

Multi-Factor Between-Subjects Designs

Prerequisites

[Introduction to ANOVA](#), [ANOVA Designs](#)

Basic Concepts and Terms

In the [Bias Against Associates of the Obese](#) case study, the researcher was interested in whether the weight of a companion of a job applicant would affect judgments of a male applicant's qualifications for a job. Two [independent variables](#) were investigated: (1) whether the companion was obese or of typical weight and (2) whether the companion was a girl friend or just an acquaintance. One approach would have been to conduct two separate studies, one with each independent variable. However, it is more efficient to conduct one study that includes both independent variables. Moreover, there is a much bigger advantage than efficiency for including two variables in the same study: it allows a test of the *interaction* between the variables. There is an interaction when the effect of one variable differs depending on the [level](#) of a second variable. For example, it is possible that the effect of having an obese companion would differ depending on the relationship to the companion. Perhaps there is more prejudice against a person with an obese companion if the companion is a girl friend than if she is just an acquaintance. If so, there would be an interaction between the obesity factor and the relationship factor.

There are three effects of interest in this experiment:

1. Weight: Are applicants judged differently depending on the weight of their companion?
2. Relationship: Are applicants judged differently depending on their relationship with their companion?
3. Weight x Relationship Interaction: Does the effect of weight differ depending on the relationship with the companion?

The first two effects (Weight and Relationship) are both *main effects*. A main effect of an independent variable is the effect of the variable averaging over the levels of the other variable(s). It is convenient to talk about main effects in terms of *marginal means*. A marginal mean for a variable is the mean of the means of all levels of that variable. For example, the marginal mean for the Obese Condition is the average of the "Girl-Friend-Obese" and the

"Acquaintance-Obese" conditions. Table 1 shows that this marginal mean is equal to the average of 5.65 and 6.15 which is 5.90. Similarly, the marginal mean for Typical is the average of 6.19 and 6.59 which is 6.39. The main effect of Weight is based on a comparison of these two marginal means. Similarly, the marginal means for Girl Friend and Acquaintance are 5.92 and 6.37.

Table 1. Means for All Four Conditions

		Companion Weight		Marginal Mean
		Obese	Typical	
Relationship	Girl Friend	5.65	6.19	5.92
	Acquaintance	6.15	6.59	6.37
	Marginal Mean	5.90	6.39	

In contrast to a main effect which is the effect of a variable averaged across levels of another variable, the *simple effect* of a variable is the effect of the variable at a single level of another variable. The effect of "Weight" at the level of "Girl Friend" is a simple effect of "Weight" and is the difference between the "Girl-Friend Typical" condition and the "Girl-Friend Obese" conditions. The difference is $6.19 - 5.65 = 0.54$. Similarly, the effect of "Weight" at the level of "Acquaintance" is also a simple effect of "Weight" and is the difference between the "Acquaintance Typical" condition and the "Acquaintance Obese" conditions. The difference is $6.59 - 6.15 = 0.44$.

Recall that there is an interaction when the effect of one variable differs depending on the level of another variable. This is equivalent to saying that **there is an interaction when the simple effects differ**. In this example, the simple effects are 0.54 and 0.44.

Tests of Significance

The important questions are not whether there are main effects and interactions in the sample data. Instead, what is important is whether the sample data allow you to conclude about the population. This is where Analysis of Variance comes in. ANOVA tests main effects and interactions for *significance*. An ANOVA Summary Table for these data is shown in Table 2.

Table 2. ANOVA Summary Table

Source	df	SSQ	MS	F	p
--------	----	-----	----	---	---

Weight	1	10.4673	10.4673	6.214	0.0136
Relation	1	8.8144	8.8144	5.233	0.0234
W x R	1	0.1038	0.1038	0.062	0.8043
Error	172	289.7132	1.6844		
Total	175	310.1818			

Consider first the effect of "Weight." The *degrees of freedom* (df) for "Weight" is 1. The degrees of freedom for a main effect is always equal to the number of levels of the variable minus one. Since there are two levels of the "Weight" variable (typical and obese) the df is $2 - 1 = 1$. We skip the calculation of the sum of squares (SSQ) not because it is difficult, but because it is so much easier to rely on computer programs to compute it. The mean square (MS) is the sum of squares divided by the df. The F ratio is computed by dividing the MS for the effect by the MS for error (MSE). For the effect of "Weight," $F = 10.4673/1.6844 = 6.214$. The last column, p, is the probability of getting an F of 6.214 or larger given that there is no effect of weight in the population. The p value is 0.0136 which is quite low and therefore the null [*hypothesis*](#) of no main effect of "Weight" is rejected. The conclusion is that the weight of the companion lowers judgments of qualifications.

