13.5.1: Practice 2x3 Factorial ANOVA on Mindset

One more time to practice with a factorial ANOVA Summary Table and interpretation!

Scenario

We are going back to one of the student research projects to see if faculty can increase understanding and belief in growth mindset. To see if students' scores on the Mindset Quiz improved so that more student held stronger growth mindset view, the outcome that we are trying to improve is the Difference between their Mindset Quiz score at the beginning of the semester (pretest), and their Mindset Quiz score at the end of the semester (posttest). A positive score means that students' mindset improved, while a negative score means that their mindset actually became more fixed. We are only going to look at students who did not start holding strong growth mindset beliefs to see if any improvement was made. In this version of the study, we had three types of interventions:

1. A comparison control group who did not complete any activities related to learning about growth mindset. They only took the Mindset Quiz at the beginning of the semester and again at the end of the semester.
2. A minimal intervention group who had some activities and discussion about growth mindset, but not every week.
3. A super-intervention group who had weekly activities and discussion about growth mindset. The concept of mindset was embedded in their curriculum.

At conference discussing prior studies in this series, the student researchers had a discussion with researcher from another school and realized that all of the professors who participated in the study were women (of varying racial identities). We then went back to see if the gender of the student influenced their changes in growth mindset. We had four students whose gender was unknown or non-binary, but they were all in the minimal intervention group, so we excluded them from the analyses and only compared men versus women. Because this study took place in real classrooms, the size of each group varied, but there was a total of 46 students who completed both the pretest and posttest of the Mindset Quiz and indicated their gender.

Let's start by seeing if you understand what's going on in this scenario.

? Exercise 13.5.1.1

Answer the following questions to understand the variables and groups that we are working with.

1. What are the IVs and their levels?
2. What is the DV (quantitative variable being measured)?
3. Is this a 2x2 factorial design? If not, what kind of design is it?
4. List out each of the *combinations* of the levels of the IVs.

Answer

1. One IV was the Mindset Intervention, with the three levels being Control, Minimal, or Super. The other IV was gender, with the levels being men or women.
2. The DV is the Difference between each students score on the first Mindset Quiz (pretest) subtracted from their score on the later Mindset Quiz (posttest).
3. This is a 2x3 factorial design because there are two IVs (two numbers) and the IV with fewer levels has 2 levels, and the IV with more levels had 3 levels.
4. Because 2x3 = 6, we will have six combinations:
    1. Women in the Control group
    2. Women in the Minimal group
    3. Women in the Super group
    4. Men in the Control group
    5. Men in the Minimal group
    6. Men in the Super group

Step 1. State the Hypothesis

Knowing nothing else, what do you predict for your research hypotheses? Will there be a main effect for gender? A main effect for the mindset intervention? An interaction of gender by mindset intervention?

? Exercise 13.5.1.2

Describe any main effects or interaction that you predict (in words only). Make sure that you predict the direction of effects by naming which group will have larger difference scores.

Answer

Dr. MO predicts that there will not be a main effect for the average Differences score from the pretest to the posttest based on gender.

Dr. MO predicts that there will be a main effect for the intervention such that the Control group will have the smallest Difference scores, the Minimal intervention group will be in the middle, and the Super intervention group will have the largest Difference score. All means will be statistically significantly different from each other.

Dr. MO predicts that there will be an interaction such that women will have a larger Difference score in in the Minimal and Super intervention scores than men, but gender won't matter in the Control group. She's basing this on the fact that all of the professors were women.

Your research hypothesis may differ! No means were provided, so we're really just guessing based on the scenario and what we know about psychology and growth mindset…

What about the null hypotheses?

? Exercise 13.5.1.3

What are the null hypotheses for this scenario (in words)?

Answer

All means will be similar.

If it makes more sense to you to mirror the research hypotheses, you can also list out the main effects and interaction:

- The null hypothesis for gender is that the mean Difference score for the men condition will be similar to the Difference score for women.
- The null hypothesis for the Mindset Intervention is that the mean Difference will be similar between the Control group, Minimal group, and Super group.
- The null hypothesis for the interaction is that the average Difference score for Gender and Mindset Intervention combined will be similar. [You could list out all of the combinations here again, too.]

Step 2. Find the Critical Values

You are welcome to complete this step after you've completed the ANOVA Summary Table so that we have all of the necessary Degrees of Freedom. Instead, we will find the Degrees of Freedom now.

Formulas for Degrees of Freedom

1. Cells: $(k_1 \times k_2) - 1$
2. Between group for one variable ($IV1_1$): $k_1 - 1$
3. Between group for the other variable (IV_2): $k_2 - 1$
4. Interaction: $df_1 \times df_2$
5. Within group: $df_{Total} - df_{Cells}$
6. Total: N-1

? Exercise 13.5.1.4

What are the Degrees of Freedom for each of the six sources if N = 46?

