

# **Week 11: Data Analysis and Interpretation Part I**

**Dr. T. Kody Frey**

Assistant Professor | School of Information Science

# Housekeeping

## Project Timeline

Only 6 weeks left! 1 group has been approved by IRB

- Goal is to launch survey within the next 2 weeks
- Final Qualtrics survey needs to be completed by 3/27. We will prepare them to launch during the workshop.
- Data collection will last *two weeks*
- Group should be working on *front end* (literature review, hypotheses, methods)

## What Do You Know Now?

- Research Design

- Internal Validity
- External Validity
- Measurement Validity

You can learn *a lot* about a quantitative study using just these!

## What Do We Still Need To Learn?

- Analyzing Group differences
- Analyzing Relationships

## What's Next?

Week 12	Data Analysis and Interpretation II
3/27	<p>In contrast to our overview of understanding group differences last week, the second sequence in our overview of data analysis focuses on relationships and associations between variables. We will closely examine what differentiates the various types of analyses, in addition to providing you with a basic understanding of correlations and linear regression techniques.</p> <p><b>Workshop Topic:</b> Connecting Surveys to CI SONA</p> <p><b>Readings &amp; Podcast</b></p> <p>*Gliner, J. A., Morgan, G. A., &amp; Leech, N. L. (2017). <i>Research methods in applied settings: An integrated approach to design and analysis</i> (3<sup>rd</sup> ed.). Routledge. *Read Chs. 19, 21*</p> <p>Quantitude Podcast: S3E09: Semi-Partially Clarifying Measures of Association in Regression. <a href="https://quantitudepod.org/s3e09-semi-partially-clarifying-measures-of-association-in-regression/">https://quantitudepod.org/s3e09-semi-partially-clarifying-measures-of-association-in-regression/</a></p> <p>Shulman, H. C., Dixon, G. N., Bullock, O. M., &amp; Colón Amill, D. (2020). The effects of jargon on processing fluency, self-perceptions, and scientific engagement. <i>Journal of Language and Social Psychology</i>, 39(5-6), 579-597. <a href="https://doi.org/10.1177/0261927X20902177">https://doi.org/10.1177/0261927X20902177</a></p>
	<p>Sherblom, J. C., &amp; Umphrey, L. R. (2023). The social cognition of hope. <i>Communication Research Reports</i>, 0(0), 1-10. <a href="https://doi.org/10.1080/08824096.2022.2164265">https://doi.org/10.1080/08824096.2022.2164265</a></p>
	<p><b>Final Survey in Qualtrics Due</b></p>
	<p><b>Reaction Paper #8</b></p>

## Overview of Today

- Research Design and Classification
- *t* tests
- Analyses of variance (ANOVA)
- Complex tests of group differences
- Discussion: Group Differences in Practice
- Workshop: Midterm Review / Workday

## GLM Ch. 18

*General Design Classifications for Selection of Difference Statistical Methods*

## Key Terms

### Between Group Designs

### Within Subjects Designs

### Mixed Designs

- Each participant in the research is in *one and only one condition* or group
- Participants receive only one level of the IV
- DV only needs to be measured once
- If power analysis says 20 in each group, then you need 60 total
- Each participant receives or experiences *ALL levels of the IV*
- If power analysis says 20 in each group, then you need 20 total
- Appeal is in need for fewer participants acting as own controls
- Impacted by *carryover effects*
- Has at least one between-groups IV and one within-subjects IV
- Has a minimum of two IVs

## Understanding Design Types

It's all about the *levels* of the IV

Designs are usually described in terms of:

- the general type of design (between groups, within subjects, or mixed)
- the number of independent variables
- the number of levels within each independent variable.

## What does this mean?

The researcher will typically state the design:

EX: A single factor design with 3 levels

EX: A 2 x 2 design with repeated measures

EX: A 3 x 3 x 2 factorial design

## Single Factor Designs

Note: *Factor* is another name for IV

Single factor designs only have one IV

EX: The effect of time of class (morning, afternoon, night) on student engagement.

EX: Students' perceptions of instructor effectiveness over time.

## Class Example

Ikeola and Marian are examining campaigns using visual images can influence HPV vaccine behavioral intentions and eventual uptake of the HPV vaccine.

- IV: Visual message (4 levels: threat, efficacy, threat x efficacy, control)
- DV: Behavioral intentions

## Between-Groups Factorial Designs

Includes more than one IV

EX: The effect of time of class (morning, afternoon, night) and student sex (male, female) on student engagement.

This is called a 3 x 2 factorial design.

- The **numbers** represent the number of levels (3 levels of time of class and 2 levels of sex).
- The **number of numbers** represents the number of IVs (time of class and sex)

## Think Fast!

If you have a  $4 \times 2 \times 3$  factorial design...

How many IVs do you have?

What are the levels of each IV?

How many groups would you need for the study?

## Class Example

Elizabeth, Madelyn, and Ansley are examining how message sender status on Instagram influences behavioral intention / adoption.

- IV: Status (3 levels: influencer, layperson, control)
- DV: Behavioral intention

This is currently a *single-subject* design. How would you make it a between-subjects design?

## Within-Subjects Factorial Designs

Two or more IVs

Participants experience all conditions

EX: The effect of teaching style (traditional or inquiry based) on student engagement over time (pretest and posttest)

## Example

To assess the CIS courses in the School of Information Science, we collect data from all course sections in August, December, January, and April to see if participation in course influences public speaking self-efficacy.

EX: The effect of class time on public speaking self-efficacy over time.

## Mixed Designs

“Such a design might have two between-groups independent variables with three and four levels, respectively, and have one within-subjects independent variable with two levels.

It would be described as a  $3 \times 4 \times 2$  factorial design with repeated measures on the third factor.”

Basically, if it has a BG component and a WS component, it is a mixed design!

## Summing it Up

**TABLE 18.1****Examples of General Design Classifications**

<b>Single factor</b>	<b>One independent variable</b>
Between	single factor design with ___ levels
Within	single factor repeated-measures design with ___ levels
Mixed	NA
<b>Two factor</b>	<b>Two independent variables</b>
Between	___ x ___ factorial design
Within	___ x ___ design with repeated measures on both factors
Mixed	___ x ___ (mixed) design with repeated measures
<b>Three factor</b>	<b>Three independent variables</b>
Between	___ x ___ x ___ factorial design
Within	___ x ___ x ___ design with repeated measures on all factors
Mixed	___ x ___ x ___ design with repeated measures on last (or last two) factors

*Note:* The dependent variable is not part of the design classification and, thus, is not mentioned. The number of levels for *an* independent variable is inserted in each blank.

If you recognize the design, you can generally determine the type of test needed.

## GLM Ch. 19

*The design classification ultimately drives the appropriate statistics.*

# Assumptions

Ask Yourself

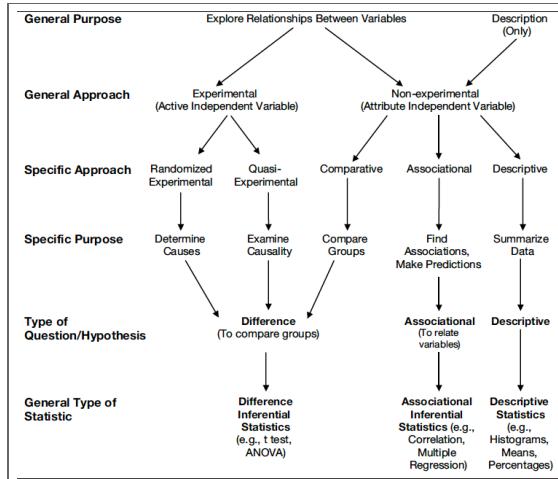
Statistical Assumptions

Remember!

- Is it difference, associational, or descriptive?
- What are the IVs and DVs?
- How many IVs do you have?
- What are the levels of the IV?
- Is it a between, within, or mixed classification?
- What is the level of measurement for each variable?

Three major statistical assumptions:

- DV comes from population that is *normally distributed*
- Variance of groups must be equal (*homogeneity of variance*)
- Independence of observations



## Choosing Your Test: Basic Differences

		Scale of Measurement of Dependent Variable ↓	COMPARE ↓	One Factor or Independent Variable with 2 Categories or Levels/Groups/Samples		One Independent Variable 2 or More Categories or Levels or Groups	
				Independent Samples or Groups (Between)	Repeated Measures or Related Samples (Within)	Independent Samples or Groups (Between)	Repeated Measures or Related Samples (Within)
Parametric Statistics	Dependent Variable Approximates Normal Distribution Data and Assumptions Not Markedly Violated	MEANS	INDEPENDENT SAMPLES t TEST or ONEWAY ANOVA Ch. 20	PAIRED SAMPLES t TEST Ch. 20	ONEWAY ANOVA Ch. 20	REPEATED MEASURES ANOVA Ch. 20	
Non Parametric Statistics	Dependent Variable Clearly Ordinal (Ranked) Data or Assumptions Markedly Violated	MEDIANS OR RANKS	MANN-WHITNEY Ch. 20	WILCOXON or SIGN TEST Ch. 20	KRUSKAL-WALLIS Ch. 20	FRIEDMAN Ch. 20	
	Dependent Variable Nominal (Categorical) Data	COUNTS	CHI SQUARE Ch. 21 or FISHER'S EXACT TEST Ch. 20	MCNEMAR Ch. 20	CHI SQUARE Ch. 21	COCHRAN Q TEST	

*Notes:* To select the appropriate statistic, locate a box based on a) the type of question, b) design, and c) scale of measurement. It is acceptable to use statistics which are in the box(es) below the appropriate statistic, but there is usually some loss of information and power. It is not acceptable to use statistics above the appropriate box.

