SOURCEBOOK SPSS ANNOTATED OUTPUT

Abstract: This chapter is intended to facilitate the connection between standard introductory statistics concepts and their implementation in SPSS. It shows the output from various types of analyses, describes how to interpret the output, and shows the link between hand calculation formulas and SPSS output. Results derive from the examples in the other sections of this project.

Keywords: SPSS output, annotation, statistical interpretation

Original: July 2009 Updated: May 2025

This document is part of an online statistics sourcebook.

A browser-friendly viewing platform for the sourcebook is available: https://cwendorf.github.io/Sourcebook

> All data, syntax, and output files are available: https://github.com/cwendorf/Sourcebook

TABLE OF CONTENTS FOR THIS CHAPTER

Frequencies and Descriptives	3
Correlations	4
Regression	5
Confidence Interval for a Mean	6
One Sample t Test	7
Paired Samples t Test	8
Independent Samples t Test	9
OneWay ANOVA	10
Post Hoc Comparisons	11
Repeated Measures ANOVA	12
Factorial ANOVA	14

Frequencies an

"N" provides the sample size for the entire data set. "Missing" refers to the number of entries that are blank, whereas "Valid" is the number of entries that are not blank.

Outcome		
N	Valid	8
	Missing	0
Mean		4.0000
Std. Deviation		3.11677
Variance		9.714
Pereentiles	25	2.2500
	50	4.0000
	75	5.5000

The "Mean", "Standard Deviation", and "Variance" are all calculated as unbiased estimates of the respective population parameter. Here, the mean is determined as the average of the scores weighted by their frequencies:

$$M = \frac{\sum (fY)}{N} = \frac{(2 \times 0) + (1 \times 3) + (2 \times 4) + (1 \times 5) + (1 \times 7) + (1 \times 8)}{8} = 4$$

The "Variance" and "Std. Deviation" are both functions of the Sum of Squares (not shown in the output) of the scores in the frequency distribution:

$$SS = \sum f(Y - M)$$

$$SS = 2(0 - 4)^2 + 1(3 - 4)^2 + 2(4 - 4)^2 + 1(5 - 4)^2 + 1(7 - 4)^2 + 1(8 - 4)^2 = 68.000$$

The "Variance" (i.e., Mean Squares) and "Std. Deviation" are calculated as:

$$MS = \frac{SS}{(N-1)} = \frac{68}{7} = 9.714$$

$$SD = \sqrt{MS} = \sqrt{9.714} = 3.117$$

Outcome

					Cumulative
	Frequency		Percent	Valid Percent	Percent
.00	2		25.0	25.0	25.0
.00	1	V	12.5	12.5	37.5
.00	2		25.0	25.0	62.5
.00	1	1	12.5	12.5	75.0
.00	1	/	12.5	12.5	87.5
.00	1		12.5	12.5	100.0
otal	8		100.0	100.0	
	.00 .00 .00 .00	00 2 00 1 00 2 00 1 00 1	00 2 00 1 00 2 00 1 00 1	00 2 25.0 00 1 12.5 00 2 25.0 00 1 12.5 00 1 12.5 00 1 12.5	00 2 25.0 25.0 00 1 12.5 12.5 00 2 25.0 25.0 00 1 12.5 12.5 00 1 12.5 12.5 00 1 12.5 12.5 00 1 12.5 12.5

"Percentiles" provide the scores associated with particular percentile ranks. For example, the 50th percentile is the score in the following position:

$$Position = PR(N + 1) = .50(8 + 1) = 4.5$$

Thus, the score at the 50th percentile is the 4.5th score in the frequency distribution – a score of 4. Similarly, a score of .75 is at the 25th percentile and a score of 6.5 is at the 75th percentile. Importantly, in some cases, the score values are non-integer interpolated values.

The "Valid" column lists all of the actual scores in the entire data set. "Frequency" indicates the number of times that score exists. For example, the score 4 was listed 2 times. The "Percent" column provides the percentage of cases for each possible score. For example, of the 8 scores in the entire data set, the score of 4 was listed 2 times and 2/8 is 25.0%.

The "Valid Percent" column provides the percentage of cases for each possible score divided by the total number of cases. Here, there were no missing scores, so the percent columns are equal.

"Cumulative Percent" is the sum of all percentages up to and including the row in question. For example, 62.5% of scores were a 4 or smaller. Similarly, 37.5% were a 3 or smaller.

