**CHAPTER 14 SOLUTIONS**

**14.1.**

Graph (1)

1. There is a main effect due to gender. On average, males score higher in reading achievement (*M* = 18) than females (*M* = 12).
2. There is no main effect due to teaching method. On average, reading achievement for both males and females is the same under whole language (*M* = 15), synthetic phonics (*M* = 15), and analytic phonics (*M* = 15).
3. There is no interaction effect. The effectiveness of the different teaching methods does not differ by gender. Said differently, the difference between males and females in terms of their average reading achievement is the same (*MDifference* = 10) under all three teaching methods.
4. In the absence of an interaction, the statistically significant main effects are used to summarize results. In this case, we may conclude that, on average, males outperform females by the same amount for all three teaching methods.

Graph (2)

1. There is no main effect due to gender. On average, reading achievement for males (*M* = 20) is the same as for females (*M* = 20).
2. There is no main effect due to teaching method. On average, reading achievement under whole language (*M* = 20) is the same as under synthetic phonics (*M* = 20) and analytic phonics (*M* = 20).
3. There is a disordinal interaction. Holding teaching method constant and comparing males with females we may note that, on average, males score 20 points higher than females on reading achievement under whole language, no differently than females under synthetic phonics, and 20 points lower than females under analytic phonics. Alternatively, we may characterize the interaction by holding gender constant and comparing the relative effectiveness of teaching methods. From this perspective, we note that males do relatively best under whole language, next best under synthetic phonics, and relatively worst under analytic phonics. By contrast, females do relatively best under analytic phonics, next best under synthetic phonics, and relatively worst under whole language.
4. Given the statistically significant interaction, the relative effectiveness of the three methods by gender is captured by the response to part (c) above.

Graph (3)

1. There is a main effect due to gender. On average, reading achievement for males (*M* = 30) exceeds that for females (*M* = 16.67).
2. There is a main effect due to teaching method. On average, reading achievement is highest under analytic phonics (*M* = 30), followed by synthetic phonics (*M* = 25), and worst under whole language (*M* = 15).
3. There is an ordinal interaction. Holding teaching method constant and comparing males with females we may note that, on average, males score 10 points higher than females on reading achievement under both whole language and synthetic phonics, and 20 points higher than females under analytic phonics. Holding gender constant and comparing the relative effectiveness of teaching methods, we may note that, on average, males perform relatively best under analytic phonics (*M* = 40), next best under synthetic phonics (*M* = 30), and worst under whole language (*M* = 20). We may also note that, on average, females perform equally under the two types of phonics instruction (*M* = 20), but relatively worse under whole language (*M* = 10).
4. Given the statistically significant interaction, the relative effectiveness of the three methods by gender is captured by the response to part (c) above.

Graph (4)

1. There is a main effect due to gender. On average, males (*M* = 35) score higher than females (*M* = 15) on reading achievement.
2. There is no main effect due to teaching method. On average, there is no difference on reading achievement (*M* = 25) across the three teaching methods.
3. There is an ordinal interaction in that, on average, males perform better than females under each teaching method by varying amounts of reading achievement. In particular, on average, males score 10 points higher than females under whole language, 20 points higher under synthetic phonics, and 30 points higher under analytic phonics. We may also characterize the interaction by comparing teaching methods for each gender. For males, the highest average reading achievement scores are associated with analytic phonics (*M* = 40), followed by synthetic phonics (*M* = 35), followed by whole language (*M* = 30). For females, the highest average reading achievement scores are associated with whole language (*M* = 20), followed by synthetic phonics (*M* = 15), followed by analytic phonics (*M* = 10).
4. Given the statistically significant interaction, the relative effectiveness of the three methods by gender is captured by the response to part (c) above.

Graph (5)

1. There is a main effect due to gender. On average, the reading achievement score for males (*M* = 23.33) is higher than it is for females (*M* = 13.33) across all teaching methods.
2. There is a main effect due to teaching method. On average, males and females both do better under whole language and analytic phonics (*M* = 25) than they do under synthetic phonics (*M* = 15).
3. There is no interaction effect. On average, males outperform females by the same amount under all three teaching methods. That is, the mean difference on reading achievement between males and females is the same (*MDifference* = 10) under all three teaching methods.
4. In the absence of an interaction, as given in the responses to parts (a) and (b) above, the main effects characterize the results of this study. s are used to summarize the effects.

**14.2.**

1. For the main effect of SEX: H0: male = female and H1: male ≠ female.

For the main effect of CURSMOKE1: H0: smoker = non-smoker and H1: smoker ≠ non-smoker.

For the interaction effect: H0: There is no interaction in the population between cigarette use and gender on body mass index. H1: There is an interaction in the population between cigarette use and gender on body mass index.

