Chapter 12 Statistical Foundations: Analysis of Variance

There are many instances when a researcher is faced with the task of examining three or more groups for meaningful (i.e., statistical) differences. Suppose we wanted to test five groups against each other in order to determine if any one or more of them were meaningfully different from each other. To do this, we would need to perform 10 pairwise t-tests. Recall Type I error. Every time a statistical test is performed, we accept a degree of error. For a pairwise t-test comparison, we set $\alpha = .05$, when all ten t-tests are performed, $\alpha = .29$. This ballooning alpha is called experiment-wise error; so, our Type I error then becomes almost 30%, which is far too high.

As the number of pairwise comparisons increase so does experiment-wise or family wise error; both terms mean the same. Since we want to keep Type I error to $\alpha = .05$, we test variance instead of group means when three or more groups are involved. The testing of variance is as accurate as testing group means. This variance test is called ANOVA.

I. ANOVA Basics

A. Recall our discussion of the standard deviation and what it reports. Remember, that variance is the standard deviation squared and represents the spread of all individual scores around a group mean. Formula 12.1 is the formula for computing variance.

$$\sigma^2 = \frac{\Sigma(X - \bar{X})^2}{n-1}$$

Where σ^2 = variance; $\Sigma(X - \bar{X})^2$ = sum of squares; $n-1$ = degrees of freedom

1. The sources of variance are the (1) effect of the independent variable(s) on the dependent variable and (2) chance (or variation that just happens).

2. As the ANOVA procedure is applied to three or more groups, we find variance between groups and within groups.
   a. The variance between groups is referred to as between-groups or effect variance as the variance is attributed to the effect of the independent variable(s) on the dependent variable, but contains some error. It is an estimate of how different the means of each set of scores are.
   b. The within group variance is attributed to chance, random variance, or the inherent variability of the measures employed, etc. and is called error variance. It is an estimate of the extent to which scores within a particular group vary.

3. If the effect variance is larger than the error variance, then we suspect that the groups differ, meaningfully or that H_0 is likely to be rejected. To make the comparison, a ratio is formed between the effect and the error variance.

$$\frac{EffectVariance}{ErrorVariance}$$

When the ratio is larger than 1.0, effect variance is greater than error variance. When the ratio is less than 1.0, error variance is greater than effect variance. The larger the ratio, the more likely the groups differ.

4. The F-ratio
   a. The F-ratio is used to examine the effect variance and error variance comparison. As with any statistical test, there are assumptions; these are
      (1) Scores are independent.
      (2) Scores are normally distributed, i.e., if plotted would fall on the SNC.
      (3) Subjects were randomly selected and assigned.
      (4) Group variances are equal.
      (5) Scores are at least on an interval scale.

   b. The F-ratio is

$$F = \frac{EffectVariance}{ErrorVariance} = \frac{MS_{Treatment}}{MS_{Error}} = \frac{MS_{BetweenGroups}}{MS_{WithinGroups}}$$

   Remember that the effect variance is composed of both treatment and error variance; while, error variance is 100% error. Depending on textbook or article author, effect variance may be referred to as Treatment variance or Between Groups variance. Error variance may also be referred to as Within Groups variance.

   c. As with other statistical tests, the F value must achieve a particular size before statistical significance is determined. A table of critical values is used. The needed size is dependent upon alpha level, numerator degrees of freedom, and denominator degrees of freedom. If the calculated F value is greater than the critical table value, then H_0 is rejected. If the F value is equal to or less than the critical table value, the H_0 is not rejected.

   d. When the F-ratio is applied, the convention is to refer to the independent variable as a factor. If three groups are involved in the analysis, then the factor is said to have three levels. If six groups are involved, the factor has six levels, etc.

B. Effect Size

1. For the <u>one-way analysis of variance (ANOVA)</u>, Cohen (1988, pp. 285-287) recommends ∫ = 0.10 (small effect), ∫ = 0.25 (medium effect) and ∫ = 0.40 (large effect).
2. As with the t-tests, the interpretation would be: “The independent variable had a (small, medium, or large) effect on the dependent variable.” A formula for computing effect size and an example are found in Stevens (1999, p. 132).

II. One-Way (Simple or Factor) Independent Groups Analysis of Variance

A. The phase, "one-way" indicates that one independent variable's effect on a single dependent variable is being tested. The one-way ANOVA involves 3+ group means or more technically, one independent variable with 3 or more "levels."

B. **Case 12.1**: We'll demonstrate the computational sequence through an example. Your organization wants to determine which management training program is most effective. New management trainee hires have been randomly assigned to one of three training groups. At the end of the four week course, a management aptitude test was administered with a score range of 1 to 20 (highest possible). Data are presented in Table 12.1.

Table 12.1
Data Table

Group One		**Group Two**		**Group Three**	
X	X^2	X	X^2	X	X^2
10	100	4	16	15	225
8	64	7	49	10	100
12	12	6	36	14	196
14	196	8	64	15	225
13	169	6	36	11	121
11	121			11	121
16	256			9	81
				13	169
$n_1 = 7$		$n_2 = 5$		$n_3 = 8$	
$\sum x_1 = 84$		$\sum x_2 = 31$		$\sum x_3 = 98$	
$\sum x_1^2 = 1050$		$\sum x_2^2 = 201$		$\sum x_3^2 = 1238$	

C. Explanation of Terms and Symbols
1. The "X" represents an individual value or score. "X^2" represents X squared.
2. There are three groups or levels $n_1 = 7$; $n_2 = 5$; and $n_3 = 8$, members, where the subscript represents the group number.
3. The sum of scores for group one is $\sum x_1 = 84$; group two, $\sum x_2 = 31$; and group three, $\sum x_3 = 98$.
4. The sum of squares for each group are $\sum x_1^2 = 1050$, $\sum x_2^2 = 201$, and $\sum x_3^2 = 1238$.
5. The total number of subjects is twenty (N = 20). The total raw score ($\sum X$) is 213 and total sum of squares ($\sum X^2$) is 2,489.

D. Now we must solve for the Sums of Squares
1. Compute Sums of Squares Total (SS_T), using Formula 12.2a

$$SS_T = \sum X^2 - \frac{\left(\sum X\right)^2}{N} \qquad SS_T = 2489 - \frac{(213)^2}{20}$$

$SS_T = 220.55$

2. Compute Sums of Squares Between Groups (SS_B), using Formula 12.2b

$$SS_B = \frac{\left(\sum X_1\right)^2}{n_1} + \frac{\left(\sum X_2\right)^2}{n_2} \bullet\bullet\bullet \frac{\left(\sum X_k\right)^2}{N_k} - \frac{\left(\sum X\right)}{N}$$

$$SS_B = \frac{(84)^2}{7} + \frac{(31)^2}{5} + \frac{(98)^2}{8} - \frac{(213)^2}{20}$$

SS_B = 132.25 [This is treatment or effect variance. Spatz (2011, p. 240) uses SS_{treat}.]

3. Compute Sums of Squares Within Groups (SS_W), using Formula 12.2c

$SS_W = SS_T - SS_B$ or 220.55 – 132.25 = 88.3

[SS_W = 88.3 is error variance. Spatz (2011, p. 240) uses SS_{error}.]

E. Now we solve for df_B and df_W
   1. df_B = 2 (K–1 or 3-1) K = # of groups

   2. df_W = 17 (N-K or 20-3)

   3. Check: $df_B + df_W$ = N-1 or 2 + 17 = 19

F. We next compute the Mean Sums of Squares for the Between and Within Group variance (MS_B and MS_W, respectively)

   1. Compute MS_B, (effect or treatment) using Formula 12.3a

$$MS_B = \frac{MS_b}{df_b} = \frac{132.25}{2} = 66.13$$

   2. Compute MS_W (error), using Formula 12.3b

$$MS_W = \frac{MS_W}{df_W} = \frac{88.3}{17} = 5.19$$

G. In the next step, we compute the F-ratio and Summary Table 12.2
   1. F-ratio = between group/within group or effect variance/error variance
   2. F Summary Table

Table 12.2
F-Ratio Summary Table

Variance	**SS**	***df***	**MS**	**F**
Between Group (Effect or Treatment)	132.25	2	66.13	12.74*
Within Group (Error)	88.30	17	5.19	
Total	220.55	19		

*$F_{(2, 17)} = 12.74$, $p < .05$

3. Interpret the F-Ratio Summary Table
   (a) We see in the table that the F-ratio of 12.74 is statistically significant at $\alpha = .05$. This is determined by locating the numerator and denominator degree of freedom at a predetermined alpha level. The convergence of the two is the F critical value, in this case 3.59. The *df* numerator is 2 and the *df* denominator is 17; see Table 12.2.
   (b) Since $F_{2, 17, .05} = 12.74$ > the critical value of $F_{2, 17, .05} = 3.59$, we reject H_o: $\mu_1 = \mu_2 = \mu_3$ and accept the H_1: $\mu_1 \neq \mu_2 \neq \mu_3$. Using the *GraphPad Software QuickCalcs* (n.d.) calculator, we see that the exact probability is 0.0004; we reject the null hypothesis as the exact probability is less than the specified alpha level. In the calculator numerator *df* is "DFn" and the denominator *df* is "DFd."
   (c) A critical F value table, for alpha levels 0.05 and 0.01 is located at <http://www.sussex.ac.uk/Users/grahamh/RM1web/F-ratio%20table%202005.pdf>

4. To determine which groups are different, we apply Tukey HSD post hoc test to determine which groups are different.

H. Multiple Comparisons, AKA, A priori or Post hoc testing
   1. If a planned number of pairwise comparisons are determined before (or a priori) the omnibus F-test is applied, then the specialized multiple comparison test can be selected. Please refer to Stevens (1999) for choices. A common strategy is to perform the multiple comparisons after the omnibus F-test is applied and statistically significant differences are found; this is called post hoc or "after the fact " testing.