The effect "Relation" is interpreted the same way. The conclusion is that being accompanied by a girl friend leads to lower ratings than being accompanied by an acquaintance.

The df for an interaction is the product of the df of variables in the interaction. For the Weight x Relation interaction (W x R), the $df = 1$ since both Weight and Relation have one df: $1 \times 1 = 1$. The p value for the interaction is 0.8043 which is the probability of getting an interaction as big or bigger than the one obtained in the experiment if there were no interaction in the population. Therefore, these data provide no evidence for an interaction. Always keep in mind that the lack of evidence for an effect does not justify the conclusion that there is no effect. In other words, you do not accept the null hypothesis just because you do not reject it.

For "Error," the degrees of freedom is equal to the total number of observations minus the total number of groups. The sample sizes for this experiment are shown in Table 3. The total number of observations is $40 + 42 + 40 + 54 = 176$. Since there are four groups, $df = 176 - 4 = 172$.

Table 3. Sample Sizes for All Four Conditions

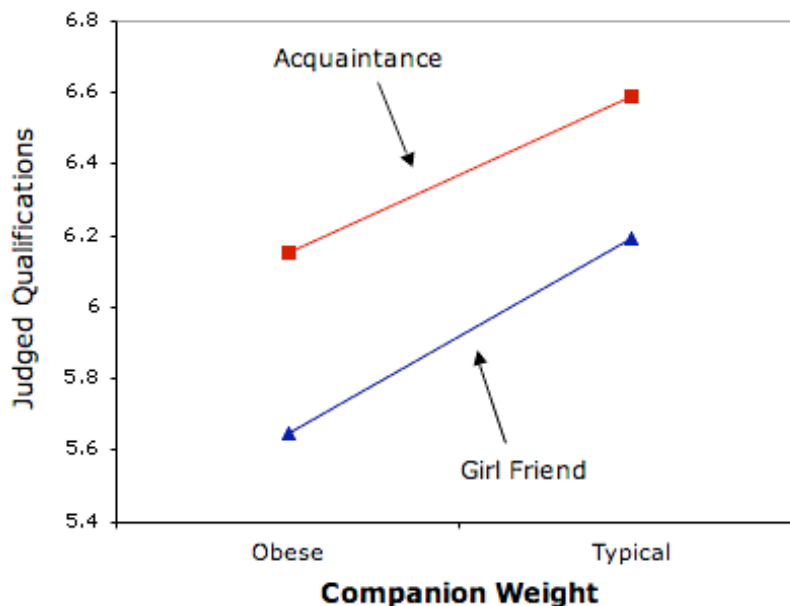
		Companion Weight	
		Obese	Typical
Relationship	Girl Friend	40	42
	Acquaintance	40	54

The final row is "Total." The degrees of freedom total is equal to the sum of all degrees of freedom. It is also equal to the number of observations minus 1, or $176 - 1 = 175$. When there are equal sample sizes, the sum of squares total will equal the sum all other sums of squares. However, when there are unequal sample sizes, as there are here, this will not generally be true. The reasons for this are complex and are discussed in the section [Unequal n](#).

Plotting Means

Although the plot shown in Figure 1 illustrates the main effects and (lack of) interaction clearly, it is often called an *interaction plot*. It is important to carefully consider the components of this plot. First, the dependent variable is on the Y axis. Second, one of the independent variables is on the X axis. In this case, the variable is "Weight." Finally, a separate line is drawn for each level of the other independent variable. It is better to label the lines right on the graph as shown here than with a legend.

Figure 1. An Interaction Plot



If you have three or more levels on the X axis, you should not use lines unless there is some numeric ordering to the levels. If your variable on the X axis is a qualitative variable, you can use a plot such as the one in Figure 2. However, as discussed in the section on [bar charts](#), it would be better to replace each bar with a [box plot](#).