- Related samples designs are also called repeated measures, matched or paired groups and are within-subjects designs
- Chi square tests for the independence of two variables. Frequency data or counts of the number of Ss in each cell or category are used rather than raw scores and means.
- ANOVA is Analysis of Variance.

## Choosing Your Test: Complex Differences

Dependent Variable(s) ↓	Two or More Independent Variables		
	All Between Groups	All Within Subjects	Mixed (Between & Within)
One Normally Distributed Dependent Variable	Factorial ANOVA Ch. 22	Factorial ANOVA with Repeated Measures on all Factors Ch. 22	Factorial ANOVA With Repeated Measures on Last or Last 2 Factors Ch. 22
Ordinal Dependent Variable	None Common	None Common	None Common
Nominal Dependent Variable	Log Linear	None Common	None Common
Several Normally Distributed Dependent Variables	MANOVA	MANOVA with Repeated Measures on all Factors	MANOVA With Repeated Measures on Last or Last Several Factors

## GLM Ch. 20

*Data Analysis and Interpretation: Basic Difference Questions*

## Disclaimer

We are going to go through A LOT of tests.

I'm not going to show you *how* to run each one (i.e., what buttons to press)

IMO - it is more important to know which test is *appropriate* and how to *interpret* it.

Once you figure out which test to run, there is a tutorial out there somewhere.

# Chi-Square Test for Goodness of Fit

Overview    Focus    Steps    Example    Write Up

## CHI-SQUARE

There are a number of different tests based on the chi-square statistic, all of which involve categorical data. For a review of chi-square, see Gravetter and Wallnau (2012).

### Chi-square test for goodness of fit

This test, which is also referred to as the one-sample chi-square, is often used to compare the proportion of cases from a sample with hypothesised values or those obtained previously from a comparison population. All that is needed in the data file is one categorical variable and a specific proportion against which you wish to test the observed frequencies. This may test that there is no difference in the proportion in each category (50%/50%) or a specific proportion obtained from a previous study.

**Example of research question:** I will test whether the number of smokers in the `survey.sav` data file is different to that reported in the literature from a previous larger nationwide study (20%).

#### What you need:

- one categorical variable, with two or more categories: smoker (Yes/No)
- a hypothesised proportion (20% smokers; 80% non-smokers or .2/.8).

**How does the proportion of smokers in our sample (observed) differ from the proportion of smokers (expected) in the population?**

**You need to know what to expect in the population (e.g., 20% are smokers; 80% are not)!**

**Example of research question:** I will test whether the number of smokers in the `survey.sav` data file is different to that reported in the literature from a previous larger nationwide study (20%).

### Procedure for chi-square test for goodness of fit

To follow along with this example, open the `survey.sav` data file.

1. From the menu at the top of the screen, click on **Analyze**, then select **Nonparametric Tests**, then **Legacy Dialogs** and then **Chi-Square**.
2. Click on the categorical variable (smoker) and click on the arrow to move it into the **Test Variable List** box.  
In the **Expected Values** section, click on the **Values** option. In the **Values** box, you need to type in two values.
  - The first value (.2) corresponds with the expected proportion for the first coded value for the variable (1=yes, smoker). Click on **Add**.
  - Type in the second value (.8), which is the expected proportion for the second coded value (2=no, non-smoker). Click on **Add**. If your variable has more than two possible values, you would need to type in as many proportions as appropriate.
3. Click on **OK** (or on **Paste** to save to **Syntax Editor**).

**Below is the syntax for running the procedure.**

The syntax from this procedure is:

```
NPAR TESTS
/CHISQUARE=smoke
/EXPECTED=.2 .8
/MISSING ANALYSIS.
```

A pink arrow points from the syntax to the output table.

The output generated from this procedure is shown below:

smoker

	Observed N	Expected N	Residual
1 YES	85	87.2	-2.2
2 NO	351	348.8	2.2
Total	436		

Test Statistics

	smoker
Chi-Square <sup>a</sup>	.069
df	1
Asymp. Sig.	.792

a. 0 cells (.0%) have expected counts less than 5. The minimum expected cell frequency is 87.2.

**Below is the sentence that goes in methods in the "participants" section.**

A pink star is placed next to the sentence. A blue arrow points from the sentence to the table.

A chi-square goodness-of-fit test indicates there was no significant difference in the proportion of smokers identified in the current sample (19.5%) as compared with the value of 20% that was obtained in a previous nationwide study,  $\chi^2 (1, n = 436) = .07, p = .79$ .

85/436 = 0.1949541284

### Chi-Square Tests

**What to Report**

- Degrees of freedom (df)
- Number of observations (N)
- Observed chi-square value ( $\chi^2$ )
- Significance level (p)
- Effect size (ES, varies with test)

*Supporting information  
(reported in text or table)*

- Number of observations per cell (n)
- Table of frequencies/percentages (optional)

**Suggested Syntax**

- $\chi^2 (df, N = XX) = \text{observed chi-square value, significance level, ES}$

**Suggested Format**

- A 2 × 3 chi-square test indicated that the relationship between gender and promotion was significant,  $\chi^2 (2, N = 112) = 13.45, p < .01, V = .29$ .

*Note:* The suggested syntax is similar for chi-square tests. This includes goodness-of-fit test, Fisher's exact test, and other chi-square tests. Briefly identifying the type of test is essential (e.g., "a Fisher's exact test revealed" or "a chi-square goodness-of-fit test indicated"). Frequencies and percentages can be reported within the text and/or a table. Tables are recommended, however, because they enhance reader comprehension.

# Chi-Square Test for Independence

## Overview    Focus    Steps    Example    Write Up

### Chi-square test for independence

This test is used when you wish to explore the relationship between *two* categorical variables. Each of these variables can have two or more categories. This test compares the observed frequencies or proportions of cases that occur in each of the categories, with the values that would be expected if there was no association between the two variables being measured. It is based on a crosstabulation table, with cases classified according to the categories in each variable (e.g. male/female; smoker/non-smoker).

When a 2 by 2 table (two categories in each variable) is encountered by IBM SPSS, the output from chi-square includes an additional correction value (*Yates' Correction for Continuity*). This is designed to compensate for what some writers feel is an overestimate of the chi-square value when used with a 2 by 2 table.

In the following procedure, using the *survey.sav* data file, I will demonstrate the use of chi-square using a 2 by 2 design. If your study involves variables with more than two categories (e.g. 2 by 3, 4 by 4), you will notice some slight differences in the output.

**Example of research question:** There are a variety of ways questions can be phrased: is there an association between gender and smoking behaviour? Are males more likely to be smokers than females? Is the proportion of males that smoke the same as the proportion of females?

**What you need:** Two categorical variables, with two or more categories in each:

- gender (Male/Female)
- smoker (Yes/No).

A Chi-Square Test of Independence is used to determine whether or not there is a significant association between two categorical variables.

How does the proportion of smokers who are female and male in our sample differ from the proportion of smokers in the population who are female and male?

### Procedure for chi-square test for independence

To follow along with this example, open the *survey.sav* data file.

1. From the menu at the top of the screen, click on **Analyze**, then **Descriptive Statistics**, and then **Crosstabs**.
2. Click on one of your variables (e.g. sex) to be your row variable and click on the arrow to move it into the box marked **Row(s)**.
3. Click on the other variable to be your column variable (e.g. smoker) and click on the arrow to move it into the box marked **Column(s)**.
4. Click on the **Statistics** button. Tick **Chi-square** and **Phi** and **Cramer's V**. Click on **Continue**.
5. Click on the **Cells** button.
  - In the **Counts** box, make sure there is a tick for **Observed**.
  - In the **Percentage** section, click on the **Row** box.
  - In the **Residuals** section click on **Adjusted standardized**.
6. Click on **Continue** and then **OK** (or on **Paste** to save to Syntax Editor).

The syntax from this procedure is:

```
CROSSTABS
/TABLES=sex BY smoke
/FORMAT=AVALUE TABLES
/STATISTICS=CHISQ PHI
/CELLS= COUNT ROW ASRESID
/COUNT ROUND CELL .
```

Selected output generated from this procedure for a 2 by 2 table is shown below. If your variables have more than two categories the printout will look a little different, but the key information that you need to look for in the output is still the same.

sex * smoker Crosstabulation					
		smoker			
		1 YES	2 NO	Total	
sex	1 MALES	Count	33	151	184
		% within sex	17.9%	82.1%	100.0%
		Adjusted Residual	-.7	.7	
	2 FEMALES	Count	52	200	252
		% within sex	20.6%	79.4%	100.0%
		Adjusted Residual	-.7	-.7	
Total		Count	85	351	436
		% within sex	19.5%	80.5%	100.0%

Chi-Square Tests					
	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2- sided)	Exact Sig. (1- sided)
Pearson Chi-Square	.494 <sup>a</sup>	1	.482		
Continuity Correction <sup>b</sup>	.337	1	.562		
Likelihood Ratio	.497	1	.481		
Fisher's Exact Test				.541	
Linear-by-Linear Association	.493	1	.483		
N of Valid Cases	436				.282

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 35.87.

b. Computed only for a 2x2 table

Symmetric Measures		
	Value	Approximate Significance
Nominal by Nominal	Phi	-.034
	Cramer's V	.034
N of Valid Cases	436	.482

A chi-square test for independence (with Yates' Continuity Correction) indicated no significant association between gender and smoking status,  $\chi^2 (1, n = 436) = .34, p = .56, \text{phi} = -.03$ .