   2. Spatz (2011, pp. 248-250) recommends the Tukey HSD (Honestly Significant Differences) which is designed for all possible pairwise comparisons. See Formula 12.4a, 12.4b, and 12.4c.

   3. Computing Tukey HSD for Equal Size Groups
      a. Compute Tukey HSD for Equal Size Groups using Formula 12.4a and Formula 12.4b (Spatz, 2011, p. 148)

$$HSD = \frac{\overline{X}_i - \overline{X}_j}{s_{\bar{x}}}$$

Formula 12.4a

$$s_{\bar{x}} = \sqrt{\frac{MS_{error}}{N_i}}$$

Formula 12.4b

where: $\bar{x}_i$ = mean of the i'th group; $\bar{x}_j$ mean of the j'th group; and N_i = number in each i'th group

b. Compute the Tukey HSD for Unequal Size Groups using Formula 12.4a and Formula 12.4c. See below.

Table 12.3
Data Table

$\bar{x}_1 = 12.0$	$\bar{x}_2 = 6.2$	$\bar{x}_3 = 12.25$
$n_1 = 7$	$n_2 = 5$	$n_3 = 8$

(1) Compute $s_{\bar{x}}$, using Formula 12.4c (Spatz, 2011, p. 250)

$$s_{\bar{x}} = \sqrt{\frac{MS_{error}}{2}\left(\frac{1}{N_1}\right)+\left(\frac{1}{N_2}\right)} = \sqrt{\frac{5.19}{2}\left(\frac{1}{7}\right)+\left(\frac{1}{5}\right)} = \sqrt{.8898} = .9433$$

(2) Compute HSD for $\overline{X}_1 - \overline{X}_2$ using Formula 12.4a

$$HSD = \frac{\overline{X}_1 - \overline{X}_2}{s_{\bar{x}}} = \frac{12 - 6.2}{.9433} = 6.1486$$

(3) Compare the computed HSD to the critical value with $df = 17$ for MS_{error} and three levels of the independent variable at $\alpha = .05$, the critical Tukey HSD value is 3.63 (Spatz, 2011. p. 398). Since the computed HSD of $6.1486 \geq 3.63$, we reject the H_0 which posits no differences between means, $\overline{X}_1$ (12) and $\overline{X}_2$ (6.2); the observed difference was due to something other than chance. We repeat this process for the other two means. See Table 12.4

Table 12.4
Tukey HSD Data Table

Comparisons	**Means**	**Computed HSD**	**HSD Critical Value**	**Decision**
$\overline{X}_1$ and $\overline{X}_2$	84-31	6.1486	3.63	Reject H_0
$\overline{X}_1$ and $\overline{X}_3$	84-98	/.2998/	3.63	Retain H_0
$\overline{X}_2$ and $\overline{X}_3$	31-98	/6.588/	3.63	Reject H_0

c. Now, we concludeusing Table 12.4
   (1) We see that groups one and two are statistically different as are groups two and three. Groups one and three are not statistically different.
   (2) We could conclude that training programs one and three are equally effective and both are more effective than training program three and the absolute value of their means are similar.

4. A shortcut, only if all group sizes are equal, to avoid having to compute all pair-wise comparisons is to pick an intermediate mean difference and test it. If it is statistically significant, then larger mean differences are going to be as well.

5. Effect Size Estimation for Pairwise Comparisons
   a. Now that we know the groups are different, we now need to determine whether the differences are small, medium or large.
   b. According to Spatz (2011, pp. 251-252), we can use Cohen's ***d*** to compute an effect size for any 2 pairwise comparisons, using Formula 12.5

$$d = \frac{\bar{X}_i - \bar{X}_j}{\hat{s}_{error}}$$

   where:
   $\bar{X}_i$ = mean for Group "i"
   $\bar{X}_j$ = mean for Group "j"
   $\hat{s} = \sqrt{MS_{error}}$

   c. Based on Table 12.4, we compute the effect sizes for $\bar{X}_1$ - $\bar{X}_2$ and $\bar{X}_2$ - $\bar{X}_3$. $\hat{s} = \sqrt{MS_{error}} = \sqrt{5.19} = 2.278$

$$d = \frac{\bar{X}_1 - \bar{X}_2}{\hat{s}_{error}} = \frac{6.1486}{2.278} = 2.70 \qquad d = \frac{\bar{X}_2 - \bar{X}_3}{\hat{s}_{error}} = \frac{|6.588|}{2.278} = 2.89$$

   Both effect sizes were quite large. The selection of either Program 1 or Program 3 should depend on other factors such as cost, ease of implementation, etc. since both effect sizes were so large.

6. The Omnibus Effect Size Estimate: The Index ∫
   a. If the ANOVA failed to produce a statistically significant result and you or a researcher expects that additional resource investment might reveal a statistically significant result, the Index ∫ might prove useful. See Formula 12.6 (Stevens, 1999, p. 132).

$$\int = \sqrt{\frac{(k-1)F}{N}}$$

(1) where: ∫ = estimated effect size; k = number of groups; F = computed F ratio and N = total number of subjects.

(2) When there are unequal group sizes, first, compute the average group size; next, to determine N, multiply the average group size by the number of groups. For the example, average group size was 6.7; so, N = 20 (6.7 x 3).

(3) Compute ∫

$$\int = \sqrt{\frac{(k-1)F}{N}} = \sqrt{\frac{(3-1)12.74}{20}} = \sqrt{\frac{25.48}{20}} = \sqrt{1.274} = 1.13$$

Spatz (2011, p. 253) provides interpretative guidelines ∫ = 0.10 (Small Effect), ∫ = 0.25 (Medium Effect), and ∫ = 0.40 (Large Effect). So, we've observed a large overall "Omnibus" effect. Of course, we expected this given the large F ratio (12.74) computed.

II. One Factor Correlated (Repeated or Randomized Block) ANOVA

A. When a researcher, evaluator, or manager, analyzes data from natural pairs, matched pairs, or repeated measures, the One Factor Correlated Measures Analysis of Variance procedure is employed. The single independent variable must have at least three levels. The dependent samples t-test uses 2-levels of on independent variable.

1. We expect variation between groups within this procedure and within groups, just as in the One Way (Factor) Independent Group ANOVA. However, a key difference is that within the Correlated ANOVA, the same subjects contribute multiple scores. Now, we have three sources of variance:
   a. The treatment or between groups variance (independent variable)
   b. Error or within group variance
   c. Subject variance
2. Remember, the F-test is based on a ratio between the treatment or between groups variance and the error or within group variance. Thus, the subject contributed variance needs to be partitioned out, which will reduce the within groups or error variance so that only it remains.
3. The computational sequence is similar to that of the One Way or One Factor ANOVA; but, there is an additional variance partition step, subjects' sums of squares or $SS_{subjects}$.

4. One Factor Correlated ANOVA Assumptions are
   a. Scores are normally distributed along the standard normal curve.

   b. The covariance matrix has sphericity.
   c. Subjects are randomly selected and assigned.
   d. Scores are correlated.

5. Advantages and Disadvantages
   a. Advantages
      (1) The One Factor Correlated ANOVA is efficient as research time and resource expenditures are reduced. There is less recruitment, administration, explanation, and debriefing investment.
      (2) Since the error (MS_w or MS_{error}) term is reduced by partitioning subjects' contributed variance, the design is more powerful.
   b. Disadvantages
      (1) The possibility of a carryover effect (i.e., a subject's performance on one IV level (e.g., pretest) affects performance on another level (e.g. posttest). If a "carry over" effect is expected, use matching.
      (2) There are different forms of the One Factor Correlated ANOVA and terminology varies depending on discipline. Sometimes confusion is caused.

6. The design described here is suitable for circumstances where the researcher or evaluator selects the independent variable levels (also called a fixed effects model) and subjects are randomly selected and assigned. This is the most utilized version used in education and training.

B. One Factor Correlated ANOVA Computation
   1. **Case 12.2**: A staff training and development specialist (trainer) is interested in assessing trainee satisfaction among three differing instructional styles (direct instruction, active learning, and self-directed study) as she is planning to develop a new sequence of courses and wants the information to inform design decisions. The trainer has developed an Instructional Satisfaction Scale (ISS) which has been validated and was found to be reliable. The ISS was administered to a group of new hires who were completing three training modules, composed of similar types of knowledge and required similar intellectual skill levels; but the instructional style was varied within each module. The ISS score range is from 0 to 100. The higher score indicates greater instructional satisfaction. Summary information is found in Table 12.5.