Correlations

These values of the statistics are identical to the values that would be provided by the "Frequencies" or "Descriptives" commands. See the earlier annotated output for details of how these are computed from frequency distributions. Note that they are calculated separately for each variable.

Descriptive Statistics

	Mean	Std. Deviation	N
Outcome 1	2.0000	2.44949	•
Outcome 2	6.0000	2.44949	4

	Correlations		
		Outcome 1	Outcome 2
Outcome 1	Pearson Correlation	1	.500
	Sig. (2-tailed)		.500
	Sum of Squares and Cross-products	18.000	9.000
	Covariance	6.000	3.000
	N	4	4
Outcome 2	Pearson Correlation	.500	
	Sig. (2-tailed)	.500	\\
	Sum of Squares and Cross-products	9.000	18.000
	Covariance	3.000	6.000
	N /		4/

^{*.} Correlation is significant at the 0,05 level (2-tailed).

This box presents information about the first variable that can be derived from the "Std. Deviation" above. The "Covariance" here actually represents the estimated Variance (or Mean Squares) of the variable: $MS = SD^2 = 2.44949^2 = 6.000$. The "Sum of Squares" for this variable is then found (knowing that $SS = MS \times df$) to be equal to 18.000. Finally, note that, by definition, the variable is perfectly correlated with itself (r = 1.0).

These boxes represent the conjunction of both variables and therefore present the statistics relevant to the relationship between the two variables. (Thus, the boxes are redundant.)

This "Sum of Cross Products" (SCP) is not easily determined from the summary statistics of the output, but rather from the data (and the calculations are therefore not shown here).

The "Covariance" is a function of the Sum of Cross Products and the sample size:

$$COV = \frac{SCP}{(N-1)} = \frac{9.000}{(4-1)} = 3.000$$

The "Pearson Correlation" coefficient is a function of the covariance and the standard deviations of both variables:

$$r = \frac{COV}{(SD_Y)(SD_Y)} = \frac{3.000}{(2.449)(2.449)} = .500$$

Though the statistic is not shown, *t* provides the standardized statistic for testing whether the correlation differs from zero:

$$t = \frac{r}{\sqrt{(1 - r^2)/(N - 2)}} = \frac{.500}{\sqrt{(1 - .500^2)/(4 - 2)}} = .816$$

The t statistic follows a non-normal (studentized or t) distribution that depends on degrees of freedom. Here, df = N - 2 = 4 - 2 = 2. A t with 2 df that equals .816 has a two-tailed probability (p) of .500, which is not a statistically significant finding.

This box presents information about the second variable that can be derived from the "Std. Deviation" above. The "Covariance" here actually represents the estimated Variance (or Mean Squares) of the variable: $MS = SD^2 = 2.44949^2 = 6.000$. The "Sum of Squares" for this variable is then found (knowing that $SS = MS \times df$) to be equal to 18.000. Finally, note that, by definition, the variable is perfectly correlated with itself (r = 1.0).

Regression

(Note that some aspects of this output have been deleted and rearranged for the sake of presentation!)

Descriptive Statistics

_			_
	Mean	Std. Deviation	14
Outcome 2	6.0000	2.44949	4
Outcome 1	2.0000	2.44949	4

These values of the statistics are identical to the values that would be provided by the "Frequencies" or "Descriptives" commands. See the earlier annotated output for details of how these are computed from frequency distributions. Note that they are calculated separately for each variable.

Variables Entered/Removeda

Model	Variables Entered	Variables Removed	Method
1	Outcome 1 ^b		Enter

- a. Dependent Variable: Outcome 2
- b. All requested variables entered.

Model Summary

Mod	del					Std. Error of the
		R	R Square	Act	sted R Square	Estimate
1		.500a	.250		125	2.59808

Unstandardized Coefficients

5.000

.500

Coefficientsa

1.785

.612

Std. Error

Standardized

Coefficients

Beta

a. Predictors: (Constant), Outcome 1

These calculations are dependent on the Covariance ("COV"), which is not determinable from the summary statistics provided, but rather the data. Therefore, the calculations for it are not shown here.

