Research Series

COMPLETE SERIES

By: Curtis Lauterbach
3-7-13
Revised 10-21-18

INTRODUCTION

Also the author of this website has made an effort to ensure the accuracy and completeness of the presented information, the author assumes no responsibility for any errors, missing information, inaccuracies, or inconsistencies.

VALIDITY & RELIABILITY

It is important to think about validity and reliability when designing your study. This will impact how they will be able to be generalized to the population.

Validity

Validity-How likely the measure you are using actually measures your dependent variable.

There are different types of validity:

1. **Internal**-How well the observed relationship accurately reflects the relationship between the intended variables (Heiman, 2001).
2. **External**-Are the results able to be generalized to population beyond the participants studied.
3. **Construct**-How well a test reflects the construct it is meant to test.
4. **Content**-How well a test measures ONLY the construct in question (Heiman, 2001).
5. **Criterion**-How well a test is able to tell participants apart on abilities. An example could be the Mini Mental States exam. It has been designed to distinguish between normal memory and the beginnings of dementia.
6. **Concurrent**-How well a test measures a participant's current abilities.
7. **Predictive**-How well a test predicts a participant's future abilities.
8. **Convergent**-How well a test correlates to another well accepted test (Heiman, 2001).
9. **Discriminant**-How well a test DOES NOT correlate to another accepted test.
10. **Face**-How well a test "appears" to measure the construct in question ("face value").
11. **Ecological**-Generalizations of what participants can do in a study to what they can do in real life.
12. **Temporal**-Generalizations to other time frames (i.e. 5 and 30 minute word recalls for memory).

Reliability

Reliability-How likely the measure you are using measures the dependent variable consistently (does it measure it each time) (Heiman, 2001).

There are different types of validity:

1. **Inter-rater**-How consistent results are between raters. An example could be the results of three different judges on a snowboarder's performance.
2. **Test-retest**-How consistent the results from a specific test are over time (Heiman, 2001). An example could be Grooved Pegboard scores for hand-eye coordination that are measured twice at six month intervals.
3. **Split-half**-There should be strong correlations between participant's scores on one half of a test compared to the other (Heiman, 2001). An example could be hand-eye coordination scores, where the participant's scores are divided in half (odd vs even trials). The summary scores of the odd numbered trials should have strong correlations to the even numbered trials.

DESIGNS

The type of research project you are doing will depict the design you will use:

1. **Within-Subjects**-Used when the researcher is interested in determining whether or not one condition is better than another, and is dependent upon having the same participants in all conditions. An example could be determining whether website A is more efficient than website B.
2. **Between-Subjects**-Used when the researcher is interested in determining whether or not one condition is better than another, but participants are placed in only one condition. An example could be determining whether using PowerPoint presentations helps students get better grades compared to overhead projects (Heiman, 2001).

Counterbalancing-The practice of randomizing the order of condition for participants so balance the number of participants that receive condition A and condition B (see Tables 1 and 2). This allows for an equal number of participants in each condition (Heiman, 2001).

Table 1. Counterbalancing for within-subjects design

Participant #	**First Condition**	**Second Condition**
1	A	B
2	B	A
3	A	B
4	B	A

Table 2. Counterbalancing for between-subjects design

Participant #	**Condition**
1	A
2	B
3	A
4	B

REFERENCE

Heiman, G. (2001). Research Methods in Psychology (3rd ed.). Cengage Learning, pgs. 71-73, 76-77, 114-115, 165, 173, 282-283.

INTRODUCTION

Also the author of this website has made an effort to ensure the accuracy and completeness of the presented information, the author assumes no responsibility for any errors, missing information, inaccuracies, or inconsistencies.

STATISTICS

When performing research it is essential that you are able to make sense of you data. This allows you to inform other researchers in your field and others what you have found. It also can be used to help build evidence for a theory. Therefore and understanding of what test to use and when is necessary. There are some good practices to do if you are not familiar with performing statistical analysis. The first is to determine your variables.

Overview of basic statistics

1. **Dependent variable (DV)**-This is the one that you are measuring. From the example of the student's score.
2. **Independent variable (IV)**-This is the one that you are manipulating with the different conditions. From the example either PowerPoint or overhead projector presentation (Aron, Aron, & Coups, 2005).