2. Calculate the Sums of Squares (Spatz, 2011, p. 260)

$$SS_{tot} = \sum X_{tot}^2 - \frac{(\sum X_{tot})^2}{N_{tot}} = 93,034 - \frac{(1164)^2}{15} = 93,034 - 90,326.4 = 2,707.6$$

   a. Calculate SS $_{tot}$ (equals SS_T or Total Variance) using Formula 12.6a

Table 12.5
Data Table

Subjects	Direct Instruction x_1	x^2	Active Learning x_2	x^2	Self-Directed Study x_3	x^2	ΣX_{tot}	$(\Sigma X_{tot})^2$
s_1	55	3025	58	3364	62	3844	175	30625
s_2	69	4761	70	4900	72	5184	211	44521
s_3	73	5329	74	5476	76	5776	223	49729
s_4	91	8281	90	8100	94	8836	275	75625
s_5	90	8100	93	8649	97	9409	280	78400
$\Sigma x/\Sigma x^2$	Σx_1=378	Σx^2=29496	Σx_2=385	Σx^2=30489	Σx_3=401	Σx^2=33049	ΣX_{tot} =1164	
$\bar{X}$	75.6		77.0		80.2		$(\sum X_{tot})^2$ = 278,9000	
ΣX_{tot} = 1164 (378+385+401)					ΣX^2_{tot} = 93,034 ($55^2 + 58^2 + 62^2 \ldots 90^2 + 93^2 + 97^2$)			

   b. Calculate SS $_{sub}$ (Subjects Sums of Squares or variance due to subjects) using Formula 12.6b. N_k = number of treatments. (Spatz, 2011, p. 260)

$$SS_{sub} = \sum\left[\frac{(\sum X_k)^2}{N_k}\right] - \frac{(\sum X_{tot})^2}{N_{tot}} = \frac{278,900}{3} - \frac{(1164)^2}{15} = 92,966.67 - 90,326.4 = 2,640.27$$

2. Calculate the Sums of Squares (Spatz, 2011, p. 260)
   a. Calculate SS $_{tot}$ (equals SS_T or Total Variance) using Formula 12.6a

$$SS_{tot} = \sum X_{tot}^2 - \frac{(\sum X_{tot})^2}{N_{tot}} = 93,034 - \frac{(1164)^2}{15} = 93,034 - 90,326.4 = 2,707.6$$

   b. Calculate SS $_{sub}$ (Subjects Sums of Squares or variance due to subjects) using Formula 12.6b. N_k = number of treatments. (Spatz, 2011, p. 260)

$$SS_{sub} = \sum\left[\frac{(\sum X_k)^2}{N_k}\right] - \frac{(\sum X_{tot})^2}{N_{tot}} = \frac{278,900}{3} - \frac{(1164)^2}{15} = 92,966.67 - 90,326.4 = 2,640.27$$

   c. Calculate SS $_{treatment}$ (Treatment Sums of Squares i.e., Between Groups Variance) using Formula 12.6c. N_t = number in each treatment group. (Spatz, 2011, p. 260)

$$SS_{treatment} = \sum\left[\frac{(\sum X_t)^2}{N_t}\right] - \frac{(\sum X_{tot})^2}{N_{tot}} = \frac{(378)^2}{5} + \frac{(385)^2}{5} + \frac{(401)^2}{5} - \frac{(1164)^2}{15} =$$

$$SS_{treatment} = 90,382 - 90,326.4 = 55.6$$

d. Calculate SS error (Error Sums of Squares or Within Groups Variance using Formula 12.6d. (Spatz, 2011, p. 260)

$$SS_{error} = SS_{tot} - SS_{sub} - SS_{treatment} = 2707.6 - 2,640.27 - 55.6 = 11.73$$

e. Calculate degrees of freedom (*df*) (Spatz, 2011, p. 261)
(1) $df_{tot} = N_{tot} - 1$ or 15-1 = 14

(2) $df_{sub} = N_t - 1$ or 5-1 = 4

(3) $df_{treat} = N_k - 1$ or 3-1 = 2

(4) $df_{error} = (N_t - 1)(N_k - 1) = (4)(2) = 8$

f. Construct the F-Ratio Summary Table

Table 12.6
F-Ratio Summary Table

Variance	**SS**	***df***	**MS**	**F**
Subjects	2,604.27	4		
Treatment (Between Group or Effect)	55.6	2	27.8	18.963*
Error (Within Group)	11.73	8	1.466	
Total	2,707.6	14		

*$F_{(2, 8)} = 18.963$, $p < .05$

g. Interpret the F-Ratio Summary Table
(1) Since the test statistic $F_{(2, 8)} = 18.963$ is greater than the critical value of $F_{(2, 8)} = 4.46$. The *df* denominator is 8 and the *df* numerator is 2.
(2) We reject the $H_o: \mu_1 = \mu_2 = \mu_3$ and accept the $H_1: \mu_1 \neq \mu_2 \neq \mu_3$. Using the *GraphPad Software QuickCalcs* (n.d.) calculator, we see that the exact probability is 0.0009; we reject the null hypothesis as the exact probability is less than the specified alpha level.

(3) To determine which groups are different, we apply Tukey HSD post hoc test.

h. Construct Tukey HSD Data Table

Table 12.7
Tukey HSD Data Table

Comparisons	Means	Computed HSD	HSD Critical Value	Decision
$\bar{X}_1$ and $\bar{X}_2$	75.6-77.0	|2.31|	4.04	Retain H_0
$\bar{X}_1$ and $\bar{X}_3$	75.6-80.2	|7.59|	4.04	Reject H_0
$\bar{X}_2$ and $\bar{X}_3$	77.0-80.2	|5.28|	4.04	Reject H_0

(1) Calculate Tukey HSD for $\bar{X}_1$ and $\bar{X}_2$ (Spatz, 2011, p. 263) using Formula 12.4 and 12.7. The remaining comparisons are computed in the same manner. $\bar{X}_1$ = Direct Instruction; $\bar{X}_2$ = Active Learning; & $\bar{X}_3$ = Self-directed Study.

$$HSD = \frac{\bar{X}_i - \bar{X}_j}{s_{\bar{x}}}$$

Formula 12.4

$$s_{\bar{x}} = \sqrt{\frac{MS_{error}}{N_t}}$$

Formula 12.7
N_t = number in each treatment group

$$s_{\bar{x}} = \sqrt{\frac{MS_{error}}{N_t}} = \sqrt{\frac{1.466}{4}} = \sqrt{0.367} = 0.606$$

$$HSD = \frac{\bar{X}_i - \bar{X}_j}{s_{\bar{x}}} = \frac{75.6 - 77.0}{0.606} = \frac{-1.4}{0.606} = -2.31$$

(2) Compare the computed HSD to the critical value with *df* = 8 for MS_{error} and three levels of the independent variable at $\alpha = .05$, the critical Tukey HSD value is 4.04 (Spatz, 2011. p. 398). Since the computed HSD of $|2.31| < 4.04$, we retain the H_0 which posits no statistically significant differences between the means, $\bar{X}_1$ (75.6) and $\bar{X}_2$ (77); the observed difference is due to chance. We repeat this process for the other two means and report the decision in Table 12.7.

(3) It would appear that trainees prefer self-directed study over direct instruction or active learning.

III. Factorial ANOVA

A. Introduction (Adapted from Spatz, 2011, pp. 268-276)

1. We have studied the independent and dependent samples t-tests and the One Way or One Factor Independent Samples and Repeated Measures ANOVA.
   a. The t-tests measured the effects one independent variable of 2 levels on a single dependent variable.
   b. The One Way or One Factor ANOVA assessed the effect of one independent variable of 3 or more levels on one dependent variable. The “One Way” or “One Factor” refers to one independent variable.
   c. The Factorial ANOVA measures the effect of two (2) independent variables of 2 or more levels on one dependent variable. Not only does, Factorial Analysis of Variance provide information on both independent variables, but also their interaction.
   d. Factorial ANOVA designs are used when an evaluator or researcher thinks the answer to a question depends on how one independent variable’s

(Factor A) effect on the dependent variable might be dependent on the effect of a second independent variable (Factor B).

2. Illustrations when we might use Factorial Analysis of Variance
   a. Generally, do the Body Mass Index (BMI) of men and women differ? The answer to this question might be partially dependent on level of exercise. So we'd set up the analysis with BMI value as the dependent variable; Factor A as gender with two levels: men and women; and Factor B as exercise with 3 levels: 0-2, 3-5, and 6-7 times weekly. This would be a 3 x 2 Factorial Design; see Table 12.8a.

Table 12.8a
3 x 2 Factorial Design

	Factor A (Gender)	
Factor B (Exercise)	Men	Women
0-2 Times Weekly	BMI Score	BMI Score
3-5 Times Weekly	BMI Score	BMI Score
6-7 Times Weekly	BMI Score	BMI Score

   b. What causes the test score gap between 11^{th} grade White and African American students? "Test scores" is the dependent variable. Factor A (Race) has 2 levels, White and African-American. Conventional wisdom suggests that family income, as an indicator of poverty, might contribute to the "test score gap;" so, we can use "Free/Reduced Lunch" status as an indicator of poverty or Factor B, with 2 levels. This is a 2 X 2 Factorial Design; see Table 12.8b.