## Chi-Square Tests

- What to Report**
- Degrees of freedom (df)
  - Number of observations (N)
  - Observed chi-square value ( $\chi^2$ )
  - Significance level (p)
  - Effect size (ES, varies with test)
- Supporting information (reported in text or table)*
- Number of observations per cell (n)
  - Table of frequencies/percentages (optional)

- Suggested Syntax**
- $\chi^2 (\text{df}, N = XX) = \text{observed chi-square value, significance level, ES}$

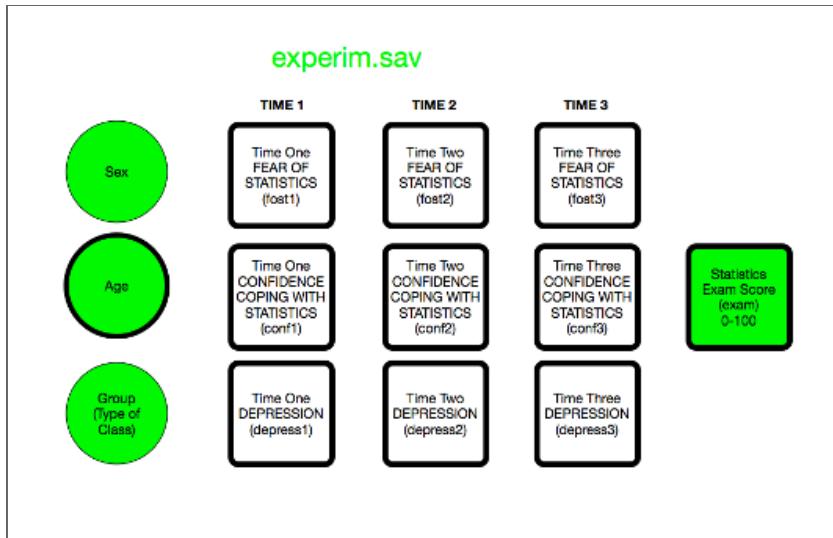
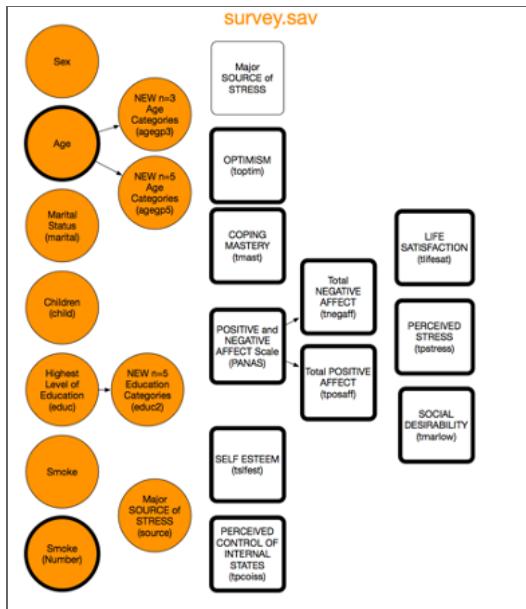
- Suggested Format**
- A 2  $\times$  3 chi-square test indicated that the relationship between gender and promotion was significant,  $\chi^2 (2, N = 112) = 13.45, p < .01, V = .29$ .

*Note:* The suggested syntax is similar for chi-square tests. This includes goodness-of-fit test, Fisher's exact test, and other chi-square tests. Briefly identifying the type of test is essential (e.g., "a Fisher's exact test revealed" or "a chi-square goodness-of-fit test indicated"). Frequencies and percentages can be reported within the text and/or a table. Tables are recommended, however, because they enhance reader comprehension.

## Example Data

### Survey Data

### Experiment Data



# t Test for Independent Samples

## Overview    Steps    Output    Example    Write Up

A statistical test conducted when there is an *independent variable with two levels* and a *scale/continuous dependent variable*.

Ex: Men will report higher levels of self esteem than women.

- Check information about groups
- Check assumptions
- Assess differences between groups
- Calculate effect size

T-TEST																																																	
GROUPS = sex(1 2)																																																	
/MISSING = ANALYSIS																																																	
/VARIABLES = tsselfest																																																	
/CRITERIA = CI(.95).																																																	
The output generated from this procedure is shown below.																																																	
Group Statistics																																																	
<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>SEX</th> <th>N</th> <th>Mean</th> <th>Std. Deviation</th> <th>Std. Error Mean</th> </tr> </thead> <tbody> <tr> <td>MALES</td> <td>184</td> <td>34.02</td> <td>4.91</td> <td>.36</td> </tr> <tr> <td>FEMALES</td> <td>252</td> <td>33.17</td> <td>5.71</td> <td>.36</td> </tr> </tbody> </table>							SEX	N	Mean	Std. Deviation	Std. Error Mean	MALES	184	34.02	4.91	.36	FEMALES	252	33.17	5.71	.36																												
SEX	N	Mean	Std. Deviation	Std. Error Mean																																													
MALES	184	34.02	4.91	.36																																													
FEMALES	252	33.17	5.71	.36																																													
Independent Samples Test																																																	
<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th rowspan="2"></th> <th colspan="2">Levene's Test for Equality of Variances</th> <th colspan="5">t-test for Equality of Means</th> </tr> <tr> <th>F</th> <th>Sig.</th> <th>t</th> <th>df</th> <th>Sig. (2-tailed)</th> <th>Mean Difference</th> <th>Std. Error Difference</th> <th>95% Confidence Interval of the Difference</th> </tr> </thead> <tbody> <tr> <td>Total self-esteem</td> <td>3.500</td> <td>.002</td> <td>1.622</td> <td>434</td> <td>.105</td> <td>.85</td> <td>.52</td> <td>.-16 1.87</td> </tr> <tr> <td>Equal variances assumed</td> <td></td> <td></td> <td>1.661</td> <td>422.349</td> <td>.098</td> <td>.85</td> <td>.51</td> <td>.-16 1.85</td> </tr> <tr> <td>Equal variances not assumed</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> </tbody> </table>								Levene's Test for Equality of Variances		t-test for Equality of Means					F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	Total self-esteem	3.500	.002	1.622	434	.105	.85	.52	.-16 1.87	Equal variances assumed			1.661	422.349	.098	.85	.51	.-16 1.85	Equal variances not assumed								
	Levene's Test for Equality of Variances		t-test for Equality of Means																																														
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference																																									
Total self-esteem	3.500	.002	1.622	434	.105	.85	.52	.-16 1.87																																									
Equal variances assumed			1.661	422.349	.098	.85	.51	.-16 1.85																																									
Equal variances not assumed																																																	

An independent-samples t-test was conducted to compare the self-esteem scores for males and females. There was no significant difference in scores for males ( $M = 34.02$ ,  $SD = 4.91$ ) and females ( $M = 33.17$ ,  $SD = 5.71$ ;  $t(434) = 1.62$ ,  $p = .11$ , two-tailed). The magnitude of the differences in the means (mean difference = .85, 95% CI: -1.80 to 1.87) was very small (eta squared = .006).

**t Test****What to Report**

- Degrees of freedom (df)
- Observed *t* value (*t*)
- Significance level (*p*)
- Effect size (ES, varies with test)

*Supporting information  
(reported in text or table)*

- Number of observations (*N*)
- Number of observations per cell (*n*)
- Descriptive statistics (*M* and *SD*)

**Suggested Syntax**

- $t$  (df) = observed *t* value, significance level, ES

**Suggested Format**

- An independent groups *t* test revealed that Group 1 ( $M = 2.45$ ,  $SD = .49$ ) differed from Group 2 ( $M = 2.29$ ,  $SD = .45$ ) as predicted,  $t$  (df) = 3.98,  $p < .01$ ,  $\eta^2 = .05$ .

*Note:* The syntax is similar for one-group, independent, and correlated groups *t* tests. For that reason, it is important to indicate clearly which test was conducted with a statement such as "a one-group *t* test revealed." Also report an appropriate effect size measure; common effect size measures associated with *t* tests include eta-squared and *d*.

## Dependent (Paired Samples) t Test

### Overview    Steps    Example    Write Up

A statistical test conducted when there is a *within-subjects, independent variable with two levels* and a *scale/continuous dependent variable*.

H: Fear of Statistics will be lower at time two than at time one.

