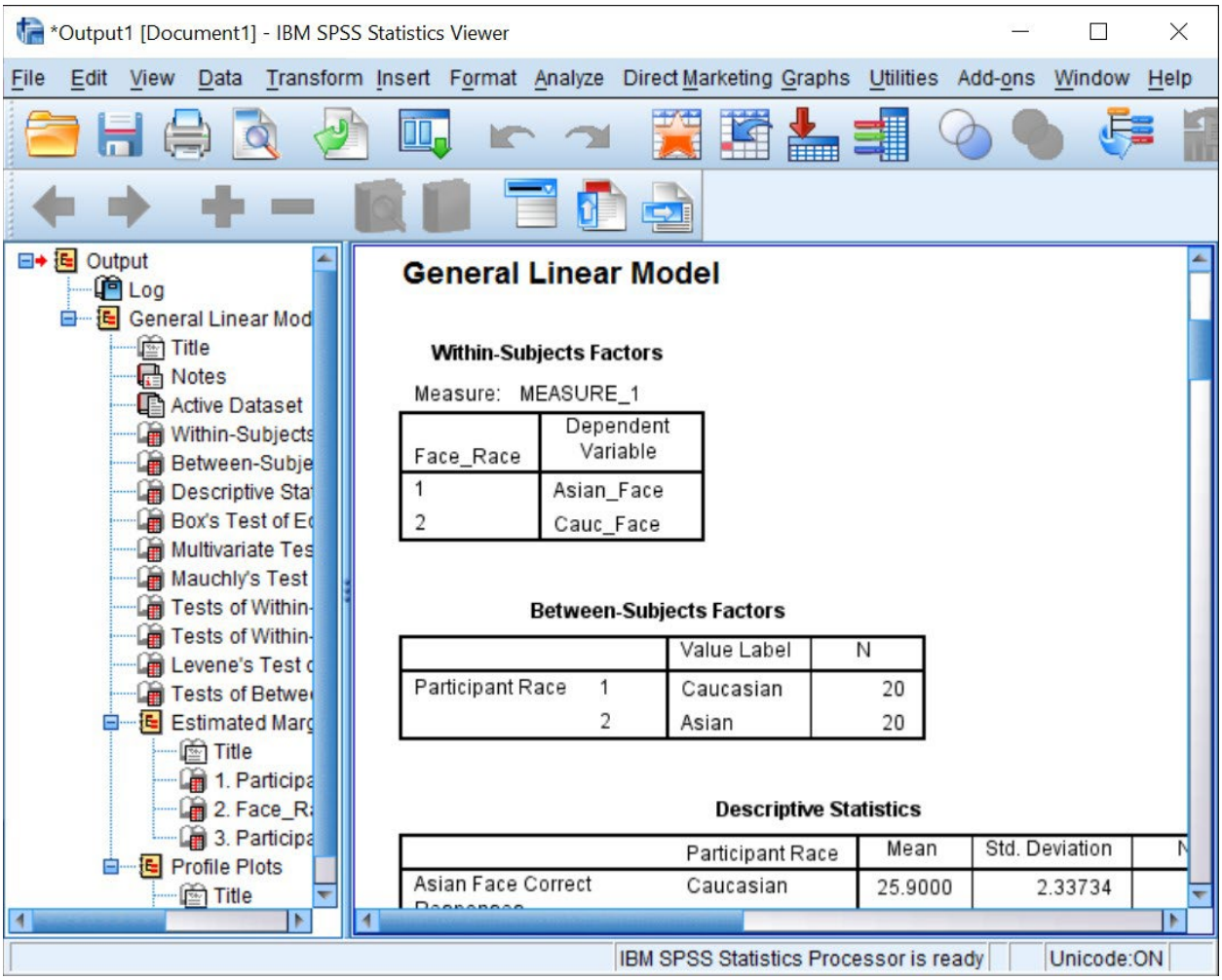
Two-Way Mixed ANOVA



There is quite a lot of output for Analysis of Variance, but don't worry - this tutorial will talk you through the output you need, box by box.

Within-Subjects Factors

This box is just here to remind you what values you have assigned the different levels of your within-participants variable, and what they mean. You may find it useful to refer back to this when interpreting your output.

Within-Subjects Factors

Measure: MEASURE_1

Face_Race	Dependent Variable
1	Asian_Face
2	Cauc_Face

From looking at the box you should be able to see that for the first factor, **Face_Race**, there are two levels, where: 1=Asian Faces and 2=Caucasian Faces.