Table 12.8b
2 x 2 Factorial Design

	Factor A (Race)	
Factor B (Poverty)	White	African-American
On Free/Reduced Lunch	Test Score	Test Score
No Free/Reduced Lunch	Test Score	Test Score

   c. Suppose undergraduate college students are faced with a "Rising Junior" basic skills competency test which must be passed before college sophomores can become juniors. The question could be, "Is there an effect on the earned test score (dependent variable) by a students' major (Factor A) and hours studied, expressed in 3 levels of hours studied (Factor B) for the examination?" This is a 3 x 2 Factorial Design; See Table 12.8c.

   d. Virtually every organization is concerned with providing superior customer service to ensure customer satisfaction. We suspect that customer service training results in varying levels of mastery, at least when expressed as a test score. But do these varying levels of mastery actually impact customer satisfaction? To answer this question, we can sort recent CSR training program completers into 3 mastery (Factor A) levels: low, medium, and high. The dependent variable can be measured using the Customer Satisfaction Index (CSI). However, a colleague argues that prior customer service experience (Factor) also impacts customer satisfaction; so, the decision is made to measure prior program completer experience. The experience is expressed in years, 1-3, 4-5, and 5-7. Program completers with no prior experience or those with 8 or more years were not included in the study. This is a 3 x 3 Factorial Design; see Table 12.8d.

Table 12.8c
3 x 2 Factorial Design

	Factor A (Major)	
Factor B (Hours Studied)	Science	English
0-10 hours	Test Score	Test Score
11-19 hours	Test Score	Test Score
20-29 hours	Test Score	Test Score

Table 12.8d
3 x 3 Factorial Design

	Factor A (Mastery)		
Factor B (Experience)	Low	Medium	High
1-3 years	CSI Score	CSI Score	CSI Score
4-5 years	CSI Score	CSI Score	CSI Score
5-7 years	CSI Score	CSI Score	CSI Score

e. A multi-national company has among its corporate values, “people centered leadership.” It is interested in learning how its hourly employees view their immediate manager. Using a consultant, the company has constructed and validated the People Centered Leadership Scale (PCLS) which measures perceived people centered leadership, the dependent variable. Two members of the evaluation team point out that culture (Factor A) influences perceptions of management and another asserts that extroverted managers (Factor B) are likely to affect respondent perceptions. Factor A has 4 levels (countries where it has manufacturing plants) and Factor B has 3 levels. This is a 3 X 4 Factorial Design; see Table 12.8e.

Table 12.8e
3 x 4 Factorial Design

Factor B	**Factor A (Culture)**			
(Extroversion)	US	Egypt	Brazil	China
Low	PCLS Score	PCLS Score	PCLS Score	PCLS Score
Medium	PCLS Score	PCLS Score	PCLS Score	PCLS Score
High	PCLS Score	PCLS Score	PCLS Score	PCLS Score

3. Factorial ANOVA Notation and Terms: Explanation of Table 12.8
   a. The notation “2 x 3” means that one factor (independent variable) has two levels and the second factor (independent variable) has three levels. So, the “2” refers to the number of rows and “3” refers to the number of columns (R x C); see the shaded portion of Table 12.9a.
   b. Note that there are six “active” cells (shaded); in Cell A_1B_1, subjects are exposed to Level 1 of each independent variable or factor. Subjects in Cell A_2B_1 are exposed to Levels A_2 and B_1 and so on. No subject is exposed to more than one dual combination of Factor A and Factor B. Level A1 produces a mean score on the dependent variable, $\bar{X}_{A1}$, as does each of the other 5 levels.

Table 12.9a
2 x 3 Factorial Design

	Factor A			
Factor B	A_1	A_2	A_3	Factor B $\bar{X}$ ’s
B_1	A_1B_1	A_2B_1	A_3B_1	$\bar{X}_{B1}$
B_2	A_1B_2	A_2B_2	A_3B_2	$\bar{X}_{B2}$
Factor A $\bar{X}$ ’s	$\bar{X}_{A1}$	$\bar{X}_{A2}$	$\bar{X}_{A3}$	

   c. Factor A, Level 1’s main effect is noted as $\bar{X}_{A1}$, while Factor B1’s **main effect** is noted as $\bar{X}_{B1}$. As you can see, the main effect in a factorial ANOVA would be produced using two, One Way ANOVA’s, one for each factor or independent variable.
   d. What is unique in factorial analysis of variance is the **interaction effect** of Factors A and B on the dependent variable: in other words, the effect of the two different levels of Factor B, might be dependent on which level of Factor A was being applied to the dependent variable along with the one of the two levels of Factor B. We would know there is no interaction effect if the scores on B_1 and B_2 are consistent across A_1, A_2, and A_3.
   e. A Factorial ANOVA conducts three (3) statistical tests.
      (1) Factor A main effect
      (2) Factor B main effect
      (3) The interaction effect of *AB*
      The object then is to determine whether or not main effects and an interaction effect are present in the study. We first start by determining whether or not there is an interaction between Factor A and Factor B.
   f. <u>Interactions and main effects are independent of each other; we may see main effects and no interaction or an interaction without main effects</u>.

4. A Study with no Interaction: Explanation of Table 12.9b
   a. Presented in Table 12.9b are scores on a hypothetical dependent variable for each for each Factor A and Factor B level paring.

Table 12.9b
2 x 3 Factorial Design

	Factor A			
Factor B	A_1	A_2	A_3	Factor B $\bar{X}$'s
B_1	A_1B_1=20	A_2B_1=40	A_3B_1=120	$\bar{X}_{B1}$=60
B_2	A_1B_2=100	A_2B_2=120	A_3B_2=200	$\bar{X}_{B2}$=140
Factor A $\bar{X}$'s	$\bar{X}_{A1}$=60	$\bar{X}_{A2}$=80	$\bar{X}_{A3}$=160	N = 100

b. Main Effects Analysis
   (1) The main effect of Factor B is found in the marginal means, $\bar{X}_{B1} = 60$ and $\bar{X}_{B2} = 140$, we could use One Way ANOVA to test the Null hypothesis H_0: $\mu_{B1} = \mu_{B2}$ to determine whether or not it should be rejected or retained.
   (2) The main effect of Factor A is found in $\bar{X}_{A1} = 60$, $\bar{X}_{A2} = 80$, and $\bar{X}_{A3} = 160$, we could use One Way ANOVA to test the Null hypothesis H_0: $\mu_{A1} = \mu_{A2} = \mu_{A3}$ to determine whether or not it should be rejected or retained.
   (3) In Factorial ANOVA, the Main Effect of one independent variable, ignores the effect of the second independent variable on the dependent variable.

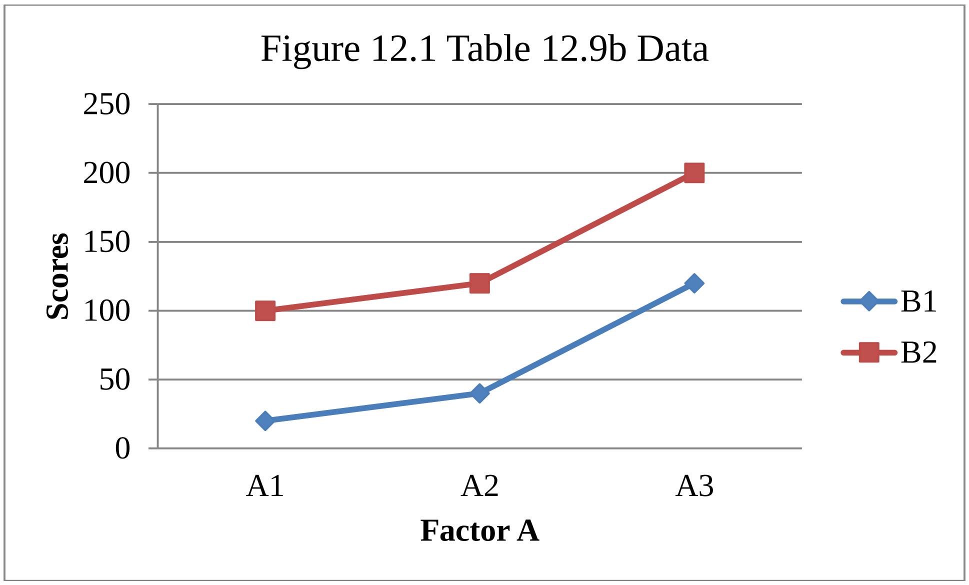
c. Interaction Analysis
   (1) Row Analysis: Examining Factor B Level B_1, there is a 20 point increase in the mean scores, moving from A_1B_1 to A_2,B_1 and a 80 point increase moving from A_2B_1 to A_3B_1. We see the same pattern across Factor B Level B_2. Changes in Factor A levels didn't change Factor B patterns; we can say changes in Factor B do not depend on changes in Factor A.
   (2) Column Analysis: Examining Column A_1, we see an 80 point increase moving from A_1B_1 to A_1B_2; we see a similar 80 increase moving from A_2B_1 to A_2B_2 and moving from A_3B_1 to A_3B_2. We can see that changes in the Factor A pattern were not affected by moving from B_1 to B_2 for any of the three columns with cell means.
   (3) Graphic Analysis: Using a line graph, we see that the Factor A scores and Factor B scores are parallel, suggesting no interaction. See Figure 12.1.