"R" is a function of the covariance and the standard deviations of both variables:

$$R = \frac{COV}{(SD_X)(SD_Y)} = \frac{3.000}{(2.45)(2.45)} = 0.500$$

$$R^2 = 0.500^2 = 0.250$$

816

The "Unstandardized Regression Coefficients" are also a function of the Covariance and the descriptive statistics:

$$B_1 = \frac{COV}{(SD_X)^2} = \frac{3.000}{(2.449)^2} = 0.500$$

.500

$$B_0 = M_Y - (B_1)(M_X) = 6.000 - (0.500)(2.000) = 5.000$$

The "Standardized Regression Coefficient" for the predictor can be similarly determined:

$$\beta_1 = B_1 \left(\frac{SD_X}{SD_Y} \right) = 0.500 \left(\frac{2.449}{2.449} \right) = 0.500$$

Outcome 1

(Constant)

Model

.500

Confidence Interval for a Mean

Case Processing Summary

Cases Valid Total Missing Percent Ν Percent Ν Ν Percent 9 8 1 Outcome 88.9% 11.1% 100.0% The values of the descriptive statistics in this tables – like the "Mean", "Variance", and "Std. Deviation" – are identical to the values that would be provided by the "Frequencies" or "Descriptives" commands. See the earlier annotated output for details of how these are computed from frequency distributions.

	Descriptives		
-		Statistic	Std. Error
Outcome	Mean	4.0000	1.10195
	95% Confidence Interval for Lower Bound	1.3943	
	Mean Upper Bound	6.6057	
	5% Trimmed Mean	 3.9444	
	Median	4.0000	
	Variance	9.714	
	Std. Deviation	3.11677	
	Minimum	.00	
	Maximum	9.00	
	Range	9.00	
	Interquartile Range	5.75	
	Skewness	.151	.752
	Kurtosis	467	1.481

The "Standard Error of the Mean" provides an estimate of how spread out the distribution of all possible random sample means would be. Here it's calculated as:

$$SE_M = \frac{SD}{\sqrt{N}} = \frac{3.117}{\sqrt{8}} = 1.102$$

This section provides a confidence interval around (centered on) the "Mean." Calculation requires the appropriate critical value. Specifically, the *t* statistic (with 7 *df*) that has a probability of .05 equals 2.365. As a result:

$$CI_M = M \pm (t_{CRITICAL})(SE_M)$$

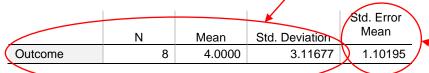
= 4.000 ± (2.365)(1.102)

Thus, the researcher estimates that the true population mean is somewhere between 1.394 and 6.606 (knowing that the estimate could be incorrect).

One Sample t Test

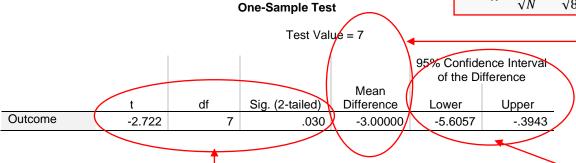
These values of the one-sample statistics are identical to the values that would be provided by the "Frequencies" or "Descriptives" commands. See the earlier annotated output for details of how these are computed from frequency distributions.

One-Sample Statistics



The "Standard Error of the Mean" provides an estimate of how spread out the distribution of all possible random sample means would be. Here it's calculated as:

$$SE_M = \frac{SD}{\sqrt{N}} = \frac{3.117}{\sqrt{8}} = 1.102$$



The "Mean Difference" is the difference between the sample mean (M = 4) and the user-specified test value (u = 7). For the example, the sample had a mean one point higher than the test value. This raw effect size is important for both the significance test and the confidence interval.

The "t", "df", and "Sig." columns provide the results of the statistical significance test. First, *t* provides the standardized statistic for the mean difference:

$$t = \frac{M - \mu}{SE_M} = \frac{-3.000}{1.102} = -2.722$$

The t statistic follows a non-normal (studentized or t) distribution that depends on degrees of freedom. Here, df = N - 1 = 8 - 1 = 7. A t with 7 df that equals -2.722 has a two-tailed probability (p) of .030, a statistically significant finding.

This section provides a confidence interval around (centered on) the "Mean Difference." Calculation requires the appropriate critical value. Specifically, the *t* statistic (with 7 *df*) that has a probability of .05 equals 2.365. As a result:

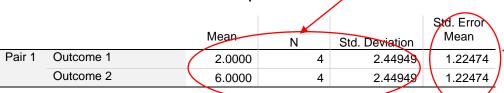
$$CI_{DIFF} = M_{DIFF} \pm (t_{CRITICAL})(SE_M) = -3.000 \pm (2.365)(1.102)$$

Thus, the researcher estimates that the true population mean difference is somewhere between -5.606 and -.394 (knowing that the estimate could be incorrect).