Levels of Measurement

Once this has been done you will want to determine the level of measurement that both variables are. This will help you later determine what test to use in your analysis. There are four levels of measurement (Aron, Aron, & Coups, 2005):

1. **Interval**-There is no true zero, there are equal intervals throughout the scale, and negative scores are possible (i.e. 80.5 degrees Fahrenheit).
2. **Ratio**-There are equal intervals throughout the scale, and has a true zero so negative scores are not possible (i.e. 1 marble).
3. **Nominal**-Categories (i.e. Male/Female).
4. **Ordinal**-Ranking (i.e. 1st, 2nd, 3rd, and 4th).

With this information you can refer to the following chart to determine what statistical test to use (see Figure 1).

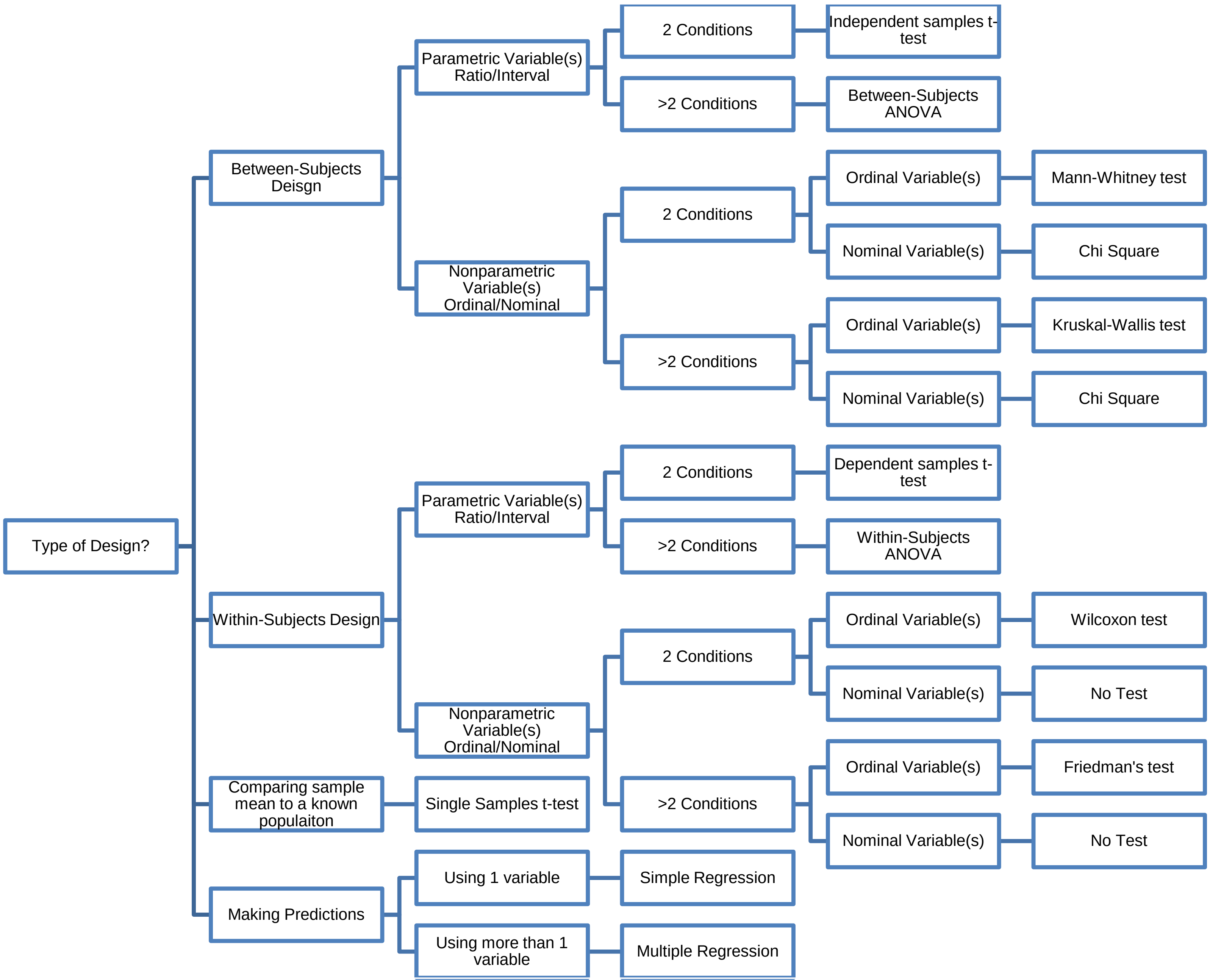

Figure 1. Flow chart for basic statistics tests.

