the critical value associated with each effect.

Source	SS	df	MS	F	$F_{Critical}$
Stress * I/E	45.00	2	22.50	16.88	3.40
I/E	67.50	1	67.5	50.63	4.26
Stress	81.57	2	40.83	30.63	3.40
Error	32.00	24	1.33		
Total	226.17	29			

It appears from the resulting ANOVA table that all three overall effects are significant.

One thing to notice about this ANOVA table is that all three F-ratios are based on the same denominator. The mean square for each effect being tested is divided by the $MS_{S/AB}$ of 1.33. This number, 1.33, represents the average of the variances of the scores in each of the six groups. In other words, if you calculated S^2 for each of the six groups and then got the average of these six numbers, you'd end up with 1.33.

A second thing to notice about the numbers in the ANOVA table is that if you add the sums of squares for the main effect for A, the main effect for B, the A X B interaction, and the error term, you get the sum of squares total. Interestingly, if you add up the degrees of freedom for the same four terms you get the total number of degrees of freedom, 29.

Okay, so what do we do now? The significant main effects tell us that (1) there's an effect of stress when you average over the two levels of I/E and that (2) there's an overall effect of I/E when you average over the three levels of stress. But doesn't it seem like these main effects are misleading? I mean, the significant main effect of stress leads you to think that there's always an effect of stress or that there's an effect of stress for everyone. But is that really true? Well…, no. When you look at the means, it seems like there's only an effect of stress when you look at the introverts. And there's no effect of stress when you only look at the extroverts. The main effect is leading you to accept a conclusion that's just not accurate. The presence of a significant interaction between stress and I/E tells you that ***the effect of stress on anxiety is not the same for introverts as it is for extroverts***. If you know that the effect of stress is not the same for introverts as it is for extroverts, WHY WOULD YOU AVERAGE OVER THE TWO LEVELS OF I/E? You wouldn't!

It turns out that when the interaction between the two independent variables is significant, the researcher has to use a great deal of caution when interpreting the main effects. When the interaction is significant it's telling you that the effect of one independent variable is not the same at the various levels of the second independent variable. Knowing this, it follows that the logical course of action would be to look at the effect of stress ***separately*** for introverts and extroverts. In other words, you do a test to see whether there's an effect of stress on anxiety when you only include the data for the introverts. Then, after you've got an answer to that question, you do a test to see if there's an effect of stress on anxiety when you only look at the data for the extroverts. When you do this, you're performing a set of simple effects. ***A simple effect is the effect of one independent variable on the dependent variable at a single level of the second independent variable***. This is in contrast to a main effect, which is the effect of an independent variable when you average over every level of the second independent variable.

Let's say that we're particularly interested in the effects of stress. So we decide to test the simple effects of stress at each level of I/E. We can do this by first using the ***Select Cases option in SPSS to include only the extroverts***. Then do a one-way ANOVA with stress as the independent variable and anxiety as the dependent variable. The output tells us that the sum of squares for this effect is 3.33. This simple effect has 2 degrees of freedom associated with it (because stress still has three levels) and the Mean Square for this simple effect is 1.67. The ANOVA table for the output also gives us an F-ratio and a significance level. The problem with using this F-ratio is that it's ***using the wrong***

denominator. Most statisticians consider the appropriate error term to be the Mean Square S/AB. Essentially, this is the average of all six group variances. The error term in the ANOVA table from SPSS for the simple effect is the average of the group variances of only the three groups that are relevant to this particular simple effect. Statisticians consider the average of six group variances to be more stable and accurate than the average of only three, so they want us to go with the same denominator that we've been using, 1.33. So, to get the correct F-ratio, all we have to do is to take the Mean Square for the simple effect that SPSS gave us (1.67) and divide it by the correct error term (1.33). We end up with an F-ratio of 1.25.