Figure 2. Plot With a Qualitative Variable on the X-axis

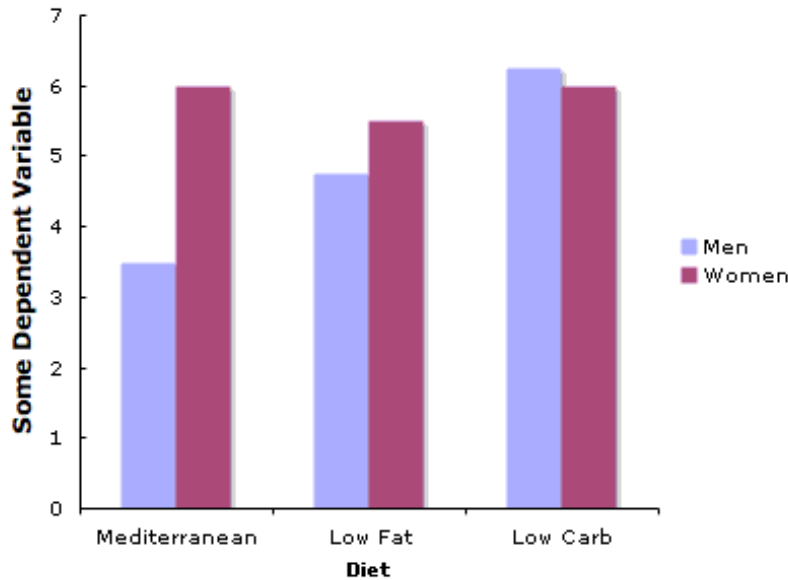
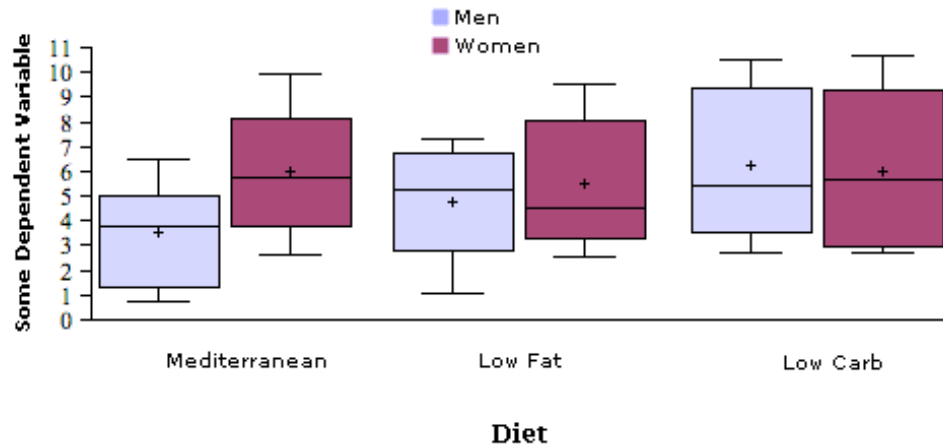


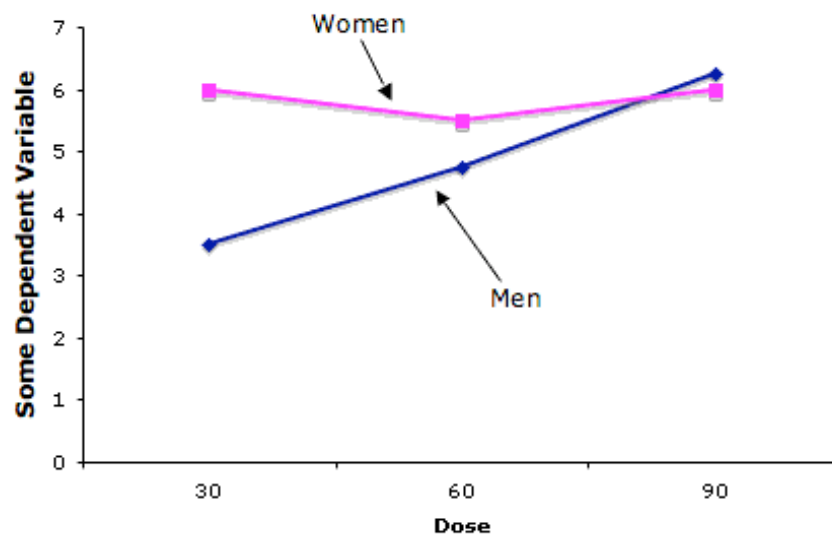
Figure 3 shows such a plot. Notice how it contains information about the medians, quantiles, and minimums and maximums not contained in Figure 2. Most important, you get an idea about how much the distributions overlap from Figure 3 that you do not get from Figure 2.

Figure 3. Box plots.



Line graphs are a common option with more than two levels if the variable is numeric. Figure 4 shows an example. A line graph has the advantage of showing the pattern of interaction clearly. Its disadvantage is that it does not convey the distributional information contained in box plots.