Statistical tests can be broken into two groups, parametric and nonparametric and are determined by the level of measurement, number of dependent and independent variables, covariates, comparing to an already known population, and type of design. Parametric tests are used to analyze interval and ratio data, and nonparametric tests analyze ordinal and nominal data. There are different tests to use in each group. We will start with the parametric tests first.

Parametric Tests: (Interval/Ratio data)

These tests assume that the data is normally distributed (bell curve) and are very strong when compared to nonparametric tests (Heiman, 2001).

1. **Within-Subjects Design**

A. **Paired-Samples t-Test (AKA: Dependent samples t-test)**: Compares means of two groups who received all conditions (Aron, Aron, & Coup, 2005). An example is having participants run a mile one day with no sleep and then with a full night's sleep, and seeing which group was the fastest.

B. **Within-subjects ANOVA (AKA: Repeated-Measures ANOVA)**: Compares means of more than two groups who received all conditions, and decreases the rate of Type I errors (described later in this section) (Aron, Aron, & Coup, 2005 & Cronk, 2011). An example is having participants run a mile one day with no sleep, a full night's sleep, and with caffeine prior to run.

2. **Between-Subjects Design**

A. **Independent samples t-test**: Compares means of two groups and participants only receive one of the conditions (Aron, Aron, & Coup, 2005). An example is determining if a PowerPoint presentation impacted grades compared to a control group that received a general lecture.

B. **Between-subjects ANOVA (AKA: One-Way ANOVA)**: Compares means of more than two groups, and participants only receive one of the conditions (Cronk, 2011). An example is determining if PowerPoint or an overhead projector impacted grades and if so which one.

C. **Pearson Correlation**: Determines if there is a relationship between two variables. It also determines the strength of the relationship, if one exists (Aron, Aron, & Coup, 2005).

The correlation coefficient (r) ranges from -1.0 to 1.0. The closer to -1.0 or 1.0, the stronger the relationship. When r is negative that means there is a negative relationship (one variable goes up or down the other does the opposite). When r is positive that means there is a positive relationship (one variable goes up or down so does the other). See the example below examining the correlation between texts sent per week and grades. Another way to determine the relationship and strength is to look at the graph. The closer the data points are to the line, the stronger the relationship (see Figure 2) (Heiman, 2001).

Caution: A strong correlation does not mean that a condition caused the outcome. It only means that the two variables are related in some way.

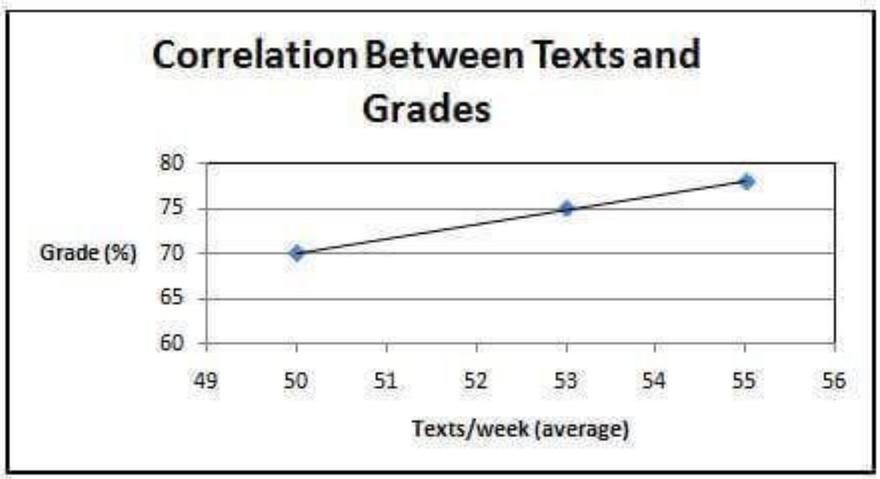

Figure 2. Graph of texts vs. grade correlation.