Answer

1. Cells: $(k_1 \times k_2) - 1 = (2 \times 3) - 1 = 6 - 1 = 5$
2. Between group for one variable ($IV1_1$): $k_1 - 1 = 2 - 1 = 1$

3. Between group for the other variable (IV_2) $k_2) - 1 = 3 - 1 = 2$
4. Interaction: \(df_1 \times df_2 = 1 \times 2 = 2\)
    1. Notice that we are multiplying the Degrees of Freedom of each, not k! It's so easy to get confused...
5. Within group (You have to find the Degrees of Freedom for the Total (#6 here) before you can find this one.): $df_{Total} - df_{Cells} = 45 - 5 = 40$
6. Total: $N - 1 = 46 - 1 = 45$

Now that we have all of the Degrees of Freedom, we can use them to find the critical values.

? Exercise 13.5.1.5

Using the Table of Critical F-Scores, what are the critical values of F for p=0.05 for:

1. Main Effect of IV 1 (Gender)?
2. Main Effect of IV 2 (Mindset Intervention)?
3. Interaction?

Answer

1. Main Effect of IV 1 (Gender) = Using the Degrees of Freedom of the numerator (IV 1) and the denominator (WG), $F_{Crit}(1, 40) = 4.08$
2. Main Effect of IV 2 (Mindset Intervention) = Using the Degrees of Freedom of the numerator (IV 2) and the denominator (WG), $F_{Crit}(2, 40) = 3.23$
3. Interaction = Using the Degrees of Freedom of the numerator (Interaction) and the denominator (WG), $F_{Crit}(2, 40) = 3.23$

Step 3. Compute the Test Statistic

We'll skip calculating the Sums of Squares; you can find them already in the ANOVA Summary Table (Table 13.5.1.1).

Table 13.5.1.1- Factorial ANOVA Summary Table

Source	SS	df	MS	F
IV 1	141.69			
IV 2	31.82			
Interaction	53.01			
Within Groups (Error)	1,072.08			
Total	1,298.60			

✓ Example 13.5.1.1

Use the Sum of Squares provided in Table 13.5.1.1to fill in the rest of the ANOVA Summary Table in Table 13.5.1.2

Solution

To complete the table, you need to need to calculate the degrees of freedom of cells $(df_{cells} = (k_1 \times k_2) - 1)$ although there's no place in the table to include this information:

$$\begin{aligned} df_{cells} &= (k_1 \times k_2) - 1 \\ df_{cells} &= (2 \times 3) - 1 \\ df_{cells} &= 6 - 1 \\ df_{cells} &= 5 \end{aligned}$$

Table 13.5.1.2-Factorial ANOVA Summary Table

Source	SS	df	MS	F

Source	SS	df	MS	F
IV 1	141.69	$k_1 - 1 = 2 - 1 = 1$	$\frac{SS_1}{df_1} = \frac{141.69}{1} = 141.69$	$\frac{MS_1}{MS_{WG}} = \frac{141.69}{26.80} = 5.29$
IV 2	31.82	$k_2 - 1 = 3 - 1 = 2$	$\frac{SS_2}{df_2} = \frac{31.82}{2} = 15.91$	$\frac{MS_2}{MS_{WG}} = \frac{15.91}{26.80} = 0.59$
Interaction	53.01	$df_1 \times df_2 = 1 \times 2 = 2$	$\frac{SS_{INT}}{df_{INT}} = \frac{53.01}{2} = 26.51$	$\frac{MS_{Interaction}}{MS_{WG}} = \frac{26.51}{26.80} = 0.99$
Within Groups (Error)	1,072.08	$df_{Total} - df_{cells} = 45 - 5 = 40$	$\frac{SS_{WG}}{df_{WG}} = \frac{1072.08}{40} = 26.80$	N/A
Total	1,298.60	$N - 1 = 46 - 1 = 45$	N/A	N/A

We did it! The completed ANOVA Summary Table for this scenario looks like Table 13.5.1.3

Table 13.5.1.3-Factorial ANOVA Summary Table

Source	SS	df	MS	F
IV 1	141.69	1	141.69	5.29
IV 2	31.82	2	15.91	0.59
Interaction	53.01	2	26.51	0.99
Within Groups (Error)	1,072.08	40	26.80	leave blank
Total	1,298.60	45	leave blank	leave blank

ANOVA Summary Tables with more than a 2x2 design are less repetitive than ANOVA Summary Tables for a 2x2, huh?

Step 4. Make the Decision

To make the decision, we need to compare each calculated F-score to the critical values that we found in Step 2.

Note

Remember:

(Critical $<$ Calculated) = Reject null = At least one mean is different from at least one other mean. = $p < .05$

(Critical $>$ Calculated) = Retain null = All of the means are similar. = $p > .05$

Exercise 13.5.1.6

For each calculated F (main effect for IV 1, main effect for IV 2, interaction), decide if the null hypothesis should be retained or rejected.

Answer

1. For the Main Effect of IV 1 (Gender) = $F_{Crit}(1, 40) = 4.08$, so we reject the null hypothesis. The higher mean is higher (because there are only two means).
2. Main Effect of IV 2 (Mindset Intervention) = $F_{Crit}(2, 40) = 3.23$, so we retain the null hypothesis. The three means for the Control group, Minimal group, and Super group were similar.
3. Interaction = $F_{Crit}(2, 40 = 3.23$, so we retain the null hypothesis. Gender and the Mindset Intervention did not interact.