Paired Samples t Test

These values of the group statistics are calculated separately for each variable.

Paired Samples Statistics

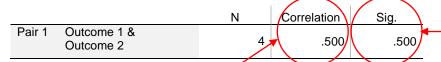


These are the standard errors for each variable calculated separately. For the first variable:

$$SE_M = \frac{SD}{\sqrt{N}} = \frac{2.449}{\sqrt{4}} = 1.225$$

Notice that the standard errors are equal because the variables have the same standard deviation.

Paired Samples Correlations



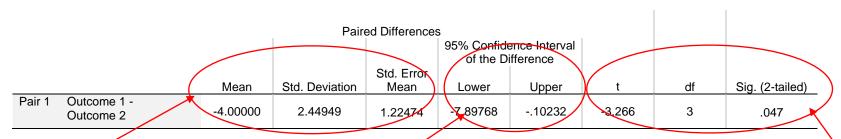
This is the correlation between the scores of the two variables. This correlation is not easily determined from the summary statistics of the output, but rather from the data (and the calculations are therefore not shown here).

Though the statistic is not shown, *t* provides the standardized statistic for testing whether the correlation differs from zero:

$$t = \frac{r}{\sqrt{(1 - r^2)/(N - 2)}} = \frac{.500}{\sqrt{(1 - .500^2)/(4 - 2)}} = .816$$

The t statistic follows a non-normal (studentized or t) distribution that depends on degrees of freedom. Here, df = N - 2 = 4 - 2 = 2. A t with 2 df that equals .816 has a two-tailed probability (p) of .500, which is not a statistically significant finding.

Paired Samples rest



The "Paired Differences" statistics are determined by taking the differences of each person's pairs of scores on the two variables. Thus, the "Std. Deviation" of these is not determinable from the summary statistics. However, the "Mean" here is the difference between the two means provided above.

This confidence interval is centered on the "Mean" of the paired differences of the two variables. Calculation requires the appropriate critical value. Specifically, the *t* statistic (with 3 *df*) that has a probability of .05 equals 3.182. As a result:

$$CI_D = M_D \pm (t_{CRITICAL})(SE_D)$$

 $CI_D = -4.000 \pm (3.182)(1.225)$

Thus, the researcher estimates that the true population mean difference is somewhere between -7.898 and -.1023 (knowing that the estimate could be incorrect).

The "t", "df", and "Sig." columns provide the results of the statistical significance test. First, *t* provides the standardized statistic for the mean difference:

$$t = \frac{M_D}{SE_D} = \frac{-4.000}{1.225} = -3.266$$

The t statistic follows a non-normal (studentized or t) distribution that depends on degrees of freedom. Here, df = N - 1 = 4 - 1 = 3. A t with 3 df that equals -3.266 has a two-tailed probability (p) of .047, a statistically significant finding.

Independent Samples t Test

These values of the group statistics are calculated separately for each group. They are not identical to the values obtained from analyzing the variable as a whole.

Group Statistics

						Std. Error	
	Lab Group	N		Mean	Std. Deviation	/ Mean	\
Outcome	Level 1		4	2.0000	2.44949	1.22474	•
	Level 2		4	6.0000	2.44949	1.22474	
						$\overline{}$	

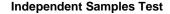
Levene's Test for

These are the standard errors for each mean separately.

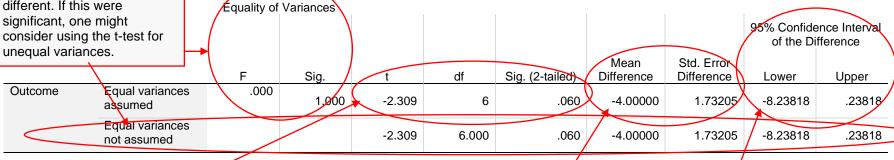
$$SE_M = \frac{SD}{\sqrt{N}} = \frac{2.449}{\sqrt{4}} = 1.225$$

Notice that the standard errors are equal because both groups have the same standard deviation and sample size.