D. **Factorial ANOVA**: Compares means of more than two groups, each with multiple levels in the independent variable (Cronk, 2011). An example is determining if patients' health improved using one of four different antibiotics either with intravenous fluids or not. In this case it would be called a 4 x 2 factorial ANOVA, because there are two independent variable (antibiotics and intravenous fluids). The first independent variable has four levels (antibiotics) and the second independent variable has two levels (intravenous fluids).

3. **Mixed Design**

A. **Mixed-Design ANOVA**: Compares means of more than two groups, with multiple independent variables (Cronk, 2011). One independent variable must be within-subjects (repeated measures) and one must be between-subjects.

4. **Comparing Means**

A. **Single-Sample t-Test (AKA: z-Test)**: Compares the mean of one sample against the mean of an already known population (Aron, Aron, & Coup, 2005). An example would be comparing the ACT scores of students in one state against that of the entire country.

5. **Covariates**

Covariates are variables which are in some way related to the dependent variable, however are not considered to be independent variables (Cronk, 2011). An example is the effect of gender on reaction time while considering age. Here gender would be the covariate as gender is related to reaction time scores.

A. **ANCOVA**: Compares the means of more than two groups, and allows researchers the ability to remove a covariate (Cronk, 2011). As stated before an example is the effect of gender and age on reaction time. The researcher could remove gender as it is known it is related to reaction time.

6. **More than One Dependent Variable**

In some cases, researchers may have more than one dependent variable. An example is determining the effect of treatment type (chemotherapy, radiation therapy, or combination on both) on tumor markers and tumor size.

A. **MANOVA**: Compares the means of more than two groups when there is more than one dependent variable, and decreases Type I error (Cronk, 2011). As stated before an example is the effect of treatment type (chemotherapy, radiation therapy, or combination on both) on tumor markers and tumor size.

Nonparametric Tests: (Ordinal/Nominal data)

These tests do not assume anything about the shape of the data. These tests are not as strong as the parametric ones (Heiman, 2001).

1. **Within-Subjects Design**

A. **Wilcoxon**: This is the nonparametric version of the dependent samples t-test as it compares the difference of ranks for groups of two with **ONLY** ordinal data (Cronk, 2011). An example is determining participants' ranking for visual preference between two iPad game apps.

B. **Friedman**: This is the nonparametric version of the within-subjects ANOVA as it compares ranks for groups of more than two with **ONLY** ordinal data (Cronk, 2011). An example is determining the rank of three different websites based on user friendliness.

2. **Between-Subjects Design**

A. **Chi Square**: Compares ranks for both two groups and more than two group designs with **ONLY** nominal data (Aron, Aron, & Coup, 2005).

1. **Chi Square Goodness of Fit**-Compares the proportion of the sample to an already existing value. An example is comparing literacy rates in central Missouri against those of the entire state.

2. **Chi Square Test of Independence**-Compares the proportions of two variables to see if they are related or not. An example would be are there similar numbers of baseball and softball athletes enrolled in this semester's general psychology course.

B. **Kruskall-Wallis**: This is the nonparametric version of the between-subjects ANOVA as it determines if multiple samples are from the same population (Cronk, 2011). An example is determining how participants rate a physician based on whether he/she used a computer, book, or no assistance in diagnosing them.

C. **Mann-Whitney**: This is the nonparametric version of the independent samples t-test as it compares ranks for groups of more than two with ONLY ordinal data (Cronk, 2011). An example would be determining if eating Snickers candy bars affects students' grades in math courses.

D. **Correlations**:

1. **Spearman's Rho**: Like the Pearson correlation, this test too determines the strength of the relationship between two variables based on their ranks. An example would be to determine the relationship between physics and English course grades.

Prediction

In some cases, a researcher may want to use one or more variables to predict another variable. In this case, the researcher would use a regression.

1. **Simple Linear Regression**: This is when one known variable is used to predict the value of another variable (Cronk, 2011). An example would be using a student's GRE scores to predict their graduate school GPA.
2. **Multiple Linear Regression**: This is when more than one variable is used to predict the value of another variable (Aron, Aron, & Coup, 2005). An example would be temperature, wind speed, and humidity to predict the amount of precipitation that will fall.

Hypothesis Types

In research there are two main hypotheses (Aron, Aron, & Coups, 2005):

1. **Null hypothesis**: This states that there is **NO** statistical significance between the groups.
2. **Alternative hypothesis**: This states that there **IS** a statistical significance between the groups.