Then all you have to do is to ***use the Select Cases option to include only the introverts*** and then run the one-way ANOVA again. The sum of squares for the simple effect is 123.33. It has 2 degrees of freedom, which gives us a Mean Square for this effect of 61.67. Now put the information from this numerator into the ANOVA table. When you divide this Mean Square by the correct error term of 1.33 you get an F-ratio of 46.37. The expanded ANOVA table that includes the two simple effects is provided below.

Source	SS	df	MS	F	$F_{Critical}$
Stress * I/E (AXB)	45.00	2	22.50	16.88	3.40
I/E (A)	67.50	1	67.5	50.63	4.26
Stress (B)	81.67	2	40.83	30.63	3.40
B at a1	3.33	2	1.67	1.25	3.40
B at a2	123.33	2	61.67	46.37	3.40
Error	32.00	24	1.33		
Total	226.17	29			

From the ANOVA table, you see that there is no significant simple effect of stress on anxiety for extroverts. However, there is a significant simple effect of stress for introverts. This combination of simple effects does its job, which is to explain why we had a significant interaction in the first place.

<u>Variability accounted for by a set of simple effects</u>

So what do these simple effects explain? I mean, the reason we did the simple effects was to explain why we had a significant interaction, so you'd think that they explain the same variability that the interaction does. But if you add the sum of squares for B at a1 (3.33) to the sum of squares for B at a2 (123.33), you get 126.66. The sum of squares for the interaction is only 45.0. It looks like the simple effects are accounting for more variability than there was to account for! It turns out that a set of simple effects is able to explain more than just the interaction between the two independent variables. Take a look at the sums of squares for the overall effects. If you take the sum of squares for the interaction (45.0) add then add the sum of squares for the main effect for stress (81.67) you get 126.66 – the same number we got when we added up the sums of squares for all of the simple effects of Stress at levels of Introversion/Extroversion (126.66).

The simple effects of B at every level of A are able to explain both the AXB interaction and the main effect for B. They explain the interaction because they show us how the effect of B looks different as you go from one level of A to the next. As far as the main effect of B goes, think of it this way. When we tested the simple effect of B at a1, we were testing the effect of B. It's just that we did this using only the 15 people who were extroverts. The test of B at a1 is a test of B for half of the participants in the study. The simple effect of B at a2 is a test of the effect of B for the 15 introverts. It's a test of the effect of B for the other half of the participants in the study. But by the time you've tested the effects of B at a1 and B at a2 you've tested the effects of B using every participant in the study. And that's the same thing that the main effect of B does.

So the simple effects of B at every level of A get two things done. They get the same thing done as the main effect for B and they get the same thing done as the AXB interaction.

Simple comparisons

The simple effect of B at a1 is not significant. Is there anything more to do in terms of investigating the effects of stress on anxiety when you only look at the extroverts? No. The non-significant simple effect tells us that there are no significant differences among the group means of 2.0, 2.0, and 3.0.

The simple effect of A at a2 is significant. What does this tell us? It tells us that there are differences among the group means of 2.0, 5.0, and 9.0. But that's a pretty general piece of information. The simple effect doesn't tell us where these differences are. Just like in a one-way ANOVA we need to do a set of comparisons among the three group means that are relevant to this simple effect.

Let's say that before the investigator had collected the data they'd anticipated that the simple effect of B at a2 would be significant. Because of this they decided to conduct two planned comparisons. They decided to test the prediction that introverts in the high stress condition would have significantly anxiety scores that introverts in the moderate stress condition. They also decided to test the prediction that introverts under moderate or high stress have significantly anxiety scores than introverts under low stress. Because these comparisons regarding the effects on one independent variable are being conducted at a single level of a second independent variable they are referred to as **simple comparisons**.

To get the Mean Squares for these simple comparisons, again, you need to use SPSS to Select Cases to include only the data for introverts. Then using the Contrasts option the researcher could enter coefficients of 0, +1. and –1 to test the first simple comparison and coefficients of 2, -1, -1 to test the second simple comparison. When the researcher runs the one-way ANOVA that includes the test of these two sets of coefficients, SPSS provides observed values for t. If the researcher wanted to reported these comparisons as F tests it seems like all they'd have to do is to square these values for t. But the problem with these values for F is that, just like with the simple effects, they are based on using the incorrect denominator.