Between-Subjects Factors

This box is just here to remind you what values you have assigned the different levels of your between-participants variable, and what they mean. You may find it useful to refer back to this when interpreting your output.

Between-Subjects Factors

		Value Label	N
Participant Race	1	Caucasian	20
	2	Asian	20

From looking at the box you should be able to see that for the second factor, **Race**, there are also two levels, where: 1 = Caucasian participants and 2 = Asian participants.

Descriptive Statistics

Descriptive Statistics

Participant Race	Mean	Std. Deviation	N

This table gives you your descriptive statistics. It's good to look at and report these, as they can give you an important insight into the pattern of your data. From this table we can see the average (mean) number of faces recognised for each of the possible variable conditions. You can also see the variation in the data (i.e. spread of scores) for the different groups from the standard deviation.

From this table we can see that on average, the overall Caucasian and Asian participants recognised similar numbers of faces (**means = 26.70 and 26.75 respectively**).

However, the breakdown of conditions showed that for the Asian facial stimuli, Asian participants got more correct responses (**mean = 27.60, SD = 2.30**) than Caucasian participants (**mean = 25.90, SD = 2.34**); but for the Caucasian faces, Caucasian participants (**mean = 27.45, SD = 2.21**) outperformed the Asian participants (**mean = 25.95, SD = 2.68**).

Box's Test of Equality of Covariance Matrices and **Multivariate Tests**

You only need the next couple of output boxes when you have multiple dependent variables (which we don't). Skip over these and go straight to the **Mauchly's Test of Sphericity** box.

Mauchly's Test of Sphericity

Mauchly's Test of Sphericity[a]

Measure: MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[b]		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
Face_Race	1.000	.000	0	.	1.000	1.000	1.000

This table tests whether the assumption of sphericity has been met. This is a bit like the assumption of homogeneity of variance for independent tests; and like Levene's test, we *do not* want Mauchly's test to be significant. However, rather than assuming equal levels of variance in the data for the different conditions, in this case we assume that the relationship between the different **pairs of conditions** is similar.

The good news is, you only need to look at this table when you have a within participants variable with more than two levels (i.e. more than one pair)... which in this case, we don't have - **Face_Race** only has two conditions!!

Tests of Within-Subjects Effects

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Face_Race	Sphericity Assumed	.050	1	.050	.010	.920	.000
	Greenhouse-Geisser	.050	1.000	.050	.010	.920	.000
	Huynh-Feldt	.050	1.000	.050	.010	.920	.000
	Lower-bound	.050	1.000	.050	.010	.920	.000
Face_Race * Race	Sphericity Assumed	51.200	1	51.200	10.588	.002	.218
	Greenhouse-Geisser	51.200	1.000	51.200	10.588	.002	.218
	Huynh-Feldt	51.200	1.000	51.200	10.588	.002	.218
	Lower-bound	51.200	1.000	51.200	10.588	.002	.218
Error(Face_Race)	Sphericity Assumed	183.750	38	4.836			
	Greenhouse-Geisser	183.750	38.000	4.836			
	Huynh-Feldt	183.750	38.000	4.836			
	Lower-bound	183.750	38.000	4.836			

This is the one of the most important tables in the output. It gives you the ANOVA results for your within-participants variable (**Face_Race**); and any interactions with that variable.

The key columns you need to interpret your analysis are:

- **df** stands for **degrees of freedom.** Degrees of freedom are crucial in calculating statistical significance, so we need to report them. For each main effect and interaction there are **two df** values of interest: the factor (or interaction) df and the error term df. In this example the **dfs** can be found in the **Sphericity Assumed** rows (you would only use the other rows if the sphericity assumption had not been met).

- **F** stands for **F-Ratio**. This column gives you the **F-Ratio** statistics calculated by the ANOVA. It is essentially calculated by dividing the systematic variance (i.e. the variation in your data that can be explained by your experimental manipulation) by the unexpected, unsystematic variance. The larger your **F-Ratio**, the more likely it is that your experimental manipulation (your IV) will have had a significant effect on the DV.

  In this example the you need to report the values in the **Sphericity Assumed** rows.

- **Sig** stands for **Significance Level.** This column gives you the probability that the results could have occurred by chance if the null hypothesis were true. The p-value should be smaller than 0.05 for the F-ratio to be significant. If this is the case (i.e. $p<.05$) we reject the null hypothesis and infer that our experimental manipulation (our IV) has had a significant effect.