Figure 3. Plot With a Quantitative Variable on the X-axis



An Example with Interaction

The following example was presented in the section on specific comparisons among means. It is also relevant here.

This example uses the made-up data from a hypothetical experiment shown in Table 4. Twelve subjects were selected from a population of high-self-esteem subjects and an additional 12 subjects were selected from a population of low-self-esteem subjects. Subjects then performed on a task and (independent of how well they really did) half were told they succeeded and the other half were told they failed. Therefore there were six subjects in each esteem/success combination and 24 subjects altogether.

After the task, subjects were asked to rate (on a 10-point scale) how much of their outcome (success or failure) they attributed to themselves as opposed to being due to the nature of the task.

Table 3. Data from Hypothetical Experiment on Attribution

		Esteem	
		High	Low
Outcome	Success	7	6
		8	5
		7	7
		8	4
		9	5
		5	6
	Failure	4	9
		6	8
		5	9
		4	8
		7	7
		3	6

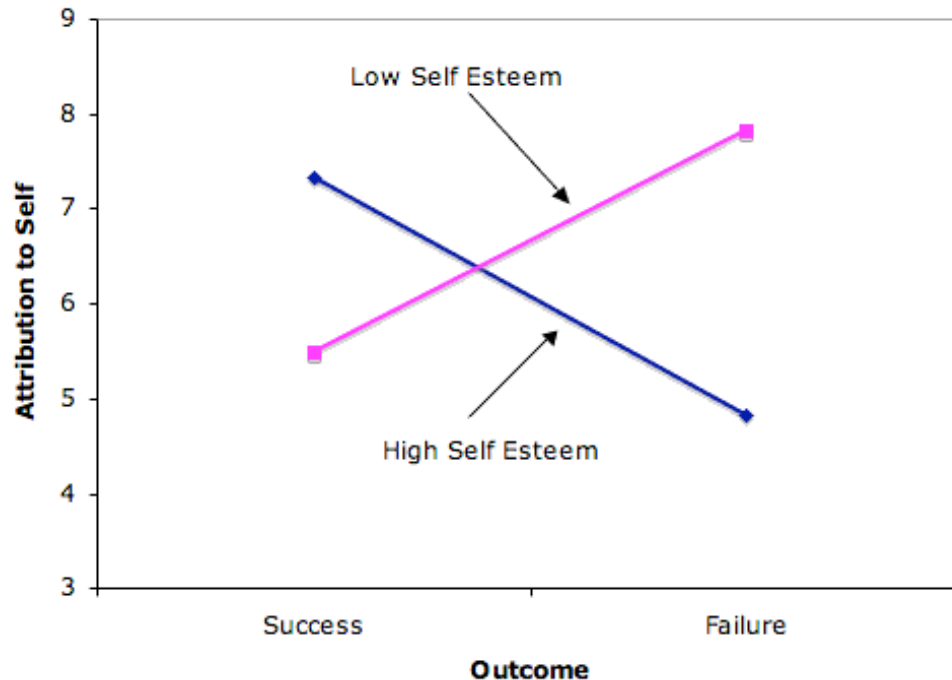
The ANOVA Summary Table for these data is shown in Table 5.

Table 5. ANOVA Summary Table for Made Up Data

Source	df	SSQ	MS	F	p
Outcome	1	0.0417	0.0417	0.0256	0.8744
Esteem	1	2.0417	2.0417	1.2564	0.2756
O x E	1	35.0417	35.0417	21.5641	0.0002
Error	20	32.5000	1.6250		
Total	23	69.6250			

As you can see, the only significant effect is the Outcome x Esteem (O x E) interaction. The form of the interaction can be seen in Figure 4.

Figure 4. Interaction Plot for Made Up Data



Clearly the effect of "Outcome" is different for the two levels of Esteem: For subjects high in self esteem, failure led to less attribution to oneself than did success. By contrast, for subjects low in self esteem, failure led to more attribution to oneself than did success. Notice that the two lines in the graph are not parallel. **Nonparallel lines indicate interaction. The significance test for the interaction determines whether it is justified to conclude that the lines in the population are not parallel.**

Three-Factor Designs

Three-factor designs are analyzed in much the same way as two-factor designs. Table 6 shows the analysis of a [study described by Franklin and Cooley](#) investigating three factors on the strength of industrial fans: (1) Hole Shape (Hex or Round), (2) Assembly Type (Stake or Spun), and (3) Barrel Shape (knurled or smooth). The dependent variable, breaking torque, was measured in foot-pounds. There were eight observations in each of the eight combinations of the three factors.