5. A Study with an Interaction: Explanation of Table 12.10
   a. Presented in Table 12.10 are hypothetical data where an interaction between Factor A and Factor B is present. We will conduct a Main Effects Analysis and then an Interaction Analysis. These analyses help a researcher to better understand his or her data and to determine if there are any possible main effects and whether or not there is or isn't a possible

interaction effect. None of these analyses are substitutes for statistical examination.

Table 12.10
2 x 3 Factorial Design

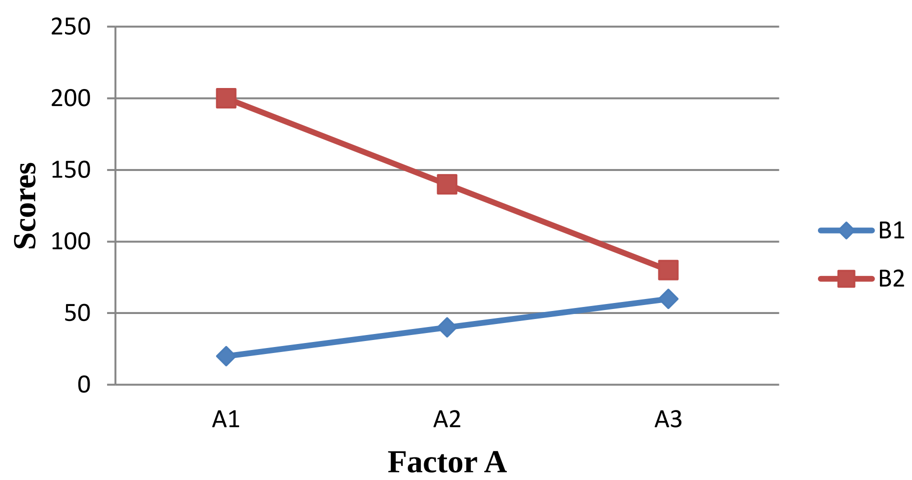
	Factor A			
Factor B	A_1	A_2	A_3	Factor B $\bar{X}$'s
B_1	A_1B_1=20	A_2B_1=40	A_3B_1=60	$\bar{X}_{B1}$=40
B_2	A_1B_2=200	A_2B_2=140	A_3B_2=80	$\bar{X}_{B2}$=140
Factor A $\bar{X}$'s	$\bar{X}_{A1}$=110	$\bar{X}_{A2}$=90	$\bar{X}_{A3}$=70	N = 90



b. Main Effects Analysis
   (1) The main effect of Factor B is found in the marginal means, $\bar{X}_{B1} = 40$ and $\bar{X}_{B2} = 140$, we could use One Way ANOVA to test the Null hypothesis H_0: $\mu_{B1} = \mu_{B2}$ to determine whether or not it should be rejected or retained. If the null is rejected, we have a main effect for Factor B.
   (2) The main effect of Factor A is found in $\bar{X}_{A1} = 110$, $\bar{X}_{A2} = 90$, and $\bar{X}_{A3} = 70$, we could use the One Way ANOVA to test the Null hypothesis H_0: $\mu_{A1} = \mu_{A2} = \mu_{A3}$ to determine whether or not it should be rejected or retained. If the null is rejected, we have a main effect for Factor A.
   (3) In Factorial ANOVA, the Main Effect of one independent variable, ignores the effect of the second independent variable on the dependent variable.

c. Interaction Analysis
(1) Row Analysis: Examining Factor B Level B_1, there is a 20 point increase in the mean scores, moving from A_1B_1 to A_2B_1, and a 20 point increase moving from A_2B_1 to A_3B_1. Inspecting Factor B Level B_2, there is a 60 point decrease in the mean scores, moving from A_1B_2 to A_2B_2, and a 60 point drop moving from A_2B_2 to A_3B_2. Changes in Factor A levels do change Factor B patterns; we can say changes in Factor B depend on changes in Factor A.
(2) Column Analysis: Examining Column A_1, we see a 180 point increase moving from A_1B_1 to A_1B_2; we see a 100 point increase moving from A_2B_1 to A_2B_2 and moving from A_3B_1 to A_3B_2, we see a 20 point increase. We can see that changes in the Factor A pattern are affected by moving from B_1 to B_2 for each of the three columns with cell means.
(3) Graphic Analysis:
(a) Using a line graph, we see that that the Factor A scores and Factor B scores are not parallel, suggesting a possible interaction. See Figure 12.2.
(b) In Figure 12.3, we see that the lines for Factor B Level B_1 and Level B_2 intersect, suggesting a possible interaction; Data in Figure 12.3 are unrelated to data in Tables 12.9 or 12.10 and are provided to show another possible pattern which may suggest an interaction.

Figure 12.2 Table 12.10 Data



B. Case: 2 x 2 Factorial Design

1. **Case 12.3**: A high school guidance counselor knew many of her students were going to attend the local university in the fall. The "summer break" was 18 weeks. She wondered about how many weeks they were able to work to save money for college. She knew from prior research that job experience and type of job affected the number of weeks a former student could work and thus, save money for college. At the end of the summer, she collected data on

weeks worked (dependent variable); prior job experience, before the summer job (Factor A); and type of industry. This is a 2 x 2 Factorial Design; see Tables 12.11a and 12.11b.

Table 12.11a
2 x 2 Factorial Design

	Factor A (Job Experience)	
Factor B (Job Type)	Little (Level A_1)	Some (Level A_2)
Service Work (Level B_1)	Weeks Worked	Weeks Worked
Manufacturing Work (Level B_2)	Weeks Worked	Weeks Worked

2. The case data are presented in Table 12.11b.

3. Partitioning Variance (Spatz, 2011, pp. 281-285)
   a. As in the One Way or One Factor ANOVA, we Partition variance.
   b. In Factorial Analysis of Variance, we partition variance into
      (1) Total Sum of Squares (SS_{Total} or SS_{Tot})
      (2) Cells Sum of Squares (SS_{Cells})
      (3) Factor A Sum of Squares ($SS_{Factor\ A}$)
      (4) Factor B Sum of Squares ($SS_{Factor\ B}$)
      (5) Interaction Sum of Squares (SS_{AB})
      (6) Error Sum of Squares (SS_{Error})
   c. Calculate Total Sum of Squares, using Formula 12.2a.

$$SS_{tot} = \sum X^2_{tot} - \frac{(\sum X_{tot})^2}{N_{tot}} = 1,946 - \frac{(172)^2}{16} = 1946 - \frac{29584}{16} = 1946 - 1849 = 97$$

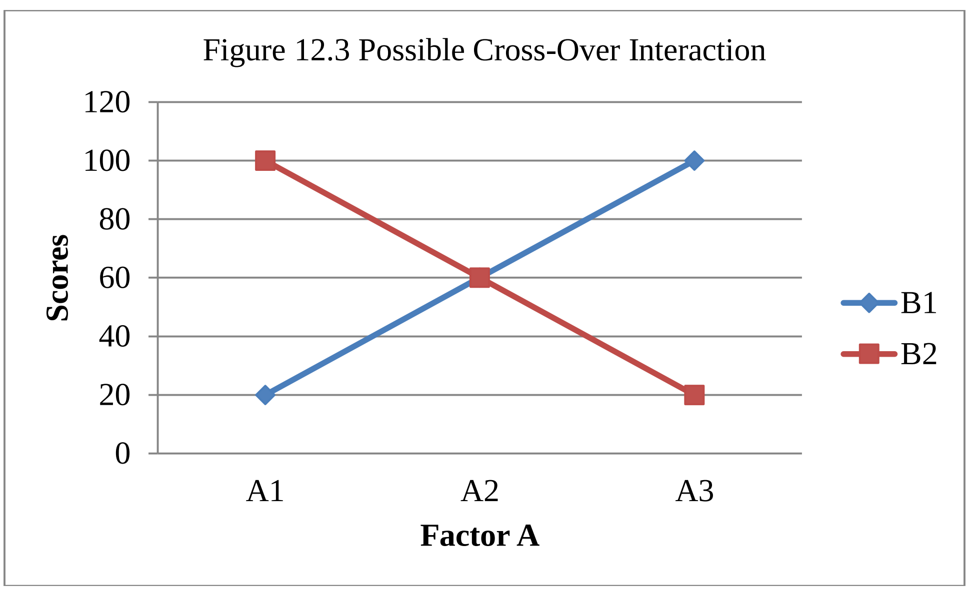


Table 12.11b
2 x 2 Factorial ANOVA Data Table

	Factor A (Job Experience)				
Factor B (Job Type)	**Level A_1** (Little) X	X^2	**Level A_2** (Some) X	X^2	**Factor B Main Effects**
Level B_1 (Service Work)	13	169	11	121	$\sum X_{B1} = 88$
	14	196	10	100	$\sum X^2_{B1} = 996$
	12	144	9	81	$\bar{X}_{B1} = 11$
	11	121	8	64	
	$\sum X_{cell} = 50$		$\sum X_{cell} = 38$		
	$\sum X^2_{cell} = 630$		$\sum X^2_{cell} = 366$		
	$\bar{X}_{cell} = 12.5$		$\bar{X}_{cell} = 9.5$		
Level B_2 (Manufacturing Work)	16	256	10	100	$\sum X_{B2} = 84$
	14	196	8	64	$\sum X^2_{B2} = 950$
	11	121	8	64	$\bar{X}_{B2} = 10.5$
	10	100	7	49	
	$\sum X_{cell} = 51$		$\sum X_{cell} = 33$		
	$\sum X^2_{cell} = 673$		$\sum X^2_{cell} = 277$		
	$\bar{X}_{cell} = 12.75$		$\bar{X}_{cell} = 8.25$		
Factor A Main Effects	$\sum X_{A1} = 101$		$\sum X_{A2} = 71$		**Total**
	$\sum X^2_{A1} = 1303$		$\sum X^2_{A2} = 643$		$\sum X_{Total} = 172$
	$\bar{X}_{A1} = 12.625$		$\bar{X}_{A2} = 8.875$		$\sum X^2_{Total} = 1,946$
					$\bar{X}_{Total} = 10.75$

d. Calculate Cells Sum of Squares, using Formula 12.8.