Factorial designs are more complex, but it's the same basic process that we've been working through this whole time.

But before we move on to the full reporting of results, let's just make sure that we know how to write the "statistical sentences".

Exercise 13.5.1.1

What would the statistical sentences look like for each effect?

Answer

1. Main Effect of IV 1 (Gender): F(1, 40)=5.29, p<.05
2. Main Effect of IV 2 (Mindset Intervention): F(2,40)=0.59, p>.05
3. Interaction: F(2, 40)=0.99, p>.05

Pay attention to the greater than/lesser than sign after the p! If p<.05, it means that there's a very small probability that the means are similar. If p>.05, there's a large probability that the means are similar (that the null hypothesis is correct).

Before we can write the conclusion, we need to know the means to see if they are in the direction that we hypothesized way back in Step 1. The means are shown in in Table 13.5.1.4 We do not need to conduct any pairwise comparisons because the only means that were statistically significantly different were for gender, which only had two levels, so we don't have to do any extra calculations to know that the higher mean is significantly higher.

Table 13.5.1.4-Mean Differences for 2x3 (Gender by Mindset Intervention) Factorial ANOVA

IV Levels	**IV1 (Gender): Men**	**IV1 (Gender): Women**	Marginal Means for IV2
IV2 (Mindset Intervention): Control	-2.33	1.56	0.50
IV2 (Mindset Intervention): Minimal	-3.50	3.16	1.34
IV2 (Mindset Intervention): Super	1.00	2.33	1.62
Marginal Means for IV1	-1.31	2.57	1.22

We are ready to do the write-up! Remember, these are means of the Difference score between each participants' pretest and posttest, and negative scores mean that the student's posttest was actually lower than their pretest; another way to say that is that the student was closer to holder growth mindset beliefs at the beginning of the semester (before the intervention) than at the end of the semester.

It can be easier to go through each research hypothesis, one at a time.

Example 13.5.1.2

Evaluate the research hypothesis for the main effect of gender by writing up the conclusion.

Solution

It was predicted that there will not be a main effect for the average Differences score from the pretest to the posttest based on gender. She was incorrect; there were gender differences in the Difference score (F(1, 40)=5.29, p<.05) such that women scored higher (M=2.57) than men (M=-1.31).

Now for the second main effect:

Example 13.5.1.3

Evaluate the research hypothesis for the main effect of the intervention by writing up the conclusion.

Solution

It was predicted that there will be a main effect for the intervention such that the Control group will have the smallest Difference scores, the Minimal intervention group will be in the middle, and the Super intervention group will have the largest Difference score. This was not supported (F(2, 40)=0.59, p>.05). Instead, the three means were similar ($M_{Control} = 0.50; M_{Minimal} = 1.34, M_{Super} = 1.62$).

And then the interaction:

Example 13.5.1.4

Evaluate the research hypothesis for the interaction by writing up the conclusion.

Solution

It was predicted predicted that there will be an interaction such that women will have a larger Difference score in in the Minimal and Super intervention scores than men, but gender won't matter in the Control group. This was not supported (F(2,40)=0.99, p>.05). The means by gender and mindset intervention did not statistically significantly interact. The means for each combination can be seen in Table 13.5.1.4

Before you move on, look at back at your own write-ups to makes sure that each of them includes all of the four required components when Reporting Results.

Summary of the Current Mindset Study

Now that we've interpreted all of the statistical analyses, let's step back a bit and think about this study. It was predicted that gender wouldn't matter on its own, but that the mindset intervention would. But the results actually are the opposite! What happened?

Well, we don't know for sure, and would have to do several more studies to try to figure out exactly what was going on. What we can see is that these interventions designed to increase Mindset Quiz scores actually made men have a more fixed mindset. We're not really sure why that is, and future research should have more than professors who are women provide the mindset interventions. We could also look at the interventions themselves, and see if those may have inadvertently encouraged women but discouraged men. But this is the fun of research! We accidently found a puzzle, and can do more and more research to find all of the puzzle pieces to understand what's going on!

Or, we could be experiencing a Type I Error and/or a Type II Error!

Example 13.5.1.5

What is a Type I Error, and what is a Type II Error (in your own words)? How could they relate to this scenario?

Solution

Type I Errors are when you reject null hypothesis when it is true. In this scenario, we could be saying that there is a difference in changes in mindset between the genders when there wouldn't be any change in a larger sample or in the population.

Type II Errors are when we retain the null hypothesis when it is false. In this scenario, we said that the intervention didn't work. If we committed a Type II Error, we could be missing how influential these interventions are because we had a wonky sample.

The good news is that conducting more research studies solves those problems, too! Win-win! If we conduct more research, we can figure out the puzzle, and rule out either of these errors.

Okay, we're almost done with this chapter, and with learning how to compare means between groups. The only next step is to decide which statistical analysis to run!