  However, if the p-value is larger than 0.05, we have to retain the null hypothesis: that there is no difference between the conditions (or no interaction).

- **Partial Eta Squared.** While the *p*-value can tell you whether your main effects and interaction terms are statistically significant, partial eta squared (η_p^2) tells you about the magnitude of these effects. As such, we refer to this as a measure of **effect size**. To interpret your effects sizes, you can use the following cut-offs:
  - 0.14 or more are large effects
  - 0.06 or more are medium effects
  - 0.01 or more are small effects

So we know which columns we need to look at, but what numbers do we use?

Which row you choose depends on whether your assumption of Sphericity has been met. If your Mauchley's test was **non-significant** (i.e. the assumption had been met), **or** if you have less than 3 levels in your within-participants IV (which we do!), then you read across from the **Sphericity Assumed** row. If Mauchley's *had* been significant, then you would need to use one of the other rows.

SPSS groups your analysis into blocks: one block for each variable and interaction.

First, let's look at **Face_Race**.

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Face_Race	Sphericity Assumed	.050	1	.050	.010	.920	.000
	Greenhouse-Geisser	.050	1.000	.050	.010	.920	.000
	Huynh-Feldt	.050	1.000	.050	.010	.920	.000
	Lower-bound	.050	1.000	.050	.010	.920	.000
Face_Race * Race	Sphericity Assumed	51.200	1	51.200	10.588	.002	.218
	Greenhouse-Geisser	51.200	1.000	51.200	10.588	.002	.218
	Huynh-Feldt	51.200	1.000	51.200	10.588	.002	.218
	Lower-bound	51.200	1.000	51.200	10.588	.002	.218
Error(Face_Race)	Sphericity Assumed	183.750	38	4.836			
	Greenhouse-Geisser	183.750	38.000	4.836			
	Huynh-Feldt	183.750	38.000	4.836			
	Lower-bound	183.750	38.000	4.836			

Using the Sphericity Assumed row, you need to report your results as:

F (*IV df*, *error df*) = *F-Ratio*, p = *Sig*, η_p^2 = *Partial Eta Squared*

...along with a sentence, explaining what you have found. In this case you might say:

There was no significant main effect of Face_Race on face recognition scores overall (F(1, 38) = .01, p = .92, η_p^2 < .001).

Next, we can use the same method to look at the interaction term. The results are in the **Face_Race * Race** block. Again, using the Sphericity Assumed row to locate the correct values, you might report your results as something like:

There was a significant interaction between Face_Race and Race in terms of recognition scores (F(1, 38) = 10.59, p = .002, η_p^2 = .22).

Tests of Within-Subjects Contrasts

You really don't need to worry about this table for interpreting your results... so skip over this and go straight to your **Levene's Test of Equality of Error Variances** which assesses the assumption of homogeneity of variance for your *between-participants* variable (Race).

Levene's Test of Equality of Error Variances

Levene's Test of Equality of Error Variances[a]

	F	df1	df2	Sig.
Asian Face Correct Responses	.157	1	38	.695
Caucasian Face Correct Responses	1.019	1	38	.319

When conducting an ANOVA with any between-participants variables, one assumption that must be met is that the groups you are comparing have a similar dispersion of scores (i.e. homogeneity of variance). The Levene's Statistic tells us whether or not this is the case.

For Mixed ANOVA, this assumption is assessed at all levels of your within-participants variable (i.e. for both Asian and Caucasian facial stimuli).

If the test is significant this indicates that there are statistically significant differences in the way in the data are dispersed, suggesting that the assumption has been violated. As such, we are looking for a ***non-significant*** result here. In this example, the *Sig.* column tells us that this *is* what we have found as both p-values are greater than .05.

Tests of Between-Subjects Effects

Tests of Between-Subjects Effects

Measure: MEASURE_1

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Intercept	57138.050	1	57138.050	8659.007	.000	.996
Race	.200	1	.200	.030	.863	.001
Error	250.750	38	6.599			

This is another one of the most important tables in the output. This is where we get our ANOVA statistics for the between-participants variable (Race).

As with the Tests of Within-Subjects Effects table, the columns we need to interpret your analysis are highlighted above: ***df***, ***F***, ***Sig*** and ***Partial Eta Squared***

In this case we need to read across the row representing our IV (labelled **Race**) and the Error row.