$$SS_{Cells} = \sum\left[\frac{\sum X_{Cells}}{N_{Cells}}\right] - \frac{(\sum X_{Total})^2}{N_{Total}} = \frac{(50)^2}{4} + \frac{(38)^2}{4} + \frac{(51)^2}{4} + \frac{(33)^2}{4} - \frac{(172)^2}{16} = 1908.5 - 1849 = 59.5$$

(Remember, once computed, the SS $_{Cells}$ is portioned into SS $_{Factor\ A}$, SS $_{Factor\ B}$, and SS $_{AB}$.)

e. Calculate Factor Main Effects Sum of Squares

(1) Compute Factor A Sum of Squares, using Formula 12.9a

$$SS_{FactorA} = \frac{(\sum X_{A1})^2}{N_{A1}} + \frac{(\sum X_{A2})^2}{N_{A2}} - \frac{(\sum X_{Total})^2}{N_{Total}} = \frac{(101)^2}{8} + \frac{(71)^2}{8} - \frac{(172)^2}{16} = 56.25$$

(2) Calculate Factor B Sum of Squares, using Formula 12.9b

$$SS_{FactorB} = \frac{(\sum X_{B1})^2}{N_{B1}} + \frac{(\sum X_{B2})^2}{N_{B2}} - \frac{(\sum X_{Total})^2}{N_{Total}} = \frac{(88)^2}{8} + \frac{(84)^2}{8} - \frac{(172)^2}{16} = 1.0$$

(3) Calculate Interaction Sum of Squares using Formula 12.10

$$SS_{AB} = SS_{Cells} - SS_{FactorA} - SS_{FactorB} = 59.5 - 56.25 - 1.0 = 2.25$$

f. Calculate Error Sum of Squares, using Formula 12.11.

$$SS_{Error} = \sum\left[\sum X^2_{Cell} - \frac{(\sum X_{Cell})^2}{N_{Cell}}\right] = \left(630 - \frac{(50)^2}{4}\right) + \left(366 - \frac{(38)^2}{4}\right) ... \left(277 - \frac{(51)^2}{4}\right) = 37.5$$

g. Check Calculations
(1) Check SS Cells Sum of Squares

$$SS_{Cells} = SS_{FactorA} + SS_{FactorB} + SS_{InteractionAB} = 56.25 + 1.0 + 2.25 = 59.5$$

(2) Check SS Total Sum of Squares

$$SS_{Total} = SS_{Cells} + SS_{Error} = 59.5 + 37.5 = 97$$

4. Establishing Degrees of Freedom (*df*) (Spatz, 2011, p. 287)
   a. $df_{\text{Total}} = N_{\text{Total}} - 1$ 16 - 1 = 15
   b. $df_{\text{Factor A}} = A - 1$ 2 – 1 = 1
   c. $df_{\text{Factor B}} = B - 1$ 2 – 1 = 1
   d. $df_{\text{AB}} = (A-1)(B-1)$ (1)(1) = 1
   e. $df_{\text{Error}} = N_{\text{Total}} - (A)(B)$ 16 - (2)(2) = 12

5. Construct the Factorial ANOVA Summary Table

Table 12.12
Factorial ANOVA Summary Table

Variance Source	**SS**	***df***	**MS**	**F**	**p**
Factor A	56.5	1	56.5	18.08	p < 0.05
Factor B	1.0	1	1.0	0.32	p > 0.05
AB	2.25	1	2.25	0.72	p > 0.05
Error	37.5	12	3.125		
Total	97	15			

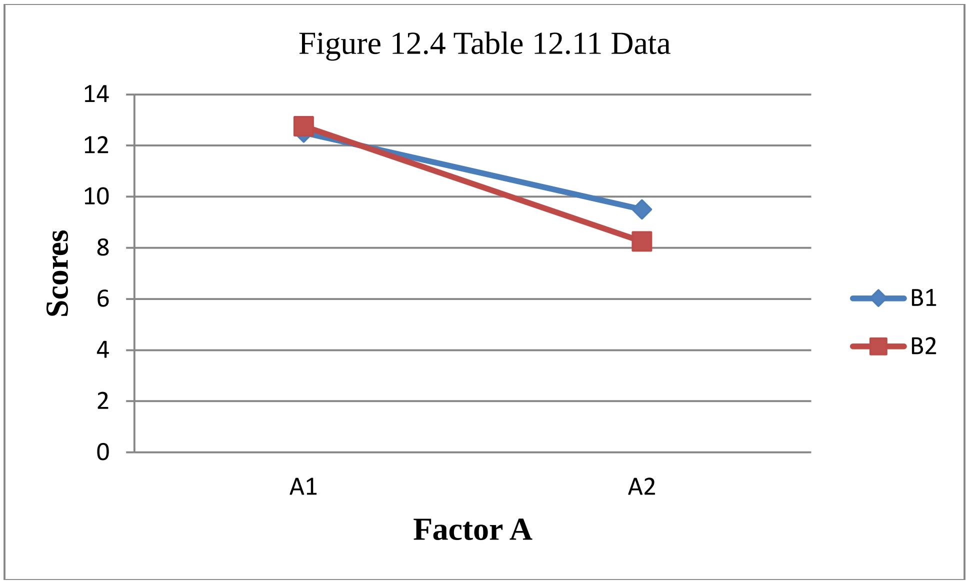
a. A <u>Mean Square (MS)</u> is computed by dividing the Variance Source by its associated degree or degrees of freedom. For example, the mean square (MS) for Factor A is 56.5/1=56.5 and the Mean Square for the Error Variance is 37.5/12=3.125.
b. The <u>F-Ratio</u> is computed by dividing Factor A, Factor B, and AB Mean Square variance by the Error Mean Square. For example, the Factor A F-Ratio is 56.5/3.125=18.08. The Factor B F-Ratio is 1.0/3.0125=.032. The AB F-Ratio is 2.25/3.125=0.72.
c. Testing the F Ratio for Statistical Significance at $\partial = 0.05$
(1) Factor A: $F_{(1, 8)} = 18.08 > 5.32$ (critical value), reject H_o: $\mu_1 = \mu_2$.
(2) Factor B: $F_{(1, 8)} = 0.32 < 5.32$ (critical value), retain H_o: $\mu_1 = \mu_2$.
(3) Interaction: $F_{(1, 8)} = 0.72 < 5.32$ (critical Value), retain the null.

Note: Any F-Ratio less than 1.0 is not going to be significant. For $F_{(1, 8)}$ the "1" is the degree of freedom (numerator) associated with a Specific Variance Source, e.g., Factor A. The "8" or dominator is the degrees of freedom of the Error Variance.

6. Interpreting the Factorial ANOVA Summary Table
   a. First, examine the interaction (AB)
      (1) The interaction (AB) which isn't significant, $F_{(1,8)} = 0.72$, $p > 0.05$.
      (2) We would write, "There is not a significant interaction between job experience and job type, $F_{(1,8)} = 0.72$, $p > 0.05$." In summarizing the interaction, we don't mention the dependent variable.
      (3) Because the interaction effect wasn't significant, we will interpret each main effect as if it was a One Way or One Factor ANOVA.
   b. Next, examine the Main Effect of Factor A
      (1) The main effect of Factor A was significant, $F_{(1, 8)} = 18.08$, $p < 0.05$.
      (2) We would write, "The main effect of job experience on weeks worked was significant, $F_{(1, 8)} = 18.08$, $p < 0.05$."
      (3) Because there was the main effect of Factor A was significant, its means must be graphed. By reviewing the marginal means ($A_1 = 12.625$ & $A_2 = 8.875$), we see that subjects in Level A_1 scored significantly higher than Level A_2. See Figure 12.4.

   c. Finally, examine the Main Effect of Factor B
      (1) The main effect of Factor B was not significant, $F_{(1, 8)} = 0.32$, $p > 0.05$.
      (2) F-Ratios less than 1.0 are not significant. So, no further analysis is needed.
   d. To summarize the information concerning Main Effects, we'd write, "The main effect of job experience was significant, $F_{(1, 8)} = 18.08$, $p < 0.05$ on weeks worked; but the main effect of job type was not, $F_{(1, 8)} = 0.32$, $p > 0.05$."

7. The plot of Factor B in Figure 12.4 might suggest an interaction. However, this is not the case. The A_1B_1 mean is 12.50 and the A_1B_2 mean is 12.75. The origins of the plot line are very close but don't actually touch. So the plot lines of Factor B Level B_1 and Factor B Level B_2 are actually fairly parallel; but, due to the scale of the line graph, otherwise is suggested. This should serve as a reminder to subject data to a full Factorial ANOVA. What we see on visual inspection may be inaccurate. Generally, the more parallel the lines between the factors (independent variables) are, the less likely there is an interaction.