1. The R code for generating the line graph is:

**interaction.plot(Framingham$CURSMOKE1,Framingham$SEX, Framingham$BMI1, xlab = "Smokes", ylab = "BMI", trace.label = "Sex",**

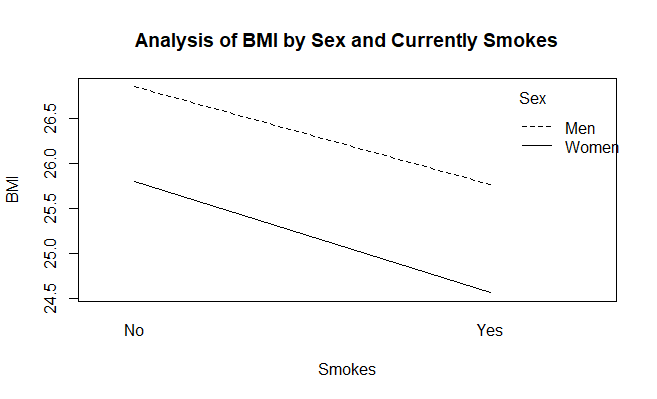
**main = "Analysis of BMI by Sex and Currently Smokes" )**

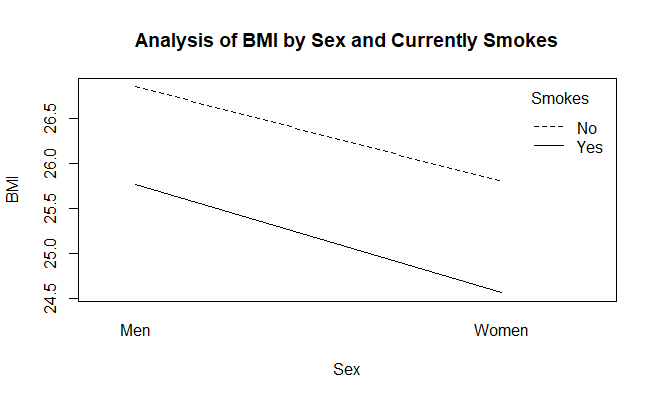
OR

**interaction.plot(Framingham$SEX,Framingham$CURSMOKE1, Framingham$BMI1, xlab = "Sex",**

**ylab = "BMI", trace.label = "Smokes", main = "Analysis of BMI by Sex and Currently Smokes" )**

Because the lines are approximately parallel, there does not appear to be an interaction between gender and cigarette use on BMI. Because the line for males is higher than that for females, on average, males have a higher BMI than females across both levels of CURSMOKE1, suggesting a main effect due to gender. Because both lines slope downward, and a higher BMI is associated with non-smokers, there appears also to be a main effect due to cigarette use.





c) The ANOVA is robust to possible violations of the normality assumption when each cell contains a large number of research participants, which is the case for these data since *n* = 100. Hence, the normality assumption is not an issue for these data.

d) The ANOVA is robust to violations of the homogeneity of variance assumption when cell sizes are equal and large, which is the case for these data. Hence, the homogeneity of variance assumption is not an issue for these data.

e) According to the ANOVA results, the interaction effect is not statistically significant, *F*(1, 396) = 0.04, *p* = .85. However, the main effect due to cigarette use is statistically significant, *F*(1, 396) = 9.20, *p* = .003, as is the main effect due to gender, *F*(1, 396) = 8.59, *p* = .004.

> aov1 <- aov(Framingham$BMI1~Framingham$SEX\*Framingham$CURSMOKE1)

> summary(aov1)

Df Sum Sq Mean Sq F value Pr(>F)

Framingham$SEX 1 127 126.75 8.591 0.00357 \*\*

Framingham$CURSMOKE1 1 136 135.66 9.195 0.00259 \*\*

Framingham$SEX:Framingham$CURSMOKE1 1 1 0.55 0.038 0.84641

Residuals 396 5842 14.75

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

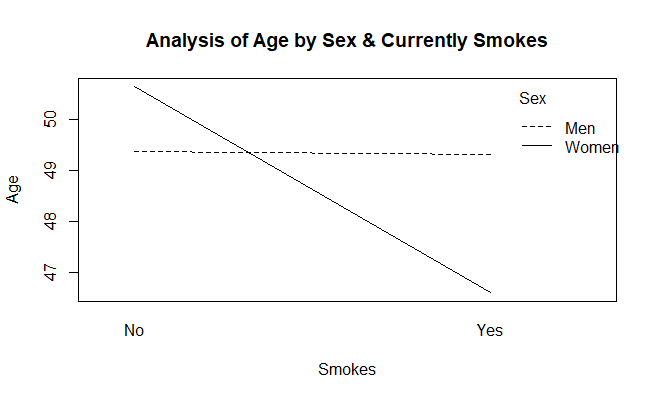
f) Because each statistically significant main effect only has two levels, *post-hoc* testing is not necessary. The sample means themselves are sufficient to indicate the direction of the difference.

On average, males (*M* = 26.31, *SD* = 3.48) have a statistically significantly higher BMI than females (*M* = 25.18, *SD* = 4.24); and, on average, non-smokers (*M* = 26.32, *SD* = 3.80) have a statistically significantly higher BMI than smokers (*M* = 25.16, *SD* = 3.95).

g) According to the value of *R2*, approximately 2.23 percent ((136/ (127+136+1+5842)) x 100) of the variance in initial BMI is explained by cigarette use, and approximately 2.08 percent (127/ (127+136+1+5842)) x 100) of the variance in initial BMI is explained by gender. Collectively, both statistically significant main effects account for approximately 4.31 percent of initial BMI variance.

We can produce two graphs, exchanging the x variable and trace variable.