- Check information about cases
- Check assumptions
- Assess differences over time
- Calculate effect size

T-TEST				
PAIRS = fost1 WITH fost2 (PAIRED)				
/CRITERIA = CI(.95)				
/MISSING = ANALYSIS.				
<b>Paired Samples Statistics</b>				
	Mean	N	Std. Deviation	Std. Error Mean
Pair 1   fost1 fear of stats time1	40.17	30	5.160	.942
fost2 fear of stats time2	37.50	30	5.151	.940
<b>Paired Samples Correlations</b>				
	N	Correlation	Sig.	
Pair 1   fost1 fear of stats time1 &   fost2 fear of stats time2	30	.862	.000	
<b>Paired Samples Test</b>				
	Paired Differences			
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference
				Lower      Upper
Pair 1   fost1 fear of stats time1 -   fost2 fear of stats time2	2.67	2.708	.494	1.655      3.678
	t	df		Sig. (2-tailed)
				.000

A paired-samples t-test was conducted to evaluate the impact of the intervention on students' scores on the Fear of Statistics Test (FOST). There was a statistically significant decrease in FOST scores from Time 1 ( $M = 40.17, SD = 5.16$ ) to Time 2 ( $M = 37.5, SD = 5.15$ ),  $t(29) = 5.39, p < .001$  (two-tailed). The mean decrease in FOST scores was 2.67 with a 95% confidence interval ranging from 1.66 to 3.68. The eta squared statistic (.50) indicated a large effect size.

**Correlated Groups t Test** This brief example from a study examining the effect of two different types of metaphors contains all of the elements necessary for a correlated group (also referred to as paired samples or paired comparison) t test, although the standard deviations appear in a table. Also note that the suggested syntax was enclosed within parentheses. This varies from APA guidelines but is common for journals in marketing. Always consult the style manual most commonly used in your discipline or the journal you will be submitting your article to for information on the precise format you should adopt. However, this difference in style does not affect the substance of what is reported.

Hypothesis 1 predicted that concrete metaphors would be easier to understand than abstract metaphors. The results of a two-tailed paired samples t test supported this hypothesis, ( $t(102) = -4.43, p < .001, \eta^2 = .16$ ). Participants were more accurate in their interpretation of concrete metaphors ( $M = 2.45$ ) than abstract metaphors ( $M = 2.26$ ). (Morgan & Reichert, 1999, p. 6)



## Single-Factor (Oneway) ANOVA

Overview    Focus    Steps    Example    Write Up

1 IV, between groups, 2 or more levels

Normally-distributed DV

EX: Does total optimism differ by age groups?

\*You COULD do this by running a whole bunch of t tests, but you inflate changes for Type I error.

ONEWAY																																																							
toptim BY agegp3																																																							
/STATISTICS DESCRIPTIVES HOMOGENEITY BROWNFORSYTHE WELCH																																																							
/PLOT MEANS																																																							
/MISSING ANALYSIS																																																							
/POSTHOC = TUKEY ALPHA(.05).																																																							
<b>Descriptives</b>																																																							
Total Optimism																																																							
<table border="1"> <thead> <tr> <th rowspan="2"></th> <th rowspan="2">N</th> <th rowspan="2">Mean</th> <th rowspan="2">Std. Deviation</th> <th rowspan="2">Std. Error</th> <th colspan="2">95% Confidence Interval for Mean</th> <th rowspan="2">Minimum</th> <th rowspan="2">Maximum</th> </tr> <tr> <th>Lower Bound</th> <th>Upper Bound</th> </tr> </thead> <tbody> <tr> <td>1 18 - 29</td> <td>147</td> <td>21.36</td> <td>4.551</td> <td>.375</td> <td>20.62</td> <td>22.10</td> <td>7</td> <td>30</td> </tr> <tr> <td>2 30 - 44</td> <td>153</td> <td>22.10</td> <td>4.147</td> <td>.335</td> <td>21.44</td> <td>22.77</td> <td>10</td> <td>30</td> </tr> <tr> <td>3 45+</td> <td>135</td> <td>22.96</td> <td>4.485</td> <td>.386</td> <td>22.19</td> <td>23.72</td> <td>8</td> <td>30</td> </tr> <tr> <td>Total</td> <td>435</td> <td>22.12</td> <td>4.429</td> <td>.212</td> <td>21.70</td> <td>22.53</td> <td>7</td> <td>30</td> </tr> </tbody> </table>										N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum	Lower Bound	Upper Bound	1 18 - 29	147	21.36	4.551	.375	20.62	22.10	7	30	2 30 - 44	153	22.10	4.147	.335	21.44	22.77	10	30	3 45+	135	22.96	4.485	.386	22.19	23.72	8	30	Total	435	22.12	4.429	.212	21.70	22.53	7	30
	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum																																															
					Lower Bound	Upper Bound																																																	
1 18 - 29	147	21.36	4.551	.375	20.62	22.10	7	30																																															
2 30 - 44	153	22.10	4.147	.335	21.44	22.77	10	30																																															
3 45+	135	22.96	4.485	.386	22.19	23.72	8	30																																															
Total	435	22.12	4.429	.212	21.70	22.53	7	30																																															
<b>Test of Homogeneity of Variances</b>																																																							
Total Optimism																																																							
<table border="1"> <thead> <tr> <th>Levene Statistic</th> <th>df1</th> <th>df2</th> <th>Sig.</th> </tr> </thead> <tbody> <tr> <td>.746</td> <td>2</td> <td>432</td> <td>.475</td> </tr> </tbody> </table>									Levene Statistic	df1	df2	Sig.	.746	2	432	.475																																							
Levene Statistic	df1	df2	Sig.																																																				
.746	2	432	.475																																																				
<b>ANOVA</b>																																																							
Total Optimism																																																							
<table border="1"> <thead> <tr> <th></th> <th>Sum of Squares</th> <th>df</th> <th>Mean Square</th> <th>F</th> <th>Sig.</th> </tr> </thead> <tbody> <tr> <td>Between Groups</td> <td>179.069</td> <td>2</td> <td>89.535</td> <td>4.641</td> <td>.010</td> </tr> <tr> <td>Within Groups</td> <td>8333.951</td> <td>432</td> <td>19.292</td> <td></td> <td></td> </tr> <tr> <td>Total</td> <td>8513.021</td> <td>434</td> <td></td> <td></td> <td></td> </tr> </tbody> </table>										Sum of Squares	df	Mean Square	F	Sig.	Between Groups	179.069	2	89.535	4.641	.010	Within Groups	8333.951	432	19.292			Total	8513.021	434																										
	Sum of Squares	df	Mean Square	F	Sig.																																																		
Between Groups	179.069	2	89.535	4.641	.010																																																		
Within Groups	8333.951	432	19.292																																																				
Total	8513.021	434																																																					
<b>Robust Tests of Equality of Means</b>																																																							
Total Optimism																																																							
<table border="1"> <thead> <tr> <th></th> <th>Statistic<sup>a</sup></th> <th>df1</th> <th>df2</th> <th>Sig.</th> </tr> </thead> <tbody> <tr> <td>Welch</td> <td>4.380</td> <td>2</td> <td>284.508</td> <td>.013</td> </tr> <tr> <td>Brown-Forsythe</td> <td>4.623</td> <td>2</td> <td>423.601</td> <td>.010</td> </tr> </tbody> </table>										Statistic <sup>a</sup>	df1	df2	Sig.	Welch	4.380	2	284.508	.013	Brown-Forsythe	4.623	2	423.601	.010																																
	Statistic <sup>a</sup>	df1	df2	Sig.																																																			
Welch	4.380	2	284.508	.013																																																			
Brown-Forsythe	4.623	2	423.601	.010																																																			
a. Asymptotically F distributed.																																																							

Post Hoc Tests						
Multiple Comparisons						
		Dependent Variable: Total Optimism				
Tukey HSD						
(I) age 3 groups	(J) age 3 groups	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
1 18 - 29	2 30 - 44	-.744	.507	.308	-1.94	.45
	3 45+	-1.595*	.524	.007	-2.83	.36
2 30 - 44	1 18 - 29	.744	.507	.308	.45	1.94
	3 45+	-.851	.519	.230	-2.07	.37
3 45+	1 18 - 29	1.595*	.524	.007	.36	2.83
	2 30 - 44	.851	.519	.230	-.37	2.07

\* The mean difference is significant at the .05 level.

### Presenting the results from one-way between-groups ANOVA with post-hoc tests

A one-way between-groups analysis of variance was conducted to explore the impact of age on levels of optimism, as measured by the Life Orientation Test (LOT). Participants were divided into three groups according to their age (Group 1: 29yrs or less; Group 2: 30 to 44yrs; Group 3: 45yrs and above). There was a statistically significant difference at the  $p < .05$  level in LOT scores for the three age groups:  $F(2, 432) = 4.6, p = .01$ . Despite reaching statistical significance, the actual difference in mean scores between the groups was quite small. The effect size, calculated using eta squared, was .02. Post-hoc comparisons using the Tukey HSD test indicated that the mean score for Group 1 ( $M = 21.36, SD = 4.55$ ) was significantly different from Group 3 ( $M = 22.96, SD = 4.49$ ). Group 2 ( $M = 22.10, SD = 4.15$ ) did not differ significantly from either Group 1 or 3.