C. Case: 3 x 2 Factorial Design

1. **Case 12.3**: Suppose undergraduate college students are faced with a "Rising Junior" basic skills competency test which must be passed before college sophomores can become juniors. The question could be, "Is there an effect on the earned test score (dependent variable) by a students' major (Factor A) and hours studied, expressed in 3 levels of hours studied (Factor B) for the examination?" Raw test scores are expressed in as standard scores, ranging for 40 to 75.

Table 12.13a
3 x 2 Factorial Design

	Factor A (Major)	
Factor B (Hours Studied)	Science	English
0-10 hours	Test Score	Test Score
11-19 hours	Test Score	Test Score
20-29 hours	Test Score	Test Score

2. The case data are presented in Table 12.13b.
3. Partitioning Variance (Spatz, 2011, pp. 281-285)
   a. As in the One Way or One Factor ANOVA, we Partition variance.
   b. In Factorial Analysis of Variance, we partition variance into
      (1) Total Sum of Squares (SS_{Total} or SS_{Tot})
      (2) Cells Sum of Squares (SS_{Cells})
      (3) Factor A Sum of Squares ($SS_{Factor\ A}$)
      (4) Factor B Sum of Squares ($SS_{Factor\ B}$)
      (5) Interaction Sum of Squares (SS_{AB})
      (6) Error Sum of Squares (SS_{Error})

Table 12.13b

3 x 2 Factorial ANOVA Data Table

	Factor A				
	Level A_1 (Science)		**Level A_2** (English)		**Factor B Main**
Factor B	X	X^2	X	X^2	**Effects**
Level B_1	45	2025	45	2025	$\sum X_{B1} = 370$
(0-10 hours)	46	2116	45	2025	$\sum X^2_{B1} = 17122$
	47	2209	47	2209	$\bar{X}_{B1} = 46.25$
	48	2304	47	2209	
	$\sum X_{cell} = 186$		$\sum X_{cell} = 184$		
	$\sum X^2_{cell} = 8654$		$\sum X^2_{cell} = 8468$		
	$\bar{X}_{cell} = 46.5$		$\bar{X}_{cell} = 46.0$		
Level B_2	57	3249	59	3481	$\sum X_{B2} = 481$
(11-19 hours)	58	3364	61	3721	$\sum X^2_{B2} = 28957$
	59	3481	62	3844	$\bar{X}_{B2} = 60.125$
	61	3721	64	4096	
	$\sum X_{cell} = 235$		$\sum X_{cell} = 246$		
	$\sum X^2_{cell} = 13815$		$\sum X^2_{cell} = 15142$		
	$\bar{X}_{cell} = 58.75$		$\bar{X}_{cell} = 61.5$		
Level B_3	67	4489	64	4096	$\sum X_{B3} = 552$
(20-29 hours)	70	4900	66	4356	$\sum X^2_{B3} = 38182$
	74	5476	67	4489	$\bar{X}_{B3} = 69.0$
	74	5476	70	4900	
	$\sum X_{cell} = 285$		$\sum X_{cell} = 267$		
	$\sum X^2_{cell} = 20341$		$\sum X^2_{cell} = 17841$		
	$\bar{X}_{cell} = 71.25$		$\bar{X}_{cell} = 66.75$		
Factor A	$\sum X_{A1} = 706$		$\sum X_{A2} = 697$		**Total**
Main Effects	$\sum X^2_{A1} = 42810$		$\sum X^2_{A2} = 41451$		$\sum X_{Total} = 1403$
	$\bar{X}_{A1} = 58.83$		$\bar{X}_{A2} = 58.08$		$\sum X^2_{Total} = 84261$
					$\bar{X}_{Total} = 58.46$

3. Partitioning Variance (Spatz, 2011, pp. 281-285)
    a. As in the One Way or One Factor ANOVA, we Partition variance.
    b. In Factorial Analysis of Variance, we partition variance into
        (1) Total Sum of Squares (SS_{Total} or SS_{Tot})
        (2) Cells Sum of Squares (SS_{Cells})
        (3) Factor A Sum of Squares ($SS_{Factor A}$)
        (4) Factor B Sum of Squares ($SS_{Factor B}$)
        (5) Interaction Sum of Squares (SS_{AB})
        (6) Error Sum of Squares (SS_{Error})

c. Calculate Total Sum of Squares, using Formula 12.2a.

$$SS_{tot} = \sum X_{tot}^2 - \frac{(\sum X_{tot})^2}{N_{tot}} = 84261 - \frac{(1403)^2}{24} = 84261 - \frac{1968409}{24} = 84261 - 82017 = 2244$$

d. Calculate Cells Sum of Squares, using Formula 12.8.

$$SS_{Cells} = \sum\left[\frac{\sum X_{Cells}}{N_{Cells}}\right] - \frac{(\sum X_{Total})^2}{N_{Total}} = \frac{(186)^2}{4} + \frac{(184)^2}{4} + \frac{(235)^2}{4} + \frac{(246)^2}{4} + \frac{(285)^2}{4} + \frac{(267)^2}{4} - \frac{(1403)^2}{16} =$$

84,176.75 - 82,017 = 2,159.75

(Remember, once computed, the SS $_{\text{Cells}}$ is portioned into SS $_{\text{Factor A}}$, SS $_{\text{Factor B}}$, and SS $_{\text{AB}}$.)

e. Calculate Factor Main Effects Sum of Squares

(1) Compute Factor A Sum of Squares, using Formula 12.9a

$$SS_{FactorA} = \frac{(\sum X_{A1})^2}{N_{A1}} + \frac{(\sum X_{A2})^2}{N_{A2}} - \frac{(\sum X_{Total})^2}{N_{Total}} = \frac{(706)^2}{12} + \frac{(697)^2}{12} - \frac{(1403)^2}{24} =$$

$$82,020.416 - 82,017 = 3.416$$

(2) Calculate Factor B Sum of Squares, using Formula 12.9b

$$SS_{FactorB} = \frac{(\sum X_{B1})^2}{N_{B1}} + \frac{(\sum X_{B2})^2}{N_{B2}} \frac{(\sum X_{B3})^2}{N_{B3}} - \frac{(\sum X_{Total})^2}{N_{Total}} =$$

$$\frac{(370)^2}{8} + \frac{(481)^2}{8} + \frac{(552)^2}{8} - \frac{(1403)^2}{24} = 84,120.625 - 82,017 = 2,103.625$$

(3) Calculate Interaction Sum of Squares using Formula 12.10

$$SS_{AB} = SS_{Cells} - SS_{FactorA} - SS_{FactorB} = 2,159.75 - 3.146 - 2,103.625 = 52.709$$

f. Calculate Error Sum of Squares, using Formula 12.11.

$$SS_{Error} = \sum\left[\sum X_{Cell}^2 - \frac{(\sum X_{Cell})^2}{N_{Cell}}\right] = \left(8654 - \frac{(186)^2}{4}\right) + \left(8468 - \frac{(184)^2}{4}\right) \ldots \left(17841 - \frac{(267)^2}{4}\right) = 84.25$$

g. Check Calculations

(1) Check SS Cells Sum of Squares

$$SS_{Cells} = SS_{FactorA} + SS_{FactorB} + SS_{InteractionAB} = 3.416 + 2,103.625 + 52.70 = 2,159.48$$

(2) Check SS Total Sum of Squares

$$SS_{Total} = SS_{Cells} + SS_{Error} = 2{,}159.48 + 84.25 = 2{,}244$$

4. Establishing Degrees of Freedom (*df*) (Spatz, 2011, p. 287)
   a. $df_{\text{Total}} = N_{\text{Total}} - 1$ — 24 - 1 = 23
   b. $df_{\text{Factor A}} = A - 1$ — 2 – 1 = 1
   c. $df_{\text{Factor B}} = B - 1$ — 3 – 1 = 2
   d. $df_{AB} = (A-1)(B-1)$ — (1)(2) = 2
   e. $df_{\text{Error}} = N_{\text{Total}} - (A)(B)$ — 24 - (3)(2) = 18

5. Construct the Factorial ANOVA Summary Table

Table 12.14
Factorial ANOVA Summary Table

Variance Source	**SS**	***df***	**MS**	**F**	**p**
Factor A	3.416	1	3.416	0.723	p > 0.05
Factor B	2,103.625	2	1051.813	224.698	p < 0.05
AB	52.709	2	26.355	5.630	p < 0.05
Error	84.25	18	4.681		
Total	2,244	23			