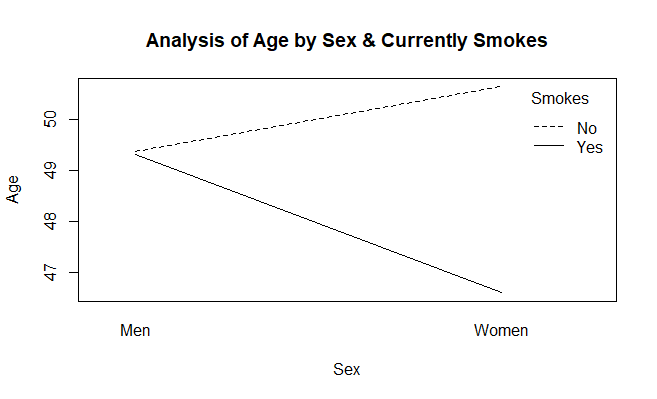
**interaction.plot(Framingham$CURSMOKE1,Framingham$SEX, Framingham$AGE1, xlab = "Smokes",ylab = "Age", trace.label = "Sex", main = "Analysis of Age by Sex & Currently Smokes")**



**interaction.plot(Framingham$SEX,Framingham$CURSMOKE1, Framingham$AGE1, xlab = "Sex",**

**ylab = "Age", trace.label = "Smokes", main = "Analysis of Age by Sex &**

**Currently Smokes" )**



Interpretation: Because the lines cross, there appears to be a disordinal interaction between gender and cigarette use on age. With respect to main effects, on average, females appear to be younger than males and, on average, smokers appear to be younger than non-smokers.

1. According to the ANOVA results, there is a statistically significant interaction, *F*(1, 396) = 5.71, *p* = .02, a statistically significant main effect due to cigarette use, *F*(1, 396) = 6.12, *p* = .01, but no statistically significant main effect due to gender, *F*(1, 396) = 0.77, *p* = .38.

> aov2 <- aov(Framingham$AGE1~Framingham$SEX\*Framingham$CURSMOKE1)

> summary(aov2)

Df Sum Sq Mean Sq F value Pr(>F)

Framingham$SEX 1 53 53.3 0.769 0.3811

Framingham$CURSMOKE1 1 424 424.4 6.122 0.0138 \*

Framingham$SEX:Framingham$CURSMOKE1 1 396 396.0 5.713 0.0173 \*

Residuals 396 27450 69.3

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

1. The R code for conducting the post hoc tests for the interaction and the related output follow. We use the ANOVA object aov2 that we created in this example.

**testInteractions(aov2, fixed = "Framingham$CURSMOKE1", pairwise = "Framingham$SEX", adjustment = "bonferroni")**

F Test:

P-value adjustment method: bonferroni

Value Df Sum of Sq F Pr(>F)

Men-Women : No -1.26 1 79.4 1.1451 0.57044

Men-Women : Yes 2.72 1 369.9 5.3365 0.04279 \*

Residuals 396 27450.3

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

**testInteractions(aov2, fixed = "Framingham$SEX", pairwise = "Framingham$CURSMOKE1", adjustment = "bonferroni")**

F Test:

P-value adjustment method: bonferroni

Value Df Sum of Sq F Pr(>F)

No-Yes : Men 0.07 1 0.2 0.0035 1.000000

No-Yes : Women 4.05 1 820.1 11.8312 0.001289 \*\*

Residuals 396 27450.3

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Looking at the Bonferroni-adjusted *p*-values associated with the *F*-tests, we can determine which groups differ from each other. Holding cigarette use constant: For non-smokers, on average, males and females are not statistically significantly different in age. For smokers, however, males are statistically significantly older than females.

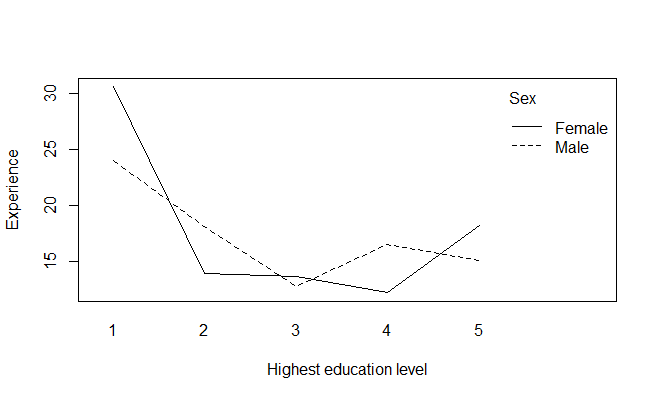
Holding gender constant: For males, smokers and non-smokers are not statistically significantly different in age. For females, however, smokers are statistically significantly younger than non-smokers.

Given that cigarette use and gender each has only two levels, additional *post-hoc* tests are not necessary.

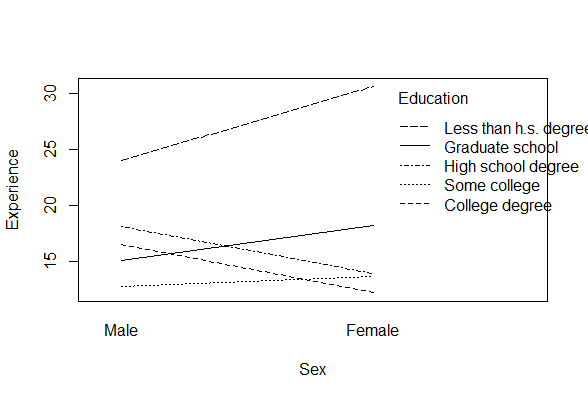
d) According to the respective values of *R2*, approximately 1.4 percent ((396/( 53+424+396+27450)) x 100) of age variance is explained by the interaction of gender and cigarette use, and approximately 1.5 percent ((424/( 53+424+396+27450)) x 100) of age variance is explained by cigarette use. Collectively, both statistically significant effects account for approximately 2.9 percent of age variance.