### Analysis of Variance (ANOVA)

- What to Report**
- Degrees of freedom (between and within) ( $df_B/df_W$ )
  - Observed  $F$  value ( $F$ )
  - Significance level ( $p$ )
  - Effect size (ES, varies with test)

- Supporting information (reported in text or table)*
- Number of observations ( $N$ )
  - Number of observations per cell ( $n$ )
  - Descriptive statistics ( $M$  and  $SD$ )
- Additional information for factorial designs*
- Variable names
  - Nature of design (within or between subjects)

- Suggested Syntax**
- $F(df_B, df_W) = \text{observed } F \text{ value}$ , significance level, ES

**Suggested Format (example is of a factorial design)**

- The means for each of the groups (shown in Table 1) indicated that influence of variable X on variable Y was in the direction predicted. The 2 (message directness)  $\times$  3 (message length) between groups ANOVA revealed that the hypothesis was supported,  $F(2, 63) = 12.47, p < .01$ ,  $\eta^2 = .17$ .

# Oneway Repeated Measures ANOVA

Overview    Focus    Steps    Example    Write Up

1 IV, within-subjects, 2 or more levels

Normally distributed DV

EX: Does confidence in coping change over time?

<pre> GLM confid1 confid2 confid3 /WSFACTOR = time 3 Polynomial /METHOD = SSTYPE(3) /EMMEANS = TABLES(time) COMPARE ADJ(BONFERRONI) /PRINT = DESCRIPTIVE ETASQ /CRITERIA = ALPHA(.05) /WSDDESIGN = time . </pre>	<p style="text-align: center;"><b>Multivariate Tests<sup>b</sup></b></p> <table border="1"> <thead> <tr> <th>Effect</th> <th>Value</th> <th>F</th> <th>Hypothesis df</th> <th>Error df</th> <th>Sig.</th> <th>Partial Eta Squared</th> </tr> </thead> <tbody> <tr> <td>time</td> <td>.749</td> <td>41.711<sup>a</sup></td> <td>2,000</td> <td>28,000</td> <td>.000</td> <td>.749</td> </tr> <tr> <td>Pillai's Trace</td> <td>.251</td> <td>41.711<sup>a</sup></td> <td>2,000</td> <td>28,000</td> <td>.000</td> <td>.749</td> </tr> <tr> <td>Wilks' Lambda</td> <td>.251</td> <td>41.711<sup>a</sup></td> <td>2,000</td> <td>28,000</td> <td>.000</td> <td>.749</td> </tr> <tr> <td>Hotelling's Trace</td> <td>2.979</td> <td>41.711<sup>a</sup></td> <td>2,000</td> <td>28,000</td> <td>.000</td> <td>.749</td> </tr> <tr> <td>Roy's Largest Root</td> <td>2.979</td> <td>41.711<sup>a</sup></td> <td>2,000</td> <td>28,000</td> <td>.000</td> <td>.749</td> </tr> </tbody> </table> <p>a. Exact statistic  b. Design: Intercept Within Subjects Design: time</p> <p style="text-align: center;"><b>Descriptive Statistics</b></p> <table border="1"> <thead> <tr> <th></th> <th>Mean</th> <th>Std. Deviation</th> <th>N</th> </tr> </thead> <tbody> <tr> <td>confidence time1</td> <td>19.00</td> <td>5.37</td> <td>30</td> </tr> <tr> <td>confidence time2</td> <td>21.87</td> <td>5.59</td> <td>30</td> </tr> <tr> <td>confidence time3</td> <td>25.03</td> <td>5.20</td> <td>30</td> </tr> </tbody> </table>	Effect	Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	time	.749	41.711 <sup>a</sup>	2,000	28,000	.000	.749	Pillai's Trace	.251	41.711 <sup>a</sup>	2,000	28,000	.000	.749	Wilks' Lambda	.251	41.711 <sup>a</sup>	2,000	28,000	.000	.749	Hotelling's Trace	2.979	41.711 <sup>a</sup>	2,000	28,000	.000	.749	Roy's Largest Root	2.979	41.711 <sup>a</sup>	2,000	28,000	.000	.749		Mean	Std. Deviation	N	confidence time1	19.00	5.37	30	confidence time2	21.87	5.59	30	confidence time3	25.03	5.20	30
Effect	Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared																																																					
time	.749	41.711 <sup>a</sup>	2,000	28,000	.000	.749																																																					
Pillai's Trace	.251	41.711 <sup>a</sup>	2,000	28,000	.000	.749																																																					
Wilks' Lambda	.251	41.711 <sup>a</sup>	2,000	28,000	.000	.749																																																					
Hotelling's Trace	2.979	41.711 <sup>a</sup>	2,000	28,000	.000	.749																																																					
Roy's Largest Root	2.979	41.711 <sup>a</sup>	2,000	28,000	.000	.749																																																					
	Mean	Std. Deviation	N																																																								
confidence time1	19.00	5.37	30																																																								
confidence time2	21.87	5.59	30																																																								
confidence time3	25.03	5.20	30																																																								

<p style="text-align: center;"><b>Pairwise Comparisons</b></p> <p>Measure: MEASURE_1</p> <table border="1"> <thead> <tr> <th>(I) time</th> <th>(J) time</th> <th>Mean Difference (I-J)</th> <th>Std. Error</th> <th>Sig.<sup>a</sup></th> <th colspan="2">95% Confidence Interval for Difference<sup>a</sup></th> </tr> <tr> <th></th> <th></th> <th></th> <th></th> <th></th> <th>Lower Bound</th> <th>Upper Bound</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>2</td> <td>-2.867*</td> <td>.868</td> <td>.008</td> <td>-5.072</td> <td>-.661</td> </tr> <tr> <td>1</td> <td>3</td> <td>-6.033*</td> <td>.833</td> <td>.000</td> <td>-8.149</td> <td>-3.918</td> </tr> <tr> <td>2</td> <td>1</td> <td>2.867*</td> <td>.868</td> <td>.008</td> <td>.661</td> <td>5.072</td> </tr> <tr> <td>2</td> <td>3</td> <td>3.167*</td> <td>.447</td> <td>.000</td> <td>-4.304</td> <td>-2.030</td> </tr> <tr> <td>3</td> <td>1</td> <td>6.033*</td> <td>.833</td> <td>.000</td> <td>3.918</td> <td>8.149</td> </tr> <tr> <td>3</td> <td>2</td> <td>3.167*</td> <td>.447</td> <td>.000</td> <td>2.030</td> <td>4.304</td> </tr> </tbody> </table> <p>Based on estimated marginal means</p> <p>* The mean difference is significant at the .05 level.</p> <p>a. Adjustment for multiple comparisons: Bonferroni.</p>	(I) time	(J) time	Mean Difference (I-J)	Std. Error	Sig. <sup>a</sup>	95% Confidence Interval for Difference <sup>a</sup>							Lower Bound	Upper Bound	1	2	-2.867*	.868	.008	-5.072	-.661	1	3	-6.033*	.833	.000	-8.149	-3.918	2	1	2.867*	.868	.008	.661	5.072	2	3	3.167*	.447	.000	-4.304	-2.030	3	1	6.033*	.833	.000	3.918	8.149	3	2	3.167*	.447	.000	2.030	4.304	<p style="text-align: center;"><b>Mauchly's Test of Sphericity<sup>b</sup></b></p> <p>Measure:MEASURE_1</p> <table border="1"> <thead> <tr> <th rowspan="2">Within Subjects Effect</th> <th rowspan="2">Mauchly's W</th> <th rowspan="2">Approx. Chi-Square</th> <th rowspan="2">df</th> <th rowspan="2">Sig.</th> <th colspan="3">Epsilon<sup>a</sup></th> </tr> <tr> <th>Greenhouse-Geisser</th> <th>Huynh-Feldt</th> <th>Lower-bound</th> </tr> </thead> <tbody> <tr> <td>time</td> <td>.592</td> <td>14.680</td> <td>2</td> <td>.001</td> <td>.710</td> <td>.737</td> <td>.500</td> </tr> </tbody> </table> <p>Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.</p> <p>a. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.</p> <p>b. Design: Intercept Within Subjects Design: time</p> <p style="text-align: center;"><b>Presenting the results from one-way repeated measures ANOVA</b></p> <p>The results of one-way repeated measures ANOVA could be presented as follows:</p> <p style="text-align: center;">A one-way repeated measures ANOVA was conducted to compare scores on the Confidence in Coping with Statistics Test at Time 1 (prior to the intervention), Time 2 (following the intervention) and Time 3 (three-month follow-up). The means and standard deviations are presented in Table 1. There was a significant effect for time, Wilks' Lambda = .25, <math>F(2, 28) = 41.17</math>, <math>p &lt; .001</math>, multivariate partial eta squared = .75.</p> <p><b>Table 1</b>  <i>Descriptive Statistics for Confidence in Coping with Statistics Test Scores for Time 1, Time 2 and Time 3</i></p> <table border="1"> <thead> <tr> <th>Time period</th> <th>N</th> <th>Mean</th> <th>Standard deviation</th> </tr> </thead> <tbody> <tr> <td>Time 1 (Pre-intervention)</td> <td>30</td> <td>19.00</td> <td>5.37</td> </tr> <tr> <td>Time 2 (Post-intervention)</td> <td>30</td> <td>21.87</td> <td>5.59</td> </tr> <tr> <td>Time 3 (3-month follow-up)</td> <td>30</td> <td>25.03</td> <td>5.20</td> </tr> </tbody> </table>	Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon <sup>a</sup>			Greenhouse-Geisser	Huynh-Feldt	Lower-bound	time	.592	14.680	2	.001	.710	.737	.500	Time period	N	Mean	Standard deviation	Time 1 (Pre-intervention)	30	19.00	5.37	Time 2 (Post-intervention)	30	21.87	5.59	Time 3 (3-month follow-up)	30	25.03	5.20
(I) time	(J) time	Mean Difference (I-J)	Std. Error	Sig. <sup>a</sup>	95% Confidence Interval for Difference <sup>a</sup>																																																																																							
					Lower Bound	Upper Bound																																																																																						
1	2	-2.867*	.868	.008	-5.072	-.661																																																																																						
1	3	-6.033*	.833	.000	-8.149	-3.918																																																																																						
2	1	2.867*	.868	.008	.661	5.072																																																																																						
2	3	3.167*	.447	.000	-4.304	-2.030																																																																																						
3	1	6.033*	.833	.000	3.918	8.149																																																																																						
3	2	3.167*	.447	.000	2.030	4.304																																																																																						
Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon <sup>a</sup>																																																																																							
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound																																																																																					
time	.592	14.680	2	.001	.710	.737	.500																																																																																					
Time period	N	Mean	Standard deviation																																																																																									
Time 1 (Pre-intervention)	30	19.00	5.37																																																																																									
Time 2 (Post-intervention)	30	21.87	5.59																																																																																									
Time 3 (3-month follow-up)	30	25.03	5.20																																																																																									