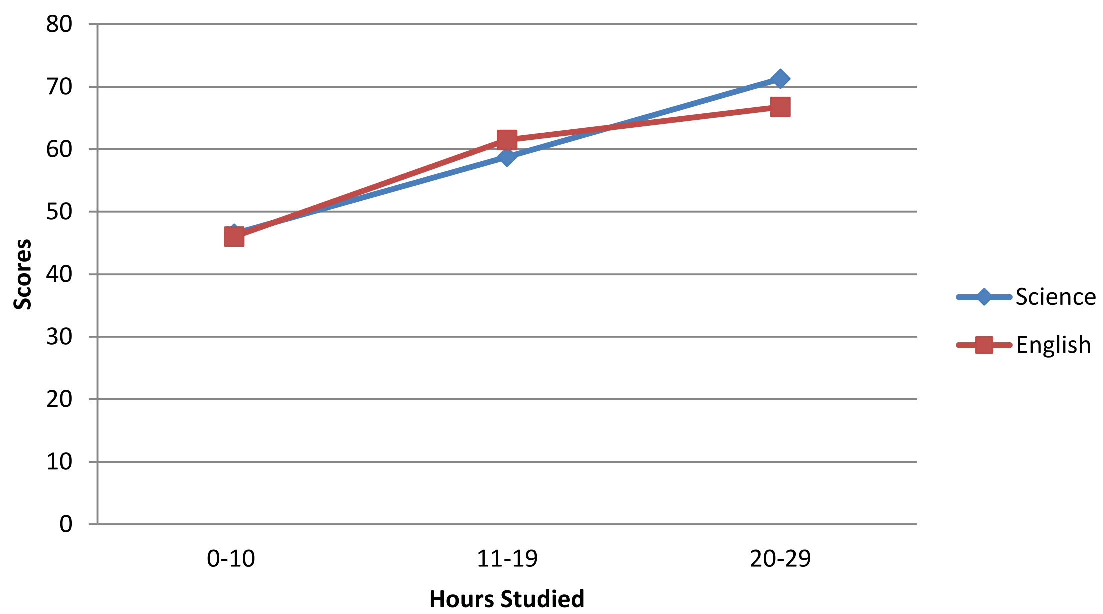
a. A <u>Mean Square (MS)</u> is computed by dividing the Variance Source by its associated degree or degrees of freedom. For example, the mean square (MS) for Factor A is 3.416/1=3.416 and the Mean Square for the Error Variance is 84.25/18=4.681.
b. The <u>F-Ratio</u> is computed by dividing Factor A, Factor B, and AB Mean Square variance by the Error Mean Square. For example, the Factor A F-Ratio is 3.416/4.681=0.723. The Factor B F-Ratio is 1051.813/4.681=224.698. The AB F-Ratio is 26.355/4.681=5.630.
c. Testing the F Ratio for Statistical Significance at $\partial = 0.05$
   (1) Factor A: $F_{(1, 18)} = 0.723 < 4.41$ (critical value), retain H_o: $\mu_1 = \mu_2$.
   (2) Factor B: $F_{(2, 18)} = 224.698 > 3.55$ (critical value), reject H_o: $\mu_1 = \mu_2 = \mu_3$.
   (3) Interaction: $F_{(2, 18)} = 5.630 > 3.55$ (critical Value), reject H_o: $\mu_1 = \mu_2$.

   Using the *GraphPad Software QuickCalcs* (n.d.) calculator, we see that the exact probability for Factor A is 0.4063 (retain H_0); for Factor B, 0.0001 (reject H_0) and Interaction AB, 0.0226 (reject H_0).

   Note: Any F-Ratio less than 1.0 is not going to be significant. For $F_{(1, 18)}$ the "1" is the degree of freedom (numerator) associated with a Specific Variance Source, e.g., Factor A. The "18" or dominator is the degrees of freedom of the Error Variance.

6. Interpreting the Factorial ANOVA Summary Table
   a. First, examine the interaction (AB)
      (1) The interaction (AB) is significant, $F_{(1, 18)} = 5.630$, $p < 0.05$.
      (2) We would write, "There is a significant interaction between major and hours studied, $F_{(1, 18)} = 5.630$, $p < 0.05$." This summary statement excludes the dependent variable.
      (3) By viewing Figure 12.5, we can see that for Science majors, as the hours studied mounted, earned test scores increased. For English majors, test scores peaked at between 11-19 hours of study and then declined with increasing hours of study. It could be that these English majors, as a group, over-studied.

Figure 12.5 Table 12.13 Data



   b. Next, examine the Main Effect of Factor A (Major)
      (1) The main effect for Factor A isn't significant, $F_{(1, 18)} = 0.723$, $p > 0.05$.
      (2) We would write, "The main effect of major on earned test score was not significant, $F_{(1, 18)} = 0.723$, $p > 0.05$.
      (3) By reviewing the marginal means ($A_1 = 58.83$ & $A_2 = 58.08$), we see that both majors produced highly similar scores.
   c. Finally, examine the Main Effect of Factor B
      (1) The main effect for Factor B is significant, $F_{(2, 18)} = 224.698$, $p < 0.05$.
      (2) We would write, "The main effect of hours studied on earned test score as significant, $F_{(2, 18)} = 224.698$, $p < 0.05$.
      (3) By reviewing the marginal means ($B_1 = 46.25$; $B_2 = 60.125$; and $B_3 = 69.0$), we see that hours studied greatly improved test scores.

d. To summarize the information on Main Effects, we'd write, "The main effect of major was not significant, $F_{(1, 18)} = 0.723$, $p > 0.05$ on test score earned; but the main effect of hours studied was, $F_{(2, 18)} = 224.698$, $p < 0.05$."

D. Testing Levels within a Factor for Statistical Significance & Effect Size

1. If an interaction isn't significant, we can compare levels within each factor to assess whether or not there are significant differences between factor levels, as we would in a One Way or One Factor ANOVA.

2. If we are working with a 2 x 2 Factorial Design, we don't need to use a post hoc test like the Tukey HSD, as the F-Ratio table gives us the information we need. Remember the 2 x 2 Factorial Design Case.
   a. For Factor A (job experience), we rejected H_o: $\mu_1 = \mu_2$, as $F_{(1, 8)} = 18.08$, p. < 0.05; there was a statistical difference between Factor A_1 (little job experience) and A_2 (some job experience) on weeks worked.
   b. For Factor B (job type), we retained H_o: $\mu_1 = \mu_2$, as $F_{(1, 8)} = 0.32$, $p > 0.05$, as there was no difference in job type (service work vs. manufacturing) on weeks worked.

3. If we are working with a 3 x 2 or higher Factorial Design, we can use the Tukey HSD post hoc test. We use Formulas 12.4 & 12.7. Since computations involving these formulas have been demonstrated using data from Table 12.7, no further computations will be presented here.

$$HSD = \frac{\overline{X}_i - \overline{X}_j}{s_{\bar{x}}}$$

Formula 12.4

$$s_{\bar{x}} = \sqrt{\frac{MS_{error}}{N_t}}$$

Formula 12.7
N_t = number in each treatment group (or a main effect margin N)

4. If the interaction isn't significant; but there are differences between pair-wise comparisons, use Formula 12.5 to compute effect size. This has been previously shown for data in Table 12.4.

E. Concluding Comments

1. The number of subjects in each cell (e.g., A_1B_1 A_1B_2, etc) must be equal; this is a balanced factorial design. For unequal cell sizes (cells with differing number of subjects in them), Investigate the "corrections" contained in a statistics package, like SPSS or SAS. Howell (2010) offers some guidance.
2. None of the factorial ANOVA formulas should be used with repeated measures designs. A subject must be included in only one data table cell. Randomly assign a subject to each cell (e.g., A_1B_1 A_1B_2, etc).
3. The Repeated Measures Factorial ANOVA can easily be applied to data using a statistical package like the Data Analysis add on in Excel (Salkind, 2010, pp. 268-280) or SPSS or SAS. The interpretation of the interaction effect and

factor main effects is the same approach as the Independent Groups Factorial ANOVA, studied here. Turner and Thayer (2001) wrote primer on analysis of variance, which the authors strongly recommend; it's still widely used and available.

a. Suppose a 2 x 2 Repeated Measures Factorial ANOVA was subjected to analysis. The design summary table would look like Table 12.15.
b. The evaluator is examining a reading skills program for at risk minority 4th grade students. A pre-test is given; students complete the program; and a posttest is given. The evaluator is also interested in assessing whether or not poverty status (indicated as being on or off "Free or Reduced Lunch" status) affects reading skill achievement.
c. Factor A is ethnicity; Factor B is poverty status; the dependent variable is reading skill, measured using a pre- and post-test. The analysis will yield an interaction assessment (ethnicity * poverty). The Factor A main effect will assess reading skills program effectiveness, based on ethnicity. Factor B will assess the same dependent variable, but based on poverty status.

Table 12.15
2 x 2 Factorial Design

	Factor A (Ethnicity)	
Factor B (Poverty)	Hispanic	Non-Hispanic
On Free/Reduced Lunch	Pre-test Score	Post-Test Score
No Free/Reduced Lunch	Pre-test Score	Post-Test Score

References

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

GraphPad Software QuickCalcs. (n.d.). Retrieved from http://graphpad.com/quickcalcs/Pvalue2.cfm

Howell, D. C. (2010). *Statistical methods for psychology* (7th ed.). Pacific Grove, CA: Cengage Wadsworth.

Salkind, N. J. (2010). *Statistics for people who (think they) hate statistics: EXCEL 2007 Edition* (2nd ed.). Thousand Oaks, CA: Sage.

Spatz, C. (2011). *Basic statistics: Tales of distributions* (10th ed.). Belmont, CA: Wadsworth.

Stevens, J. P. (1999) *Intermediate statistics: A modern approach* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Turner, J. R. & Thayer, J. F. (2001). *Introduction to analysis of variance*. Thousand Islands. CA: Sage.