1. The R code for generating the ANOVA analysis and the line graph (using one of two possible ways given) are:

**interaction.plot(as.numeric(Wages$ed), Wages$sex, Wages$exper, xlab = "Highest education level", ylab = "Experience", trace.label = "Sex")**



**interaction.plot(Wages$sex, Wages$ed, Wages$exper, xlab = "Sex", ylab = "Experience", trace.label = "Education")**



Because the lines cross several times, this is a disordinal interaction between education level and gender on years of work experience. In particular, on average, the number of years of work experience appears to be greater for females than males when both have either less than a high school degree, some college, or a graduate degree; and, greater for males than females when both have either a high school degree or a college degree. With respect to main effects, on average, the number of years of work experience for females appears to be approximately equal to that for males; and, on average, the number of years of work experience appears to be highest for those with less than a high school degree, next highest for those with a graduate school degree, and so on.

1. By running the R code below, we obtain a table of cell means, a table of standard deviations, and table of cell sizes. The ANOVA is robust to possible violations of the normality assumption when each cell contains a large number of research participants, which is the case for these data with *n* = 40. Hence, the normality assumption is not an issue for these data. We reproduce the tables here as they may be used to determine the direction of the mean differences detected in part e).

**aggregate(Wages$exper ~ Wages$sex + Wages$ed, FUN = mean)**

Wages$sex Wages$ed Wages$exper

1 Male Less than h.s. degree 24.00

2 Female Less than h.s. degree 30.60

3 Male High school degree 18.15

4 Female High school degree 13.90

5 Male Some college 12.80

6 Female Some college 13.65

7 Male College degree 16.55

8 Female College degree 12.20

9 Male Graduate school 15.10

10 Female Graduate school 18.20

**aggregate(Wages$exper ~ Wages$sex + Wages$ed, FUN = sd)**

Wages$sex Wages$ed Wages$exper

1 Male Less than h.s. degree 12.253728

2 Female Less than h.s. degree 11.405801

3 Male High school degree 13.916509

4 Female High school degree 9.896386

5 Male Some college 8.348806

6 Female Some college 9.333013

7 Male College degree 8.761659

8 Female College degree 8.274769

9 Male Graduate school 10.357210

10 Female Graduate school 10.754129

**table(Wages$sex, Wages$ed)**

Less than h.s. degree High school degree Some college College degree Graduate school

Male 40 40 40 40 40

Female 40 40 40 40 40

c) The ANOVA is robust to violations of the homogeneity of variance assumption when cell sizes are equal and large, which is the case for these data. Hence, the homogeneity of variance assumption is not an issue for these data.

d) According to the ANOVA results there is a statistically significant interaction, *F*(4, 390) = 4.11, *p* = .003, a statistically significant main effect due to education level, *F*(4, 390) = 23.15, *p* < .0005, but not a statistically significant main effect due to gender, *F*(1, 390) = 0.14, *p* = .71.

> aov3 <- aov(Wages$exper~Wages$sex\*Wages$ed)

> summary(aov3)

Df Sum Sq Mean Sq F value Pr(>F)

Wages$sex 1 15 15.2 0.139 0.70981

Wages$ed 4 10158 2539.6 23.154 < 2e-16 \*\*\*

Wages$sex:Wages$ed 4 1802 450.6 4.108 0.00284 \*\*

Residuals 390 42776 109.7

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

1. To conduct the two tests of simple effects due to education level, we use the **testInteractions** function from the **phia** package. We will not install or load the phia package into working memory as we have done that earlier in the exercise.

We start by fixing sex and comparing across ed levels. We first check for the simple effects and then if they are significant, move on to the pairwise comparisons using the Bonferroni adjustment. Based on the following table, we see that for both males and females, there are statistically significant mean differences in experience between the education levels.

> testInteractions(aov3, fixed = "Wages$sex", across="Wages$ed", adjustment = "none")

F Test:

P-value adjustment method: none

Wages$ed1 Wages$ed2 Wages$ed3 Wages$ed4 Df Sum of Sq F Pr(>F)

Male 8.9 3.05 -2.30 1.45 4 2851 6.4972 4.536e-05 \*\*\*

Female 12.4 -4.30 -4.55 -6.00 4 9110 20.7647 1.571e-15 \*\*\*

Residuals 390 42776

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Based on the following table, holding gender constant: For males, those with less than a high school degree have statistically significantly more years of work experience, on average, than those with some college or a college or graduate degree. For females, those with less than a high school degree, on average, have statistically significantly more years of work experience than those at other education levels.