**Analysis of Variance (ANOVA)**

- What to Report**
- Degrees of freedom (between and within) ( $df_B/df_W$ )
  - Observed  $F$  value ( $F$ )
  - Significance level ( $p$ )
  - Effect size (ES, varies with test)

*Supporting information  
(reported in text or table)*

- Number of observations ( $N$ )
- Number of observations per cell ( $n$ )
- Descriptive statistics ( $M$  and  $SD$ )

*Additional information for factorial designs*

- Variable names
- Nature of design (within or between subjects)

- Suggested Syntax**
- $F(df_B, df_W) = \text{observed } F \text{ value, } p \text{ significance level, ES}$

**Suggested Format (example is of a factorial design)**

- The means for each of the groups (shown in Table 1) indicated that influence of variable  $X$  on variable  $Y$  was in the direction predicted. The 2 (message directness)  $\times$  3 (message length) between groups ANOVA revealed that the hypothesis was supported,  $F(2, 63) = 12.47, p < .01, \eta^2 = .17$ .

## Complex Questions: Incorporating More than One IV

2x3 Design		IV 1		2x3 Design		Time of Day	
IV 2	IV2: Level 1	IV1: Level 1	IV1: Level 2	Caffeine	1 coffee	dv	dv
		dv	dv		2 coffees	dv	dv
		dv	dv		3 coffees	dv	dv
	IV2: Level 3	dv	dv				

## Factorial ANOVA

Overview    Focus    Steps    Example    Write Up

Performed in designs with more than one independent variable, two or more levels, and between groups design.

Normally distributed DV

EX: Do gender and age interact to influence reports of optimism?

Tests of Between-Subjects Effects

Dependent Variable: total optimism

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	238.647 <sup>a</sup>	5	47.729	2.475	.032	.028
Intercept	206790.069	1	206790.069	10721.408	.000	.962
agegp3	150.863	2	75.431	3.911	.021	.018
sex	5.717	1	5.717	.296	.586	.001
agegp3 * sex	55.709	2	27.855	1.444	.237	.007
Error	8274.374	429	19.288			
Total	221303.000	435				
Corrected Total	8513.021	434				

a. R Squared = .028 (Adjusted R Squared = .017)

Multiple Comparisons

Dependent Variable: toptim

Tukey HSD

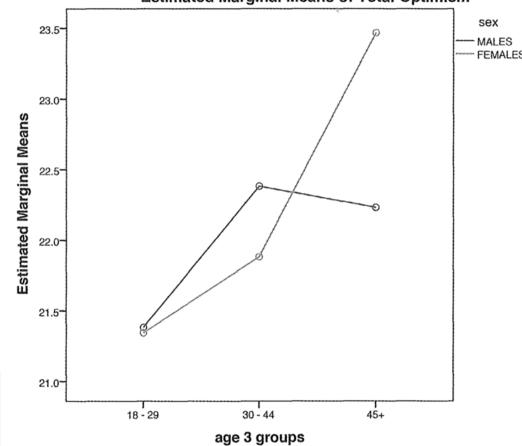
(I) age 3 groups	(J) age 3 groups	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1 18 - 29	2 30 - 44	-.74	.507	.308	-1.94	.45
	3 45+	-.160*	.524	.007	-2.83	-.36
2 30 - 44	1 18 - 29	.74	.507	.308	-.45	1.94
	3 45+	-.85	.519	.230	-2.07	.37
3 45+	1 18 - 29	1.60*	.524	.007	.36	2.83
	2 30 - 44	.85	.519	.230	-.37	2.07

Based on observed means.

The error term is Mean Square(Error) = 19.288.

\* The mean difference is significant at the .05 level.

Estimated Marginal Means of Total Optimism



## PRESENTING THE RESULTS FROM TWO-WAY ANOVA

The results of the analysis conducted above could be presented as follows:

A two-way between-groups analysis of variance was conducted to explore the impact of sex and age on levels of optimism, as measured by the Life Orientation Test (LOT). Participants were divided into three groups according to their age (Group 1: 18–29 years; Group 2: 30–44 years; Group 3: 45 years and above). The interaction effect between sex and age group was not statistically significant,  $F(2, 429) = 1.44, p = .24$ . There was a statistically significant main effect for age,  $F(2, 429) = 3.91, p = .02$ ; however, the effect size was small (partial eta squared = .02). Post-hoc comparisons using the Tukey HSD test indicated that the mean score for the 18–29 years age group ( $M = 21.36, SD = 4.55$ ) was significantly different from the 45+ group ( $M = 22.96, SD = 4.49$ ). The 30–44 years age group ( $M = 22.10, SD = 4.15$ ) did not differ significantly from either of the other groups. The main effect for sex,  $F(1, 429) = .30, p = .59$ , did not reach statistical significance.

## Mixed between-within subjects ANOVA

[Overview](#)
[Focus](#)
[Step 1](#)
[Step 2](#)
[Example](#)
[Write Up](#)

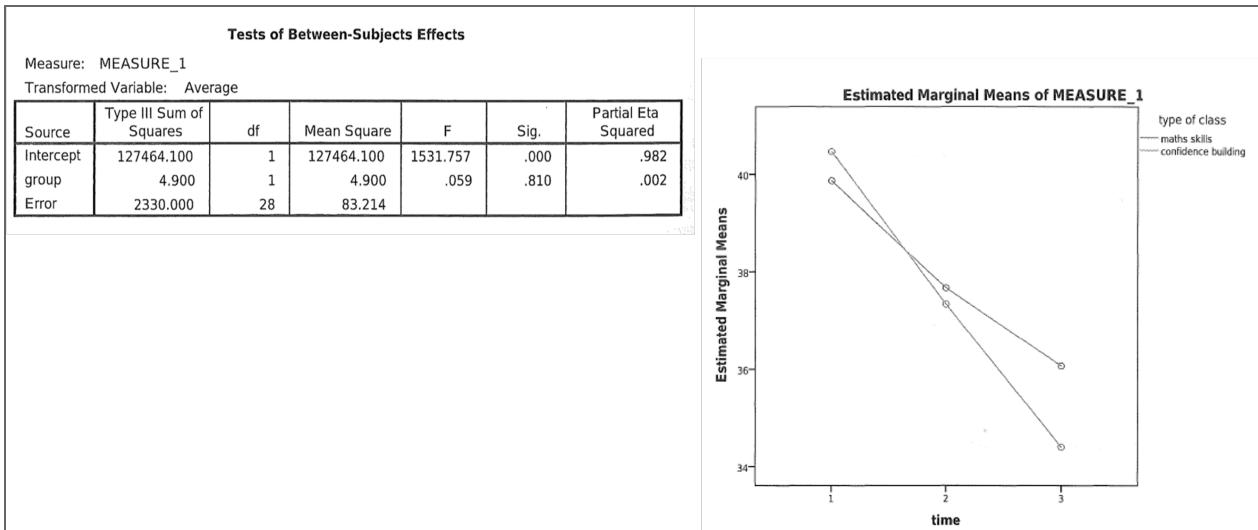
Features one between-groups IV and one within-subjects IV

Normally distributed DV.

EX: Is there an interaction between *condition* and *time* on perceptions of fear of statistics?