"Levene's Test" determines whether the variability from the two groups is significantly different. If this were significant, one might consider using the t-test for



t-test for Equality of Means



The "t", "df", and "Sig." columns provide the results of the statistical significance test. First, t provides the standardized statistic for the mean difference:

$$t = \frac{M_1 - M_2}{SE_{DIFF}}$$
$$t = \frac{-4.000}{1.732} = -2.309$$

The t statistic follows a non-normal (studentized or t) distribution that depends on degrees of freedom. Here, df = N - 2 = 8 - 2 = 6. A t with 6 df that equals -2.309 has a two-tailed probability (p) of .060, a finding that is not statistically significant.

The "Mean Difference" is the difference between the two group means. For the example, group one's mean was 4 points lower.

The "Standard Error of the Difference" is a function of the two groups' individual standard errors. When sample sizes are equal:

$$SE_{DIFF} = \sqrt{SE_1^2 + SE_2^2}$$

 $SE_{DIFF} = \sqrt{1.225^2 + 1.225^2} = 1.732$

This value is important for both the significance test and the confidence interval. [Importantly, the computation of the standard error of the difference is more complex for unequal sample sizes.]

This section provides a confidence interval around (centered on) the "Mean Difference." Calculation requires the appropriate critical value. Specifically, the t statistic (with 6 df) that has a probability of .05 equals 2.447. As a result:

$$CI_{DIFF} = M_{DIFF} \pm (t_{CRITICAL})(SE_{DIFF})$$

 $CI_{DIFF} = -4 \pm (2.447)(1.732)$

Thus, the researcher estimates that the true population mean difference is somewhere between -8.238 and 0.238 (knowing that the estimate could be incorrect).

OneWay ANO

These values of the group statistics are calculated separately for each group. They are not identical to the values obtained from analyzing the variable as a whole.

These are the standard errors for each mean separately. For example, for the first group.

95% Confidence Interval for

$$SE_M = \frac{SD}{\sqrt{N}} = \frac{2.449}{\sqrt{4}} = 1.225$$

Outcome

					Me				
	N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound	Minimum	Maximum	
Level 1	4	2.0000	2.44949	1.22474	-1.8977	5.8977	.00	5.00	١
Level 2	4	6.0000	2.44949	1.22474	2.1023	9.8977	4.00	9.00	
Level 3	4	7.0000	2,44949	1.22474	3.1023	10,8977	4.00	9.00	
Total	12	5.0000	3.16228	.91287	2.9908	7.0092	.00	9.00	\geq
		•		•					

"Minimum" and "Maximum" values are the lowest and highest scores in each group.

These "Total" values are all calculated for the set of data as a whole (i.e., not separately for each group). Because the mean and standard deviation are the same for each of the group values, the SE and CI will also be the same:

$$SE_M = \frac{SD}{\sqrt{N}} = \frac{3.162}{\sqrt{12}} = 0.913$$

 $CI_M = M \pm (t_{CRITICAL})(SE_M) = 5.000 \pm (2.179)(0.913)$

Thus, the researcher estimates that the true population grand (or overall) mean is somewhere between 2.991 and 7.009 (knowing it may not).

This section provides a confidence interval around (centered on) each mean separately. Calculation requires the appropriate critical value. Specifically, the *t* statistic (with 3 *df*) that has a probability of .05 equals 3.182. For example, for the first group:

$$CI_M = M \pm (t_{CRITICAL})(SE_M) = 2.000 \pm (3.182)(1.225)$$

Thus, the researcher estimates that the true population mean is somewhere between -1.898 and 5.898 (knowing it may not).

ANOVA

O			

	Sum of				-	
	Squares	df	Mean Square	λ		Sig.
Between Groups /	56.000	2	28.000		4.667	.041
Within Groups	54.000	9	6.000	Х		
Total	110.000	11				

"Mean Squares" are estimates of the variance for each source. For each in turn:

$$MS_{BETWEEN} = \frac{SS_{BETWEEN}}{df_{BETWEEN}} = \frac{56.000}{2} = 28.000$$

 $MS_{WTIHIN} = \frac{SS_{WITHIN}}{df_{WITHIN}} = \frac{54.000}{9} = 6.000$

"Within Groups" statistics are a function of the group variabilities. Because SS for each group is calculable ($SS = SD^2 \times df$):

$$SS_{WITHIN} = SS_1 + SS_2 + SS_3$$

 $SS_{WITHIN} = 18 + 18 + 18 = 54.0$
 $df_{WITHIN} = df_1 + df_2 + df_3 = 9$

"Between Groups" statistics are a function of the group means and sample sizes:

$$SS_{BETWEEN} = \sum n_{GROUP} (M_{GROUP} - M_{TOTAL})^2$$

 $SS_{BETWEEN} = 4(2-5)^2 + 4(6-5)^2 + 4(7-5)^2$
 $SS_{BETWEEN} = 56.000$
 $df_{BETWEEN} = \#groups - 1 = 2$