> testInteractions(aov3, fixed = "Wages$sex", pairwise="Wages$ed", adjustment = "bonferroni")

F Test:

P-value adjustment method: bonferroni

Value Df Sum of Sq F Pr(>F)

Less than h.s. degree-High school degree : Male 5.85 1 684 6.2403 0.25797

Less than h.s. degree-Some college : Male 11.20 1 2509 22.8733 4.915e-05 \*\*\*

Less than h.s. degree-College degree : Male 7.45 1 1110 10.1206 0.03167 \*

Less than h.s. degree-Graduate school : Male 8.90 1 1584 14.4435 0.00335 \*\*

High school degree-Some college : Male 5.35 1 572 5.2192 0.45752

High school degree-College degree : Male 1.60 1 51 0.4668 1.00000

High school degree-Graduate school : Male 3.05 1 186 1.6963 1.00000

Some college-College degree : Male -3.75 1 281 2.5642 1.00000

Some college-Graduate school : Male -2.30 1 106 0.9646 1.00000

College degree-Graduate school : Male 1.45 1 42 0.3834 1.00000

Less than h.s. degree-High school degree : Female 16.70 1 5578 50.8541 9.730e-11 \*\*\*

Less than h.s. degree-Some college : Female 16.95 1 5746 52.3881 4.881e-11 \*\*\*

Less than h.s. degree-College degree : Female 18.40 1 6771 61.7347 7.726e-13 \*\*\*

Less than h.s. degree-Graduate school : Female 12.40 1 3075 28.0373 3.984e-06 \*\*\*

High school degree-Some college : Female 0.25 1 1 0.0114 1.00000

High school degree-College degree : Female 1.70 1 58 0.5270 1.00000

High school degree-Graduate school : Female -4.30 1 370 3.3716 1.00000

Some college-College degree : Female 1.45 1 42 0.3834 1.00000

Some college-Graduate school : Female -4.55 1 414 3.7750 1.00000

College degree-Graduate school : Female -6.00 1 720 6.5644 0.21557

Residuals 390 42776

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Based on the following table, holding education level constant: For those with less than a high school degree, males have statistically significantly less years of experience, on average, than females. No other statistically significant differences are noted between males and females at any of the other education levels.

> testInteractions(aov3, fixed = "Wages$ed", across ="Wages$sex", adjustment = "none")

F Test:

P-value adjustment method: none

Value Df Sum of Sq F Pr(>F)

Less than h.s. degree -6.60 1 871 7.9429 0.005073 \*\*

High school degree 4.25 1 361 3.2936 0.070318 .

Some college -0.85 1 14 0.1317 0.716827

College degree 4.35 1 378 3.4504 0.063989 .

Graduate school -3.10 1 192 1.7523 0.186359

Residuals 390 42776

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

f) According to the respective values of *R2*, approximately 3.29 percent ((1802/(1802+10158+15+42776))\*100)of work experience variance is explained by the interaction of gender and education level, and approximately 18.55 percent ((10158/(1802+10158+15+42776))\*100) ((10158.26/54751.91) x 100) of work experience variance is explained by education level. Collectively, both statistically significant effects account for approximately 21.84 percent of work variance.

* 1. We create a dataset called newdata by entering data for the three variables with the **data.frame** function.

newdata = data.frame(Time=c("morning","morning","morning","morning",

"morning","morning","morning","morning","morning","morning","morning",

"morning","morning","morning","morning","morning","morning","morning",

"morning","morning","afternoon","afternoon","afternoon","afternoon",

"afternoon","afternoon","afternoon","afternoon","afternoon","afternoon",

"afternoon","afternoon","afternoon","afternoon","afternoon","afternoon",

"afternoon","afternoon","afternoon","afternoon"),

Year=c("freshman","freshman","freshman","freshman",

"freshman","sophomore","sophomore","sophomore","sophomore","sophomore",

"junior","junior","junior","junior","junior","senior","senior","senior",

"senior","senior","freshman","freshman","freshman","freshman","freshman",

"sophomore","sophomore","sophomore","sophomore","sophomore","junior",

"junior","junior","junior","junior","senior","senior","senior","senior",

"senior"),

Score=c(80,80,75,70,70,85,80,80,83,82,93,90,89,87,87,100,98,95,93,90,70,70,65,60,60,75,71,70,69,65,85,84,80,73,72,88,83,80,79,75))

1. We use the **aggregate** function to produce cell and marginal means and standard deviations.

> # Cell mean and standard deviations

> aggregate(newdata$Score ~ newdata$Time + newdata$Year, FUN=mean)

newdata$Time newdata$Year newdata$Score

1 afternoon freshman 65.0

2 morning freshman 75.0

3 afternoon junior 78.8

4 morning junior 89.2

5 afternoon senior 81.0

6 morning senior 95.2

7 afternoon sophomore 70.0

8 morning sophomore 82.0

> aggregate(newdata$Score ~ newdata$Time + newdata$Year, FUN=sd)

newdata$Time newdata$Year newdata$Score

1 afternoon freshman 5.000000

2 morning freshman 5.000000

3 afternoon junior 6.058052

4 morning junior 2.489980

5 afternoon senior 4.847680

6 morning senior 3.962323

7 afternoon sophomore 3.605551

8 morning sophomore 2.121320

>

> # Marginal means and standard deviations for Time

> aggregate(newdata$Score ~ newdata$Time, FUN=mean)

newdata$Time newdata$Score

1 afternoon 73.70

2 morning 85.35

> aggregate(newdata$Score ~ newdata$Time, FUN=sd)

newdata$Time newdata$Score

1 afternoon 8.066174

2 morning 8.449696

>

> # Marginal means and standard deviations for Year

> aggregate(newdata$Score ~ newdata$Year, FUN=mean)

newdata$Year newdata$Score

1 freshman 70.0

2 junior 84.0

3 senior 88.1

4 sophomore 76.0

> aggregate(newdata$Score ~ newdata$Year, FUN=sd)

newdata$Year newdata$Score

1 freshman 7.071068

2 junior 7.007932

3 senior 8.569325

4 sophomore 6.912147

1. Evaluate the tenability of assumptions using the following R code:

The skewness statistics were calculated using the **skew**, **se.skew** and **skew.ratio** functions, output below.