As with the results for the within-participants variables and interaction, you need to report your results using the following formula:

F (*IV df*, *error df*) = *F-Ratio*, p = *Sig*, η_p^2 = *Partial Eta Squared*

...along with a sentence, explaining what you have found. For example:

There was no significant main effect of Race on face recognition scores overall (F(1, 38) = .03, p = .86, η_p^2 = .001).

We now have our ANOVA results for both of our IVs (**Face_Race** and **Race**) and the interaction between these variables. But the question is, what does it all mean?

We know that while neither of our IVs produced a significant main effect, we did find an interaction between our IVs... but we now need to interpret and explain what this actually means in English!! To do that, we need to refer back to our **Descriptive Statistics** to see what is happening in each of our different conditions.

Estimated Marginal Means

These boxes show you the breakdown of the means for the different levels of your IVs, and all of the different condition combinations. For a Mixed ANOVA, the mean values are the same as those displayed the **Descriptive Statistics** box at the beginning of your output. Note that these boxes report the Standard Error rather than the Standard Deviation through.

1. Participant Race

Measure: MEASURE_1

Participant Race	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Caucasian	26.675	.406	25.853	27.497
Asian	26.775	.406	25.953	27.597

This first box tells you the overall mean recognition scores for both levels of your first IV: **Participant Race**. Here we can see that both Caucasian (mean = 26.68) and Asian (mean = 26.78) participants performed similarly.

2. Face_Race

Measure: MEASURE_1

Face_Race	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	26.750	.367	26.007	27.493
2	26.700	.389	25.913	27.487

The second box tells you the overall mean recognition scores for the levels of your second IV: **Face_Race**. From the very first box in the SPSS output, we saw that 1=Asian Faces and 2=Caucasian Faces.

As such, we can see that overall participants performed similarly for both Asian (mean = 26.75) and Caucasian (mean = 26.70) faces.

3. Participant Race * Face_Race

Measure: MEASURE_1

Participant Race	Face_Race	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
Caucasian	1	25.900	.519	24.850	26.950
	2	27.450	.550	26.337	28.563
Asian	1	27.600	.519	26.550	28.650
	2	25.950	.550	24.837	27.063

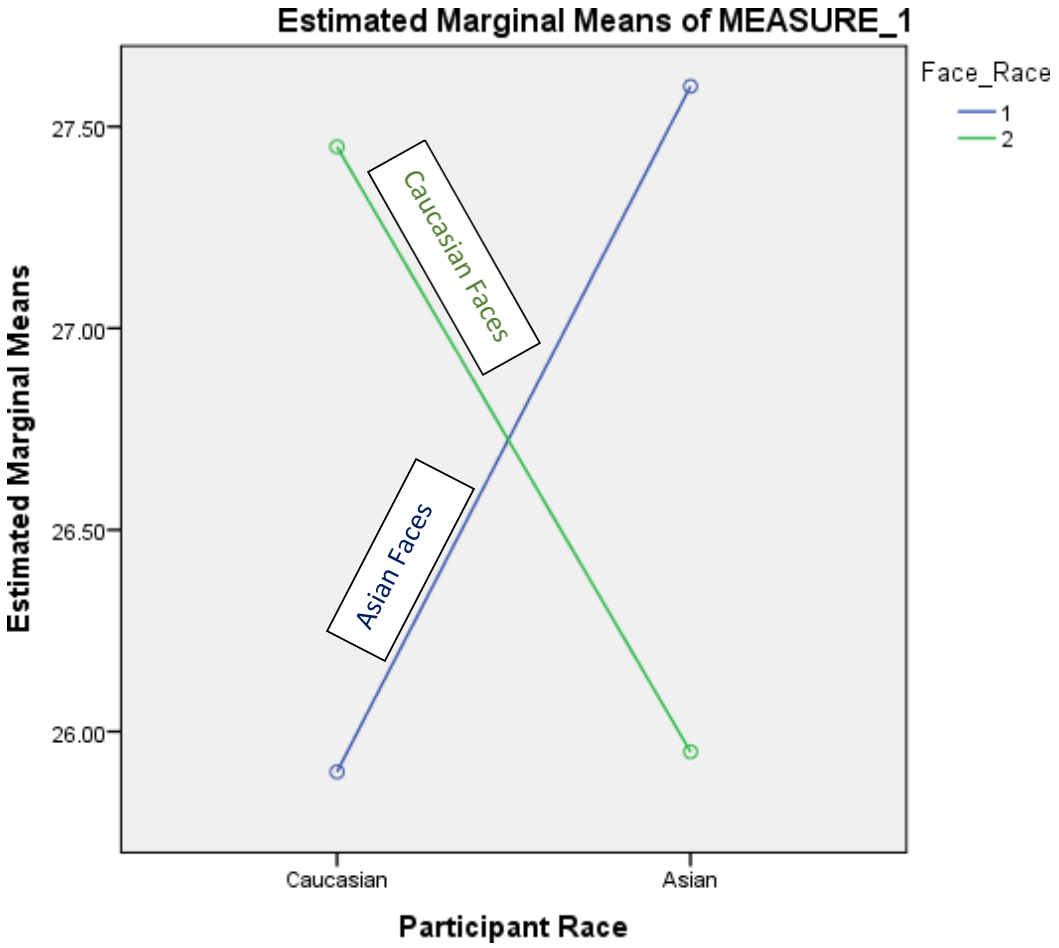
The final box gives you the breakdown of facial recognition scores for all of your condition combinations. This can help you to interpret your interaction. Again, remembering that the number codes are 1=Asian Faces and 2=Caucasian Faces, we can interpret the pattern of results accordingly.