The procedure for calculating the Mean Square for each simple comparison is described in the SPSS handout, but basically what you have to do is remember that the Mean Square we need must have been divided by the same incorrect error term that was used in the SPSS output for the simple effect of B at a2. When we multiply the error term in the output for this simple effect by the preliminary value for F we got by squaring the value for t we get a value of 39.99. 39.99 is the correct Mean Square for that simple comparison. When we divide it by the correct error term we get an observed value for F of 30.07, which is greater than the critical value of 4.26. This tells us that introverts under high stress have significantly higher levels of anxiety than introverts under moderate stress

Using the same procedure for the simple comparison of (b1 vs b2 + b3) at a2 we get a Mean Square for the simple comparison of 83.32 and an observed value for F of 62.65. Both simple comparisons are significant. This tells us that introverts under moderate or high stress have significantly higher levels of anxiety than introverts under low stress.

The expanded ANOVA table that includes the two simple comparisons is presented below.

Source	SS	df	MS	F	$F_{Critical}$
Stress * I/E (AXB)	45.00	2	22.50	16.88	3.40
I/E (A)	67.50	1	67.5	50.63	4.26
Stress (B)	81.67	2	40.83	30.63	3.40
B at a1	3.33	2	1.67	1.25	3.40
B at a2	123.33	2	61.67	46.37	3.40
(b2 vs b3) at a2	39.99	1	39.99	30.07	4.26
(b1 vs b2 + b3) at a2	83.32	1	83.32	62.65	4.26
Error	32.00	24	1.33		
Total	226.17	29			

One more thing about this ANOVA table. Notice that when you add up the sums of squares for the two simple comparisons you get 123.33. This is the same number as the sum of squares for the simple effect of B at a2. The two comparisons we did are orthogonal to each other and together they account for all of the variability associated with that simple effect. One way of thinking about it is that the job of a set of simple comparisons is to explain why the simple effect was significant. And that's what this set of simple comparisons does.

<u>Sample paragraph for a results section</u>

If one were to take the information contained in our ANOVA table and write it up the way it would go in an APA-style results section it might look something like this…

A significant interaction between Stress Level and Personality Type was observed, F (2,24) = 16.88, $p < .05$. Significant main effects were observed for both personality type, (1,24) = 50.63, $p < .05$, and stress level, F (2,24) = 30.63, $p < .05$. Simple effects indicated the presence of a significant effect of stress on self-reported anxiety for introverted participants, F (2, 24) = 46.37, $p < .05$. No significant simple effect of stress level for extroverts was observed, p>.05. Simple comparisons examining the effect of stress level for introverts revealed that participants under moderate or high stress reported significantly higher levels of self-reported anxiety than did participants under low stress, F (1,24) = 62.65, $p < .05$. In addition, for introverts, participants under high stress reported significantly higher levels of self-reported anxiety than did participants under moderate levels of stress, F (1,24) = 30.07, $p < .05$.

Essentially, each row of the ANOVA table provides a single piece of information about the data. You do one test, you learn one thing. Each sentence represents what the researcher has learned from each F-test. The researcher starts with the most general piece of information, which is the interaction. Then they describe the result for a simple effect. If the simple effect is significant they report the results for the simple comparisons that were done to explain why the simple effect was significant.

If the interaction hadn't been significant then there was no need to conduct a set of simple effects or do simple comparisons. If the interaction isn't significant that means that when you're looking at the effect of a particular independent variable, it's okay to average over the levels of the second independent variable – it's okay to interpret the main effect. If a main effect is significant then the researcher should test a set of comparisons that examines differences among the levels of this independent variable. In this situation these comparisons are referred to as **main comparisons**.

The flowchart of how the analyses are conducted is presented below.