> skew(newdata$Score[newdata$Year=="freshman" & newdata$Time=="morning"])

[1] 0

> se.skew(newdata$Score[newdata$Year=="freshman" & newdata$Time=="morning"])

[1] 0.9128709

> skew.ratio(newdata$Score[newdata$Year=="freshman" & newdata$Time=="morning"])

[1] 0

>

> skew(newdata$Score[newdata$Year=="sophomore" & newdata$Time=="morning"])

[1] 0.3513642

> se.skew(newdata$Score[newdata$Year=="sophomore" & newdata$Time=="morning"])

[1] 0.9128709

> skew.ratio(newdata$Score[newdata$Year=="sophomore" & newdata$Time=="morning"])

[1] 0.3849002

>

> skew(newdata$Score[newdata$Year=="junior" & newdata$Time=="morning"])

[1] 0.6170311

> se.skew(newdata$Score[newdata$Year=="junior" & newdata$Time=="morning"])

[1] 0.9128709

> skew.ratio(newdata$Score[newdata$Year=="junior" & newdata$Time=="morning"])

[1] 0.6759237

>

> skew(newdata$Score[newdata$Year=="senior" & newdata$Time=="morning"])

[1] -0.08411073

> se.skew(newdata$Score[newdata$Year=="senior" & newdata$Time=="morning"])

[1] 0.9128709

> skew.ratio(newdata$Score[newdata$Year=="senior" & newdata$Time=="morning"])

[1] -0.09213868

>

> skew(newdata$Score[newdata$Year=="freshman" & newdata$Time=="afternoon"])

[1] 0

> se.skew(newdata$Score[newdata$Year=="freshman" & newdata$Time=="afternoon"])

[1] 0.9128709

> skew.ratio(newdata$Score[newdata$Year=="freshman" & newdata$Time=="afternoon"])

[1] 0

>

> skew(newdata$Score[newdata$Year=="sophomore" & newdata$Time=="afternoon"])

[1] 0

> se.skew(newdata$Score[newdata$Year=="sophomore" & newdata$Time=="afternoon"])

[1] 0.9128709

> skew.ratio(newdata$Score[newdata$Year=="sophomore" & newdata$Time=="afternoon"])

[1] 0

>

> skew(newdata$Score[newdata$Year=="junior" & newdata$Time=="afternoon"])

[1] -0.1620248

> se.skew(newdata$Score[newdata$Year=="junior" & newdata$Time=="afternoon"])

[1] 0.9128709

> skew.ratio(newdata$Score[newdata$Year=="junior" & newdata$Time=="afternoon"])

[1] -0.1774893

>

> skew(newdata$Score[newdata$Year=="senior" & newdata$Time=="afternoon"])

[1] 0.3091461

> se.skew(newdata$Score[newdata$Year=="senior" & newdata$Time=="afternoon"])

[1] 0.9128709

> skew.ratio(newdata$Score[newdata$Year=="senior" & newdata$Time=="afternoon"])

[1] 0.3386526

Because there are only five students per cell we check the tenability of the normality assumption by computing the skewness ratio for each cell. As the skewness ratio for each cell is less than 2 in absolute value, the data appear to be reasonably symmetric and the normality assumption may be considered to be tenable.

According to the results of Levene’s test, shown below, the homogeneity of variance assumption is tenable, *F*(7, 32) = 1.60, *p* = .17.

> levenes.test(newdata$Score, newdata$Year:newdata$Time)

Levene's Test for Homogeneity of Variance

Df F value Pr(>F)

group 7 1.5952 0.1726

32

1. There is no statistically significant interaction effect, *F*(3, 32) = 0.49, *p* = .69. There is a statistically significant main effect due to time, *F*(1, 32) = 72.58, *p* < .0005 as well as a statistically significant main effect due to academic year, *F*(3, 32) = 35.06, *p* < .0005.

> aov4 <- aov(newdata$Score~newdata$Time\*newdata$Year)

> summary(aov4)

Df Sum Sq Mean Sq F value Pr(>F)

newdata$Time 1 1357.2 1357.2 72.579 9.80e-10 \*\*\*

newdata$Year 3 1967.1 655.7 35.064 3.13e-10 \*\*\*

newdata$Time:newdata$Year 3 27.3 9.1 0.486 0.694

Residuals 32 598.4 18.7

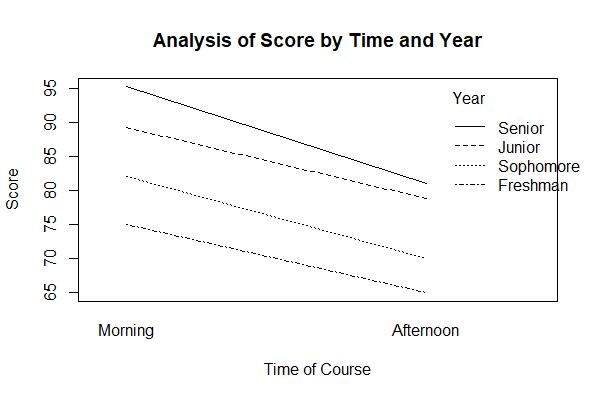
---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

1. According to the value of *R2*, approximately (1357.2/(1357.2+1967.1+27.3+598.4)\*100) = 34.4 percent of the final exam variance is explained by time, and (1967.1/(1357.2+1967.1+27.3+598.4)\*100) = 49.8 percent by year.
2. According to the line graph, the nearly parallel line segments suggest the absence of an interaction. The difference in the heights of the two lines by academic year suggest a main effect due to time; and in particular, that students who are taking the course in the morning, on average, perform better on the final exam than those who are taking the course in the afternoon. There is a main effect due to academic year, as depicted by the positive slopes of the line segments depicting morning and afternoon times, suggest a main effect for academic year; and, in particular, that performance on the final exam increases with academic year.