As you can see in Table 6, there are three main effect, three two-way interactions, and one three-way interaction. The degrees of freedom for the main effects are, as in a two-factor design, equal to the number of levels of the factor minus one. Since all the factors here have two levels, all the main effects

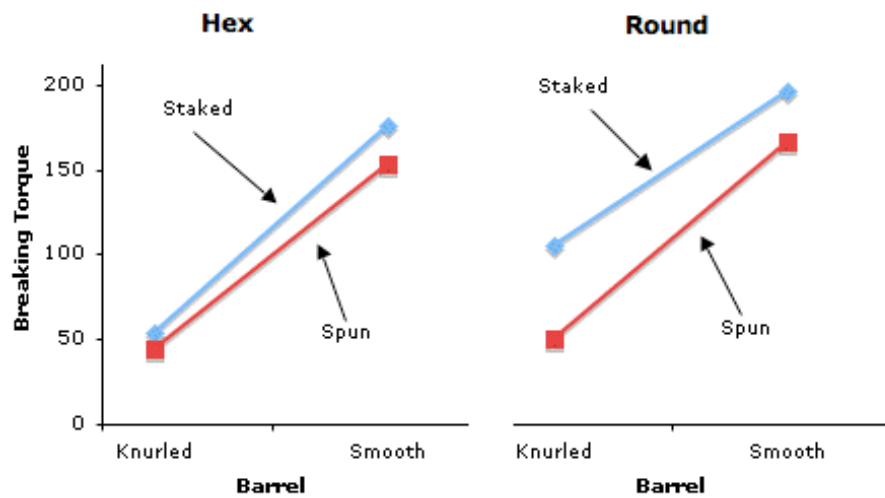
have one degree of freedom. The interaction degrees of freedom are always equal to the product of the degrees of freedom of the component parts. This holds for the three-factor interaction as well as the two-factor interactions. The error degrees of freedom is equal to the number of observations (64) minus the number of groups (8) and is 56.

Table 6. ANOVA Summary Table for Fan Data

Source	df	SSQ	MS	F	p
Hole	1	8258.27	8258.27	266.68	<0.0001
Assembly	1	13369.14	13369.14	431.73	<0.0001
H x A	1	2848.89	2848.89	92.00	<0.0001
Barrel	1	35.0417	35.0417	21.5641	<0.0001
H x B	1	594.14	594.14	19.1865	<0.0001
A x B	1	135.14	135.14	4.36	0.0413
H x A x B	1	1396.89	1396.89	45.11	<0.0001
Error	56	1734.12	30.97		
Total	63	221386.91			

A three-way interaction means that the two-way interactions differ as a function of the level of the third variable. The usual way to portray a three-way interaction is to plot the two-way interactions separately. Figure 5 shows the Barrel (knurled or smooth) x Assembly (Staked or Spun) separately for the two levels of Hole Shape (Hex or Round). For the Hex Shape, there is very little interaction with the lines being close to parallel with a very slight tendency for the effect of Barrel to be bigger for Staked than for Spun. The two-way interaction for the Round Shape is different: The effect of Barrel is bigger for Spun than for Staked. The finding of a significant three-way interaction indicates that this difference in two-way interactions is not due to chance.

Figure 5. Plot of the three-way interaction.



Formatting data for Computer Analysis

The data in Table 3 have been reformatted in Table 7. Note how there is one column to indicate the level of outcome and one column to indicate the level of esteem. The coding is as follows:

High-self-esteem: 1
Low self esteem: 2

Success: 1
Failure: 2

Table 7. Attribution Data
Reformatted

outcome	esteem	attrib
1	1	7
1	1	8
1	1	7
1	1	8
1	1	9
1	1	5
1	2	6
1	2	5
1	2	7
1	2	4
1	2	5
1	2	6
2	1	4
2	1	6
2	1	5
2	1	4

2	1	7
2	1	3
2	2	9
2	2	8
2	2	9
2	2	8
2	2	7
2	2	6

To use [Analysis Lab](#) to do the calculations, you would copy the data and then

1. Click the "Enter/Edit User Data" button (You may be warned that for security reasons you must use the keyboard shortcut for pasting data).
2. Paste your data.
3. Click "Accept Data"
4. Click the "Advanced" button next to the "Anova button."
5. Select "attrib" as the dependent variable and both outcome and esteem as "group" variables.
6. Click the "do ANOVA button."