The "F" statistic is a ratio of the between and within group variance estimates:

$$F = \frac{MS_{BETWEEN}}{MS_{WITHIN}} = \frac{28.000}{6.000} = 4.667$$

An *F* with 2 and 9 *df* that equals 4.667 has a two-tailed probability (*p*) of .041, a statistically significant finding.

Post Hoc Comparisons

Tukey's HSD procedure is appropriate for post-hoc pairwise comparisons between groups. The output lists all possible pairwise comparisons, including those that are redundant.

Dependent Variable: Outcome

/	Tukey HSD		Mean Difference			95% Confide	ence Interval
	(I) Factor	(J) Factor	(I-J)	Std. Error	Sig.	Lower Bound	Upper Bound
	Level 1	Level 2	-4.00000	1.73205	.106	-8.8359	.8359
		Level 3	-5.00000*	1.73205	.043	-9.8359	1641
	Level 2	Level 1	4.00000	1.73205	.106	8359	8.8359
/		Level 3	-1.00000	1.73205	.835	-5.8359	3.8359
	Level 3	Level 1	5.00000	1.73205	.043	.1641	9.8350
		Level 2	1.00000	1.73205	.835	-3.8359	5.8359
	4 TI 1100			\ / W	· /		

^{*} The mean difference is significant at the .05 level.

"Mean Difference (I-J)" is the difference between the means for the "I" and "J" groups. Even though half of the listed comparisons are redundant, the mean differences will have the opposite signs because of subtraction order. This will also change the signs of the associated confidence intervals.

Homogeneous Subsets

These "Standard Errors" are for the difference between the two group means. The values are a function of the MS_{WITHIN} (from the ANOVA) and the sample sizes:

$$SE_{DIFF} = \sqrt{\left(\frac{MS_{WITHIN}}{n_{GROUP}}\right) + \left(\frac{MS_{WITHIN}}{n_{GROUP}}\right)}$$
$$SE_{DIFF} = \sqrt{\left(\frac{6.000}{4}\right) + \left(\frac{6.000}{4}\right)} = 1.732$$

In this case, because all groups are of the same size, the standard error for each comparison is the same.

The "Sig." column provides the probability of the HSD statistic (which is not listed). The HSD statistic is a function of the "Mean Difference" and the "Std. Error". For the first comparison in the example:

$$HSD = \frac{M_1 - M_2}{SE_{DIFF}} = \frac{-4.000}{1.732} = 2.309$$

An HSD of 2.309 (with 2 *df*_{BETWEEN} and 9 *df*_{WITHIN} like in the ANOVA source table) has a two-tailed probability (*p*) of .106, which is not a statistically significant finding.

This section provides confidence intervals around (centered on) the "Mean Differences." Calculation requires the appropriate critical value. Specifically, the HSD statistic (with 2 *df*_{BETWEEN} and 9 *df*_{WITHIN}) that has a probability of .05 equals 3.068. For the first comparison in the example:

$$CI_{DIFF} = M_{DIFF} \pm (HSD_{CRITICAL})(SE_{DIFF})$$

 $CI_{DIFF} = -4.000 \pm (2.792)(1.732)$

Thus, the estimates that the true population mean difference is somewhere between -8.836 and 0.836 (knowing that the estimate could be incorrect).

Tukey HSD

,		Subset for alpha = .05			
Factor	N	1	2		
Level 1	4/	2.0000			
Level 2	4	6.0000	6.0000		
Level 3	4		7.0000		
Sig.		.106	.835		

Means for groups in homogeneous subsets are displayed.

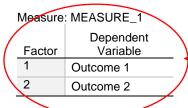
a. Uses Harmonic Mean Sample Size = 3.000.

"Homogeneous Subsets" provide groupings for the means. Means within the same subset are not significantly different from each other (note the "Sig." value at the bottom of the column for the subset). This offers a useful summary of the comparisons as analyzed above.

Repeated Measures ANOVA

(Note that some aspects of this output have been deleted and rearranged for the sake of presentation!)