The following line graph is created using the following R code:

**interaction.plot(newdata$Time, newdata$Year, newdata$Score, xlab = "Time of Course", ylab = "Score", trace.label = "Year", main = "Analysis of Score by Time and Year")**

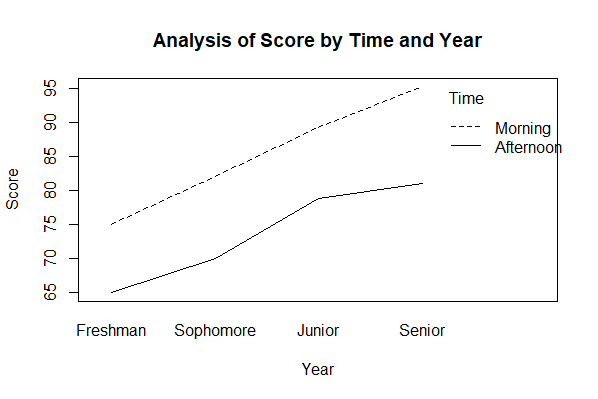


The following code would place Year on the x-axis instead.

**interaction.plot(newdata$Year, newdata$Time, newdata$Score, xlab = "Year",**

**ylab = "Score", trace.label = "Time", main = "Analysis of Score by**

**Time and Year")**



1. Because time has only two levels (morning and afternoon), the sample means themselves are sufficient to indicate the nature of this main effect. We note that students taking the course in the morning, on average, perform better on the final exam than those taking the course in the afternoon.
2. Because academic year has more than two levels, we carry out a *post-hoc* test to understand the nature of the main effect due to academic year. According to the results of the Tukey HSD *post-hoc* test, on average, juniors and seniors perform best on the final exam followed by sophomores, with freshmen performing statistically significantly worse than all other groups, on average. No statistically significant difference in performance is observed between juniors and the seniors, on average.

The R code for performing the relevant post hoc test of the main effect is:

**install.packages("DescTools")**

**library(DescTools)**

**PostHocTest(aov4, method = "hsd")$`newdata$Year`**

diff lwr.ci upr.ci pval

Sophomore-Freshman 6.0 0.7603439 11.239656 1.980570e-02

Junior-Freshman 14.0 8.7603439 19.239656 1.863262e-07

Senior-Freshman 18.1 12.8603439 23.339656 6.608534e-10

Junior-Sophomore 8.0 2.7603439 13.239656 1.299252e-03

Senior-Sophomore 12.1 6.8603439 17.339656 3.001973e-06

Senior-Junior 4.1 -1.1396561 9.339656 1.684836e-01

1. For normality: Because there are only five students per cell we check the tenability of the normality assumption by computing the skewness ratio for each cell. As the skewness ratio for each cell is less than 2 in absolute value, the data appear to be reasonably symmetric and the normality assumption may be considered to be tenable.

The skewness statistics were calculated using the following code:

> skew(Stepping$HRFinal[Stepping$Freq=="Slow" & Stepping$Height=="Low"])

[1] -0.1514485

> se.skew(Stepping$HRFinal[Stepping$Freq=="Slow" & Stepping$Height=="Low"])

[1] 0.9128709

> skew.ratio(Stepping$HRFinal[Stepping$Freq=="Slow" & Stepping$Height=="Low"])

[1] -0.1659036

>

> skew(Stepping$HRFinal[Stepping$Freq=="Slow" & Stepping$Height=="High"])

[1] 0.6826736

> se.skew(Stepping$HRFinal[Stepping$Freq=="Slow" & Stepping$Height=="High"])

[1] 0.9128709

> skew.ratio(Stepping$HRFinal[Stepping$Freq=="Slow" & Stepping$Height=="High"])

[1] 0.7478315

>

> skew(Stepping$HRFinal[Stepping$Freq=="Medium" & Stepping$Height=="Low"])

[1] 0.5936005

> se.skew(Stepping$HRFinal[Stepping$Freq=="Medium" & Stepping$Height=="Low"])

[1] 0.9128709

> skew.ratio(Stepping$HRFinal[Stepping$Freq=="Medium" & Stepping$Height=="Low"])

[1] 0.6502567

>

> skew(Stepping$HRFinal[Stepping$Freq=="Medium" & Stepping$Height=="High"])

[1] 0.176017

> se.skew(Stepping$HRFinal[Stepping$Freq=="Medium" & Stepping$Height=="High"])

[1] 0.9128709

> skew.ratio(Stepping$HRFinal[Stepping$Freq=="Medium" & Stepping$Height=="High"])