<pre> GLM fost1 fost2 fost3 BY group /WSFACTOR = time 3 Polynomial /METHOD = SSTYPE(3) /PLOT = PROFILE(time*group) /PRINT = DESCRIPTIVE ETASQ HOMOGENEITY /CRITERIA = ALPHA(.05) /WSDESIGN = time /DESIGN = group . </pre>	<b>Box's Test of Equality of Covariance Matrices<sup>a</sup></b> <table border="1" style="margin-bottom: 5px;"> <tr> <td>Box's M</td><td>1.520</td></tr> <tr> <td>F</td><td>.224</td></tr> <tr> <td>df1</td><td>6</td></tr> <tr> <td>df2</td><td>5680.302</td></tr> <tr> <td>Sig.</td><td>.969</td></tr> </table> <small>Tests the null hypothesis that the observed covariance matrices of the dependent variables are equal across groups.</small> <small>a. Design: Intercept + group Within Subjects Design: time</small>	Box's M	1.520	F	.224	df1	6	df2	5680.302	Sig.	.969																																								
Box's M	1.520																																																		
F	.224																																																		
df1	6																																																		
df2	5680.302																																																		
Sig.	.969																																																		
Selected output generated from this procedure is shown below:																																																			
<b>Descriptive Statistics</b> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th></th> <th>type of class</th> <th>Mean</th> <th>Std. Deviation</th> <th>N</th> </tr> </thead> <tbody> <tr> <td>fear of stats time1</td> <td>1 maths skills</td> <td>39.87</td> <td>4.596</td> <td>15</td> </tr> <tr> <td></td> <td>2 confidence building</td> <td>40.47</td> <td>5.817</td> <td>15</td> </tr> <tr> <td></td> <td>Total</td> <td>40.17</td> <td>5.160</td> <td>30</td> </tr> <tr> <td>fear of stats time2</td> <td>1 maths skills</td> <td>37.67</td> <td>4.515</td> <td>15</td> </tr> <tr> <td></td> <td>2 confidence building</td> <td>37.33</td> <td>5.876</td> <td>15</td> </tr> <tr> <td></td> <td>Total</td> <td>37.50</td> <td>5.151</td> <td>30</td> </tr> <tr> <td>fear of stats time3</td> <td>1 maths skills</td> <td>36.07</td> <td>5.431</td> <td>15</td> </tr> <tr> <td></td> <td>2 confidence building</td> <td>34.40</td> <td>6.631</td> <td>15</td> </tr> <tr> <td></td> <td>Total</td> <td>35.23</td> <td>6.015</td> <td>30</td> </tr> </tbody> </table>			type of class	Mean	Std. Deviation	N	fear of stats time1	1 maths skills	39.87	4.596	15		2 confidence building	40.47	5.817	15		Total	40.17	5.160	30	fear of stats time2	1 maths skills	37.67	4.515	15		2 confidence building	37.33	5.876	15		Total	37.50	5.151	30	fear of stats time3	1 maths skills	36.07	5.431	15		2 confidence building	34.40	6.631	15		Total	35.23	6.015	30
	type of class	Mean	Std. Deviation	N																																															
fear of stats time1	1 maths skills	39.87	4.596	15																																															
	2 confidence building	40.47	5.817	15																																															
	Total	40.17	5.160	30																																															
fear of stats time2	1 maths skills	37.67	4.515	15																																															
	2 confidence building	37.33	5.876	15																																															
	Total	37.50	5.151	30																																															
fear of stats time3	1 maths skills	36.07	5.431	15																																															
	2 confidence building	34.40	6.631	15																																															
	Total	35.23	6.015	30																																															

<b>Mauchly's Test of Sphericity<sup>a</sup></b> <small>Measure: MEASURE_1</small> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th rowspan="2">Within Subjects Effect</th> <th rowspan="2">Mauchly's W</th> <th rowspan="2">Approx. Chi-Square</th> <th rowspan="2">df</th> <th rowspan="2">Sig.</th> <th colspan="3">Epsilon<sup>b</sup></th> </tr> <tr> <th>Greenhouse-Geisser</th> <th>Huynh-Feldt</th> <th>Lower-bound</th> </tr> </thead> <tbody> <tr> <td>time</td> <td>.348</td> <td>28.517</td> <td>2</td> <td>.000</td> <td>.605</td> <td>.640</td> <td>.500</td> </tr> </tbody> </table> <small>Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.</small> <small>a. Design: Intercept + group Within Subjects Design: time</small> <small>b. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.</small>	Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon <sup>b</sup>			Greenhouse-Geisser	Huynh-Feldt	Lower-bound	time	.348	28.517	2	.000	.605	.640	.500	<b>Levene's Test of Equality of Error Variances<sup>a</sup></b> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th></th> <th>F</th> <th>df1</th> <th>df2</th> <th>Sig.</th> </tr> </thead> <tbody> <tr> <td>fear of stats time1</td> <td>.893</td> <td>1</td> <td>28</td> <td>.353</td> </tr> <tr> <td>fear of stats time2</td> <td>.767</td> <td>1</td> <td>28</td> <td>.389</td> </tr> <tr> <td>fear of stats time3</td> <td>.770</td> <td>1</td> <td>28</td> <td>.388</td> </tr> </tbody> </table> <small>Tests the null hypothesis that the error variance of the dependent variable is equal across groups.</small> <small>a. Design: Intercept + group Within Subjects Design: time</small>		F	df1	df2	Sig.	fear of stats time1	.893	1	28	.353	fear of stats time2	.767	1	28	.389	fear of stats time3	.770	1	28	.388	<b>Tests of Between-Subjects Effects</b> <small>Measure: MEASURE_1 Transformed Variable: Average</small> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>Source</th> <th>Type III Sum of Squares</th> <th>df</th> <th>Mean Square</th> <th>F</th> <th>Sig.</th> <th>Partial Eta Squared</th> </tr> </thead> <tbody> <tr> <td>Intercept</td> <td>127464.100</td> <td>1</td> <td>127464.100</td> <td>1531.757</td> <td>.000</td> <td>.982</td> </tr> <tr> <td>group</td> <td>4.900</td> <td>1</td> <td>4.900</td> <td>.059</td> <td>.810</td> <td>.002</td> </tr> <tr> <td>Error</td> <td>2330.000</td> <td>28</td> <td>83.214</td> <td></td> <td></td> <td></td> </tr> </tbody> </table>	Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Intercept	127464.100	1	127464.100	1531.757	.000	.982	group	4.900	1	4.900	.059	.810	.002	Error	2330.000	28	83.214			
Within Subjects Effect						Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon <sup>b</sup>																																																											
	Greenhouse-Geisser	Huynh-Feldt	Lower-bound																																																																		
time	.348	28.517	2	.000	.605	.640	.500																																																														
	F	df1	df2	Sig.																																																																	
fear of stats time1	.893	1	28	.353																																																																	
fear of stats time2	.767	1	28	.389																																																																	
fear of stats time3	.770	1	28	.388																																																																	
Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared																																																															
Intercept	127464.100	1	127464.100	1531.757	.000	.982																																																															
group	4.900	1	4.900	.059	.810	.002																																																															
Error	2330.000	28	83.214																																																																		



A mixed between-within subjects analysis of variance was conducted to assess the impact of two different interventions (Maths Skills, Confidence Building) on participants' scores on the Fear of Statistics Test, across three time periods (pre-intervention, post-intervention and three-month follow-up). There was no significant interaction between program type and time, Wilks' Lambda = .87,  $F(2, 27) = 2.03, p = .15$ , partial eta squared = .13. There was a substantial main effect for time, Wilks' Lambda = .34,  $F(2, 27) = 26.59, p < .001$ , partial eta squared = .66, with both groups showing a reduction in Fear of Statistics Test scores across the three time periods (see Table 1). The main effect comparing the two types of intervention was not significant,  $F(1, 28) = .059, p = .81$ , partial eta squared = .002, suggesting no difference in the effectiveness of the two teaching approaches.

**Table 1**

*Fear of Statistics Test Scores for the Maths Skills and Confidence Building Programs Across Three Time Periods*

	Maths Skills			Confidence Building		
	n	M	SD	n	M	SD
Pre-intervention	15	39.87	4.60	15	40.47	5.82
Post-intervention	15	37.67	4.51	15	37.33	5.88
3-mth follow-up	15	36.07	5.43	15	34.40	6.63

## Multivariate ANOVA (MANOVA)

[Overview](#)
[Focus](#)
[Example](#)
[Write Up](#)
[Other Example](#)

One between-groups IV with 2 or more levels Two or more normally-distributed DVs

\*Becomes repeated-measures MANOVA if you have within-subjects IV

EX: Do men and women differ in their reports of psychological wellbeing (3 variables: positive affect, negative affect, perceived stress)?

### PRESENTING THE RESULTS FROM MANOVA

The results of this multivariate analysis of variance could be presented as follows:

A one-way between-groups multivariate analysis of variance was performed to investigate sex differences in psychological wellbeing. Three dependent variables were used: positive affect, negative affect and perceived stress. The independent variable was gender. Preliminary assumption testing was conducted to check for normality, linearity, univariate and multivariate outliers, homogeneity of variance-covariance matrices, and multicollinearity, with no serious violations noted. There was a statistically significant difference between males and females on the combined dependent variables,  $F(3, 428) = 3.57, p = .014$ ; Wilks' Lambda = .98; partial eta squared = .02. When the results for the dependent variables were considered separately, the only difference to reach statistical significance, using a Bonferroni adjusted alpha level of .017, was perceived stress,  $F(1, 430) = 8.34, p = .004$ , partial eta squared = .02. An inspection of the mean scores indicated that females reported slightly higher levels of perceived stress ( $M = 27.42, SD = 6.08$ ) than males ( $M = 25.79, SD = 5.41$ ).