Error Types

No matter how well a study is designed there is always the probability that chance is what caused the results, not the different treatments. When a researcher accepts results that are not accurate that is called an error. There are two main types of error (Aron, Aron, & Coups, 2005):

1. **Type I error**: The alternative hypothesis is accepted when the null hypothesis is actually true
2. **Type II error**: The null is accepted when in fact the alternative hypothesis is true (see Table 3).

Table 3. Accepted/rejected the null when it was either true/false

	True	**False**
Accepted Null	Good decision	Type I
Rejected Null	Type II	Good decision

REFERENCES

Aron, A., Aron, E., & Coups, E. (2005). Statistics for psychology (4th ed.). Pearson, pgs 4, 176-177, 223, 236, 270, 506, 539, 634.

Cronk, B. (2011). How to use SPSS (7th ed.). Pyrczak, pgs 50, 59, 65, 69, 74, 77, 81, 85, 87, 101, 105, 108, 113, 131.

Heiman, G. (2001). Research Methods in Psychology (3rd ed.). Cengage Learning, pgs 206-207 & 271-272.

INTRODUCTION

Also the author of this website has made an effort to ensure the accuracy and completeness of the presented information, the author assumes no responsibility for any errors, missing information, inaccuracies, or inconsistencies.

SPSS

Before you start running participants it is important to set up your data collection sheets. There are a lot of different ways to do a statistical analysis, but in our lab we use SPSS which does a lot of work for you (see Figure 3).

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	Group	Numeric	8	2		{1.00, No Ai...	None	8	Right	Scale
2	Thoroughness	Numeric	8	2		{-3.00, Extr...	None	8	Right	Ordinal
3	Wait	Numeric	8	2		{-3.00, Very...	None	8	Right	Ordinal
4	DiagnosticC...	Numeric	8	2		{-3.00, Very...	None	8	Right	Ordinal
5	Professional...	Numeric	8	2		{-3.00, Extr...	None	8	Right	Ordinal
6	Satisfaction	Numeric	8	2		{-3.00, Extr...	None	8	Right	Ordinal
7	Liability	Numeric	8	2		{-3.00, Extr...	None	8	Right	Ordinal
8	Age	Numeric	8	2		None	None	8	Right	Scale
9	Gender	Numeric	8	2		None	None	8	Right	Nominal

Figure 3. SPSS allows you to set up your spreadsheet by naming variables and also determining the level of measurement. Taken from SPSS.

It is set up similar to an excel spreadsheet in that the DV's you are examining are across the top (see Figure 4).

	Group	Thoroughness	Wait	DiagnosticCapabilities	Professionalism	Satisfaction	Liability	Age	Gender
1	1.00	1.00	-1.00	-1.00	2.00	-2.00	-3.00	19.00	2.00
2	2.00	-1.00	3.00	-3.00	3.00	-2.00	3.00	19.00	2.00
3	3.00	2.00	1.00	2.00	2.00	2.00	1.00	18.00	1.00
4	1.00	-2.00	.00	2.00	2.00	2.00	.00	29.00	2.00
5	2.00	2.00	1.00	1.00	2.00	2.00	-1.00	18.00	2.00

Figure 4. Once your spreadsheet is set up in SPSS you can view it and model your excel spreadsheet after it. Taken from SPSS.

It is good practice to set up both your SPSS spreadsheet and your excel spreadsheet to be the same (see Figure 5). This will allow you to copy and paste your data into SPSS for quick statistical analysis.

	A	B	C	D	E	F	G	H	I
1	Group	Thorough	Wait	Diagnosti	Professio	Satisfact	Liability	Age	Gender
2	1	1	-1	-1	2	-2	-3	19	2
3	2	-1	3	-3	3	-2	3	19	2
4	3	2	1	2	2	2	1	18	1
5	1	-2	0	2	2	2	0	29	2
6	2	2	1	1	2	2	-1	18	2

Figure 5. Model your excel spreadsheet after the one in SPSS. You can simply copy and paste your data into SPSS and quickly perform your analysis.

REFERENCE

The following reference is useful with conducting statistical analysis as it tells you which options to use and the basics for each test covered.

Cronk, B. (2011). How to use SPSS (7th ed.). Pyrczak.