Within-Subjects Factors



This provides a description of the variable levels (i.e., columns in the data set) that are linked by being separate instances of the dependent variable.

Descriptive Statistics

	Mean	Std. Deviation	N
Outcome 1	2.0000	2.44949	4
Outcome 2	6.0000	2.44949	4

Estimated Marginal Means

Factor

	Measure	MEASURE_	1		
				95% Confide	ence Interval
/	Factor	Mean	Std. Error	Lower Bound	Upper Bound
	1	2.000	1.225	-1.898	5.898
`	2	6.000	1.225	2.102	9.898

These values of the descriptive statistics are calculated separately for each level or condition of the within-subjects factor. They are identical to what would be obtained if the "Frequencies" or "Descriptives" procedure had been used separately for each.

Because sample sizes are equal, a grand mean can be determined by averaging these two level means:

$$M_{TOTAL} = (M_{LEVEL} + M_{LEVEL})/2 = (2.000 + 6.000)/2 = 4.000$$

This section provides a confidence interval around (centered on) each condition's mean separately. Calculation requires the appropriate critical value. Specifically, the *t* statistic (with 3 *df*) that has a probability of .05 equals 3.182. For example, for the first time:

$$CI_M = M \pm (t_{CRITICAL})(SE_M) = 4.000 \pm (3.182)(1.225)$$

Thus, the researcher estimates that the true population mean for the measure for the Outcome 1 is somewhere between -1.898 and 5.898 (knowing that the estimate could be incorrect).

Tests of Between-Subjects Effects

Measure: MEASURE_1

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	
Intercept	128.000	1	128.000	14.222	.033	*
Error	27.000	3	9.000	\rightarrow		

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source

Source	C		Type III Sum				tne degrees d
			of Squares	df	Mean Square	F	Sig.
Factor	r	Sphericity Assumed	32.000	1	32.000	10.667	.047
		Greenhouse-Geisser	32.000	1.000	32.000	10.667	.047
		Huynh-Feldt	32.000	1.000	32.000	10.667	.047
		Lower bound	32.000	1.000	32.000	10.667	.047
Error(1	time)	Sphericity Assumed	9.000	3	3.000		
1	_	Greenhouse-Geisser	9.000	3.000	3.000		
		Huynh-Feldt	9.000	3.000	3.000		
	_ `	Lower-bound	9.000	3.000	3.000		
							1

These rows provide statistics adjusted for the "Sphericity" test (not shown). Because that test showed no violation of the assumption, these statistics show the exact same results as those in which sphericity is assumed.

The "Within-Subjects Error" is a function of variabilities of the separate levels or conditions of the factor and the "between-subjects error" given above. Because SS for each level can be determined ($SS = SD^2 \times df$, which equals 18.000 for each of the two outcomes):

$$SS_{ERROR} = SS_1 + SS_2 - SS_{SUBJECTS}$$

$$SS_{ERROR} = 18.000 + 18.000 - 27.000 = 9.000$$

$$df_{ERROR} = df_1 + df_2 - df_{SUBJECTS} = 3 + 3 - 3 = 3.000$$

The "Mean Square" is the usual ratio of the Sum of Squares to the degrees of freedom.

The "Between-Subjects Intercept" here refers to the average score of the participants in the study and the significance test determines whether that average is different from zero. This is often not an informative test.

"Between-Subjects Error" refers to the average differences across the participants of the study. This Sum of Squares is not easily determined from the summary statistics output, but rather from the data (and the calculations are therefore not shown here). However:

$$df_{SUBJECTS} = \#subjects - 1 = 3$$

The "Mean Square" is the usual ratio of the Sum of Squares to the degrees of freedom.

The statistics for the effect (or change) on the "Factor" are functions of the means of the levels or conditions and the sample sizes:

$$SS_{EFFECT} = \sum_{LEVEL} (M_{LEVEL} - M_{TOTAL})^2$$

 $SS_{EFFECT} = 4(2.0 - 4.0)^2 + 4(6.0 - 4.0)^2$
 $SS_{EFFECT} = 32.000$
 $df_{EFFECT} = \#levels - 1 = 1$

The "Mean Square" is the usual ratio of the Sum of Squares to the degrees of freedom.

$$MS_{EFFECT} = \frac{SS_{EFFECT}}{df_{EFFECT}} = \frac{32.000}{1} = 32.000$$

The "F" statistic is a ratio of the effect and within-subjects error variance estimates:

$$F = \frac{MS_{EFFECT}}{MS_{FRROR}} = \frac{32.000}{3.000} = 10.667$$

An *F* with 1 and 3 *df* that equals 10.667 has a two-tailed probability of .047, a statistically significant finding.