### Multivariate Analysis of Variance (MANOVA)

- What to Report**
- Multivariate statistic (Wilks lambda, Pillai's (varies))
  - Degrees of freedom (between and within) ( $df_B, df_W$ )
  - Observed  $F$  value ( $F$ )
  - Significance level ( $p$ )
  - Effect size (ES, varies with test)
  - Univariate effects (see ANOVA)

- Supporting information  
(reported in text or table)*
- Number of observations ( $N$ )
  - Number of observations per cell ( $n$ )
  - Descriptive statistics ( $M$  and  $SD$ )

- Suggested Syntax**
- Wilks lambda = observed  $\lambda$  value,  $F(df_B, df_W)$  = observed  $F$  value, significance level, ES

- Suggested Format**
- The  $2 \times 3$  between groups MANOVA revealed that the multivariate main effect was supported, Wilks lambda = .24,  $F(2, 63) = 12.47, p < .01, \eta^2 = .17$ .

#### Example from the Literature

This example is well organized and includes the important elements necessary for a complete results section. Of the multivariate statistics used to test main effects and interaction in MANOVA, Wilks lambda is the most commonly reported. Although effect size was reported for the multivariate main effect, it was not reported for either of the univariate effect tests. It is rec-

Your reading by Schrodt and Witt for today!

# One-Way Analysis of Covariance

Overview    Focus    Example    Write Up

One between-groups IV with 2 or more levels

One normally-distributed DVs

One or more theoretically-justified control variables

\*Becomes repeated-measures ANCOVA if you have within-subjects IV

EX: Are there differences on fear of statistics for those who receive a math skills course versus those who receive a confidence building course while controlling for initial fear of statistics?

The effect of the intervention at T2 while controlling for values of the DV at T1.

## Presenting the results from one-way ANCOVA

The results of this one-way analysis of covariance could be presented as follows:

A one-way between-groups analysis of covariance was conducted to compare the effectiveness of two different interventions designed to reduce participants' fear of statistics. The independent variable was the type of intervention (maths skills, confidence building), and the dependent variable consisted of scores on the Fear of Statistics Test administered after the intervention was completed. Participants' scores on the pre-intervention administration of the Fear of Statistics Test were used as the covariate in this analysis.

Preliminary checks were conducted to ensure that there was no violation of the assumptions of normality, linearity, homogeneity of variances, homogeneity of regression slopes, and reliable measurement of the covariate. After adjusting for pre-intervention scores, there was no significant difference between the two intervention groups on post-intervention scores on the Fear of Statistics Test,  $F(1, 27) = .76, p = .39$ , partial eta squared = .03. There was a strong relationship between the pre-intervention and post-intervention scores on the Fear of Statistics Test, as indicated by a partial eta squared value of .75.

## ANCOVA

### What to Report (in addition to ANOVA)

- Adjusted means (adj M)
- Covariate information (regression format)

### Suggested Syntax

- For main effects and interactions:  
 $F(df_{\text{fp}}, df_{\text{w}})$  = observed F value,  
 significance level, ES (accompanied by  
 a table of adjusted group and cell means  
 and standard deviations).
- For covariate analyses:  $F(df_{\text{v}}, df_{\text{w}})$  =  
 observed F value, significance level,  
 partial ES (after covariates are removed).  
 Accompanying table of pooled within-  
 cell intercorrelations among the  
 covariates and the dependent variable.

Note: Analysis of covariance (ANOVA) allows for the comparison of group means on a dependent variable after the group means have been adjusted on a relevant covariate variable. Reporting ANCOVA results is similar to reporting ANOVA results, except that adjusted means and F ratio derived from those means need to be stated. Notice that as analyses or the research designs increase in complexity, reports of analyses necessarily increase dramatically in length.



## Two-Way Analysis of Covariance

Overview    Focus    Example    Write Up

Two or more between-groups IV with 2 or more levels

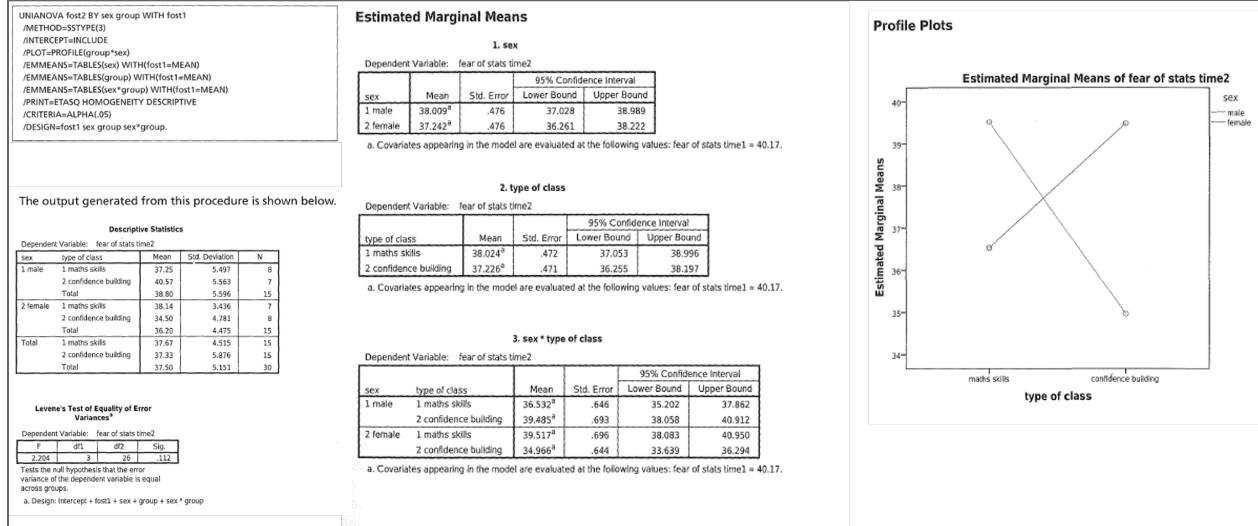
One normally-distributed DVs

One or more theoretically-justified control variables

\*Becomes 2 factor repeated-measures ANCOVA if you have within-subjects IV (or mixed design!)

EX: Do sex and type of course interact to influence fear of statistics while controlling for initial fear of statistics?

The effect of the intervention at T2 while controlling for values of the DV at T1.



## Presenting the results from two-way ANCOVA

The results of this analysis could be presented as follows:

A 2 by 2 between-groups analysis of covariance was conducted to assess the effectiveness of two programs in reducing fear of statistics for male and female participants. The independent variables were the type of program (maths skills, confidence building) and gender. The dependent variable was scores on the Fear of Statistics Test (FOST), administered following completion of the intervention programs (Time 2). Scores on the FOST administered prior to the commencement of the programs (Time 1) were used as a covariate to control for individual differences.

Preliminary checks were conducted to ensure that there was no violation of the assumptions of normality, linearity, homogeneity of variances, homogeneity of regression slopes, and reliable measurement of the covariate. After adjusting for FOST scores at Time 1, there was a significant interaction effect,  $F(1, 25) = 31.7, p < .001$ , with a large effect size (partial eta squared = .56). Neither of the main effects were statistically significant, program:  $F(1, 25) = 1.43, p = .24$ ; gender:  $F(1, 25) = 1.27, p = .27$ . These results suggest that males and females respond differently to the two types of interventions. Males showed a more substantial decrease in fear of statistics after participation in the maths skills program. Females, on the other hand, appeared to benefit more from the confidence-building program.

## Help Links For Each Test: Basic Difference

- Chi Square Goodness of Fit test in [SPSS / R](#)
- Chi Square Test of Independence in [SPSS / R](#)
- Independent samples *t* test in [SPSS / R](#)
- Paired samples *t* test in [SPSS / R](#)
- Oneway ANOVA in [SPSS / R](#)

## Help Links For Each Test: Complex Difference

- Factorial ANOVA in [SPSS / R](#)
- Mixed between-within subjects ANOVA in [SPSS / R](#)
- MANOVA in [SPSS / R](#)
- Factorial MANOVA in [SPSS / R](#)

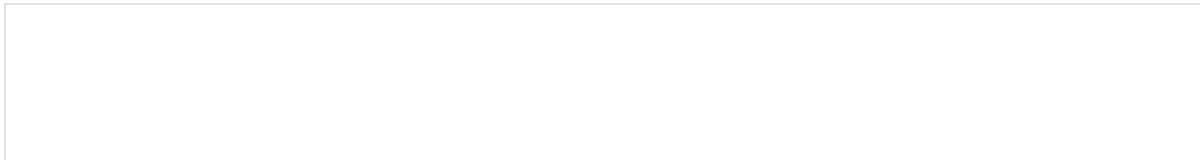
## Summary and Review

MOST designs are going to be simple; Complex does not mean BETTER

You now have a Fisher-Price stats toolkit :)

Rely on mentors, examples to guide you through more complex stats

## Error



versity of  
ntucky