Here we can see that while Asian participants recognised more Asian faces (**mean = 27.60**) than Caucasian faces (**mean = 29.95**); Caucasian participants showed the opposite pattern (**Asian faces mean = 25.90; Caucasian faces mean = 27.45**), suggestive of an own-race bias in face recognition.

Profile Plots

The final part of the output is the **Profile Plot**. Graphs can be really useful in helping you to visualise and understand your interaction. When looking at the graph, it may help you to consider the following:

- Are the lines doing the same thing… If not, what are they doing?
- Are some points on the graph closer together than others?



Here, the blue line represents **Face_Race 1** = *Asian Faces*. From the way the line slopes we can see that Asian Participants performed much better for this stimuli than Caucasian Participants.

In contrast, the green line (which represents **Face_Race 2** = *Caucasian Faces*) slopes in the opposite direction. For this stimuli type, Caucasian participants performed better Asian Participants.

As well as looking at the slopes of the lines, we can also look at the vertical distance between the dots for each race of participant. Doing this we can see that both races performed better for faces of their own race, compared to those of the other race. And this own-race advantage was of a similar magnitude in both cases.

The trick now is to put **all** of the information from your output together to make a results section that is sensible and meaningful!!

How do we write up our results?

When writing up the findings from your analysis in APA format, you need to include all of the relevant information covered by the previous slides.

- What were the inferential statistics for your IVs
    - i.e. what were the ANOVA results
- For significant findings, where did the significant difference(s) lie
    - i.e. what were the results of your post hoc tests, if you did them
- What was the pattern and direction of these differences
    - i.e. what were the means and descriptive statistics for your conditions
- If the interaction term was significant, what does it mean in English
    - i.e. describe what was happening in the different conditions

It is usually helpful to the reader of your results if you **include a table of the means and standard deviations** for all of the conditions in your results. It also helps to **include a graph of your results.** You can edit your graph by double clicking on it in the output in SPSS.

For this example, you might end up writing a results section that looks a bit like this:

> **The results of the Two-Way Mixed ANOVA showed that there was no significant main effect of Participant Race ($F(1, 38) = .03$, $p = .86$, $\eta_p^2 = .001$) on face recognition scores, with Caucasian (mean = 26.68) and Asian (mean = 26.78) performing similarly overall.**
>
> **In addition, there was also no significant main effect of Face Race on recognition scores ($F(1, 38) = .01$, $p = .92$, $\eta_p^2 < .001$), with participants showing similar average recognition scores for Asian (mean = 26.75) and Caucasian (mean = 26.70) faces.**
>
> **In contrast, there was a significant interaction between Race of Participant and Face Race ($F(1, 38) = 10.59$, $p = .002$, $\eta_p^2 = .22$). Descriptive statistics showed that while Asian participants performed better for Asian facial stimuli (mean = 27.60, SD = 2.30) compared to Caucasian stimuli (mean = 25.95, SD = 2.68); Caucasian participants showed the opposite pattern (Asian faces mean = 25.90, SD = 2.34; Caucasian faces mean = 27.45, SD = 2.21). From looking at the graph we can see that both races showed a similarly sized recognition advantage for own-race faces.**
>
> **These findings support the notion of an own-race bias in face recognition.**

This brings us to the end of this tutorial. Why not download the data file used in this tutorial and see if you can run the analysis yourself?