Factorial ANOVA

Descriptive Statistics

Dependent Variable: Outcome

	FactorA	FactorB	Mean	Std. Deviation	N	
_	Level 1	Level 1	2.0000	2.44949	4	,
		Level 2	7.0000	2.44949	4	7
<		Total	4.5000	3.50510	8	>4
	Level 2	Levei 1	6.0000	2.44949	4	
		Level 2	5.0000	2.44949	4	>
<		Total	5.5000	2.32993	8	\geq
_	Total	Level 1	4.0000	3.11677	8	
\		Level 2	6.0000	2.50713	8	<u>ر</u>
<		Total	5.0000	2.92119	16	اح

528.000

These descriptive statistics are calculated separately for each condition as defined by the factors. They are not identical to the values obtained from analyzing the variable as a whole.

These descriptive statistics are calculated separately for each factor. They represent the marginal means of one factor collapsing across the levels of the other factor. They are not identical to the values obtained from analyzing the variable as a whole.

These descriptive statistics represent the grand (or overall) values obtained from analyzing the variable as a whole. They are identical to what would be obtained if the "Frequencies" or "Descriptives" procedure had been used.

Dependent Variable: Outcome

Total

Corrected Total

Dependent variable.	not intol	mative.			
Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	56.000 ^a	3	18.667	3.111	.067
Intercept	400.000	1	400.000	66.667	.000
Factor A	4.000	1	4.000	.667	.430
Factor B	16.000	1	16.000	2.667	.128
Factor A * Factor B	36.000	1	36.000	6.000	.031
Error	72.000	12	6.000		

The "Corrected Model" statistics reflect the overall between-group variability. They are a function of the group means and sample sizes.

$$\begin{split} SS_{MODEL} &= \sum n_{GROUP} (M_{GROUP} - M_{TOTAL})^2 \\ SS_{MODEL} &= 4(2.000 - 5.000)^2 + 4(7.000 - 5.000)^2 + \\ 4(6.000 - 5.000)^2 + 4(5.000 - 5.000)^2 \\ SS_{MODEL} &= 56.000 \\ df_{MODEL} &= \#groups - 1 = 3 \end{split}$$

The "Factor A * Factor B" (interaction) statistics reflect the 128.000 between- group variability not accounted for by the factors a. R Squared = .438 (Adjusted R Squared = .297) taken individually:

$$SS_{INTER} = SS_{MODEL} - SS_{FACTORA} - SS_{FACTORB}$$

 $SS_{INTER} = 56.000 - 4.000 - 16.000 = 36.000$
 $df_{INTER} = df_{MODEL} - df_{FACTORA} - df_{FACTORB} = 1$

The "Intercept"

statistics are generally

The "Factor A" and "Factor B" statistics are a function of the level (marginal) means and sample sizes. For "Factor B":

$$SS_{FACB} = \sum_{ACTORB} n_{LEVEL} (M_{LEVEL} - M_{TOTAL})^2$$

$$SS_{FACTORB} = 8(4-5)^2 + 8(6-5)^2$$

$$SS_{FACTORB} = 16.000$$

 $df_{FACTORB} = \#levels - 1 = 1$

"Error" statistics are a function of the within group variabilities. Because SS for each group can be determined ($SS = SD^2 x df$):

$$SS_{ERROR} = SS_1 + SS_2 + SS_3 + SS_4$$

 $SS_{ERROR} = 18.000 + 18.000 + 18.000 + 18.000 = 72.000$
 $df_{ERROR} = df_1 + df_2 + df_3 + df_4 = 12$

"Mean Squares" are estimates of the variances associated with each source. For the "Factor A * Factor B" interaction:

$$MS_{INTER} = \frac{SS_{INTER}}{df_{INTER}} = \frac{36.000}{1} = 36.000$$

The "F" statistic is a ratio of the effect and within group (error) variance estimates. For the "Factor A * Factor B" interaction:

$$F_{INTER} = \frac{MS_{INTER}}{MS_{ERROR}} = \frac{36.000}{6.000} = 6.000$$

An F with 1 and 12 df that equals 6.000 has a two-tailed probability of .031, which is a statistically significant finding.