[1] 0.192817

>

> skew(Stepping$HRFinal[Stepping$Freq=="Fast" & Stepping$Height=="Low"])

[1] 0.4080872

> se.skew(Stepping$HRFinal[Stepping$Freq=="Fast" & Stepping$Height=="Low"])

[1] 0.9128709

> skew.ratio(Stepping$HRFinal[Stepping$Freq=="Fast" & Stepping$Height=="Low"])

[1] 0.4470372

>

> skew(Stepping$HRFinal[Stepping$Freq=="Fast" & Stepping$Height=="High"])

[1] 0.01065949

> se.skew(Stepping$HRFinal[Stepping$Freq=="Fast" & Stepping$Height=="High"])

[1] 0.9128709

> skew.ratio(Stepping$HRFinal[Stepping$Freq=="Fast" & Stepping$Height=="High"])

[1] 0.01167689

For homogeneity of variance: The results of Levene’s test indicate that the homogeneity of variances assumption is tenable, *F*(5, 24) = 2.41, *p* = .07.

The following R code was used to generate Levene’s test:

**levenes.test(Stepping$HRFinal, Stepping$Freq:Stepping$Height)**

1. The interaction is not statistically significant, *F*(2, 24) = .54, *p* = .59.

The following R code was used to generate the ANOVA results:

**aov6 <- aov(Stepping$HRFinal~Stepping$Height\*Stepping$Freq)**

**summary(aov6)**

Df Sum Sq Mean Sq F value Pr(>F)

Stepping$Height 1 3499 3499 17.931 0.000291 \*\*\*

Stepping$Freq 2 3728 1864 9.551 0.000888 \*\*\*

Stepping$Height:Stepping$Freq 2 211 105 0.540 0.589898

Residuals 24 4684 195

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

c) The main effect due to stepping rate is statistically significant, *F*(2, 24) = 9.55, *p* < .0005.

d) According to the Bonferroni *post-hoc* test, the average heart rate is statistically significantly higher under fast stepping than it is under either slow or medium stepping. There is no statistically significant difference in heart rate between slow and medium stepping.

The following R code was used to generate the post-hoc Bonferroni test for the main effect:

**testInteractions(aov6, pairwise = "Stepping$Freq", adjustment = "bonferroni")**

F Test:

P-value adjustment method: bonferroni

Value Df Sum of Sq F Pr(>F)

Slow-Medium -8.4 1 352.8 1.8078 0.5740155

Slow-Fast -26.7 1 3564.5 18.2652 0.0007898 \*\*\*

Medium-Fast -18.3 1 1674.4 8.5803 0.0220130 \*

Residuals 24 4683.6

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

e) There is a statistically significant main effect due to step height, *F*(1, 24) = 17.93, *p* < .0005.

f) Based on the sample means, the average heart rate is statistically significantly higher when using the high step than when using the low step.



a) The design is a 2 x 3 balanced ANOVA with five participants per cell.

b) Weight Loss by Treatment

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Sum of Squares | df | Mean Square | F | Sig. |
| Between Groups | 453.8 | 2 | 226.9 | 18.28 | P < .0005 |
| Within Groups | 335 | 27 | 12.41 |  |  |
| Total | 788.8 | 29 |  |  |  |

c) Because the obtained *F*-statistic is larger for the two-way design (*F* = 33.53) than for the one-way design (*F* = 18.28), the two-way design provides a more powerful test of the treatment effect than the one-way design. By adding gender to the design as a second factor, we lose one degree of freedom, but explain enough dependent variable variance to offset this loss. As a result, the MS within (unexplained variance) associated with the two-way design is smaller than it is for the one-way design, producing a more powerful test of the treatment effect as noted by the larger *F*-test associated with the treatment effect in the two-way design.

1. Yes. Non-smokers weigh more, on average, than smokers.
2. No. People with no CHD have the same mean weight as those with CHD.
3. No. The lines are parallel.
4. (ii)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Source of Variation | Sum of Squares | df | Mean Square | F | Sig. |
| CHD | 0 | 1 | 0 | 0 | 1 |
| Cigarette Use | 41,000 | 1 | 41,000 | 463.88 | < .0005 |
| Interaction | 0 | 1 | 0 | 0 | 1 |
| Error | 35,000 | 396 | 88.38 |  |  |
| Total | 76,000 | 399 |  |  |  |

* 1. No. In both cases, there are four sources of variance: two sources due to main effects, one due to an interaction effect, and the fourth due to remaining, unexplained error.
  2. The two-way ANOVA is more powerful than the one-way ANOVA if there is a reduction in the MSerror term resulting from the addition of the second independent variable and if that reduction is large enough to compensate for the loss of degrees of freedom associated with adding that second independent variable. A reduction in the MSerror term comes about when the second independent variable is able to explain a large enough part of the dependent variable variance left unexplained by the first independent variable.
  3. We can use the **ss.2way** function in the **pwr2** package to answer this question:

**install.packages("pwr2")**

**library(pwr2)**

Using the R code **ss.2way(a = 3, b = 5, alpha = 0.05, beta = 0.2, f.A = 0.25, f.B = 1, B = 100)**,we get an *n* = 11 per cell for a total of *N* = 165. Using the R code **ss.2way(a = 3, b = 5, alpha = 0.05, beta = 0.2, f.A = 1, f.B = 0.25, B = 100)**, we get an *n* = 14 per cell for a total of *N* = 210.