Practical 6

Trevor Atkins

10/13/2020

Part A: Exposure and vocabulary scores

1

```
Practical6a <- readRDS(file="Practical6A.rds")
str(Practical6a)
```

```
## 'data.frame':    84 obs. of  3 variables:
##  $ Subject: Factor w/ 42 levels "1","2","3","4",..: 1 2 3 4 5 6 7 8 9 10 ...
##  $ Time   : Factor w/ 2 levels "post","pre": 2 2 2 2 2 2 2 2 2 2 ...
##  $ Score  : int  9 10 8 3 8 9 4 12 10 13 ...
```

2

Subject and time are nominal variables and score is a interval variable.

3

The null hypothesis is that there is no difference in score between the pre-test and post-test. The alternative hypothesis is that there is a difference in score between the pre-test and post-test.

4

```
library("psych")
Pretest <- subset(Practical6a, Time == "pre")
Posttest <- subset(Practical6a, Time == "post")
aggregate(Score~Time, Practical6a, max)
```

```
##   Time Score
## 1 post    17
## 2  pre    15
```

```
aggregate(Score~Time, Practical6a, min)
```

```
##   Time Score
## 1 post     3
## 2  pre     3
```

```
aggregate(Score~Time, Practical6a, mean)
```

```
##   Time    Score
## 1 post 9.857143
## 2  pre 8.857143
```

```
aggregate(Score~Time, Practical6a, sd)
```

```
##   Time    Score
## 1 post 3.189705
## 2  pre 3.024868
```

```
describe(Practical6a$Score)
```

```
##    vars  n mean   sd median trimmed  mad min max range skew kurtosis   se
## X1    1 84 9.36 3.13      9    9.31 2.97   3  17    14 0.14    -0.53 0.34
```

Table X Descriptives for score as a function of test type

value	Overall	Pretest	Posttest
mean	9.36	9.86	8.86
max	17	17	15
min	3	3	3
sd	3.13	3.19	3.02

5

A statistical test that could be used to check whether there is a difference between the pre-test and the post-test is a paired-samples t-test.

6

```
shapiro.test(Practical6a$Score)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  Practical6a$Score
## W = 0.98488, p-value = 0.4328
```

```
t.test(Score~Time, data=Practical6a, paired=TRUE)
```

```
##
##  Paired t-test
##
## data:  Score by Time
## t = 2.2913, df = 41, p-value = 0.02716
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.1186001 1.8813999
## sample estimates:
## mean of the differences
##                       1
```

7

```
library('effsize')
```

```
##
## Attaching package: 'effsize'
```

```
## The following object is masked from 'package:psych':
##
##     cohen.d
```

```
cohen.d(Practical6a$Score,Practical6a$Time)
```

```
##
## Cohen's d
##
## d estimate: 0.3217111 (small)
## 95 percent confidence interval:
##      lower      upper
## -0.1151929  0.7586150
```

The effect size is 0.32 and is a small effect size.

8

```
library('pwr')
pwr.t.test(n=42, d=0.32, sig.level= 0.05)
```

```
##
##      Two-sample t test power calculation
##
##               n = 42
##               d = 0.32
##       sig.level = 0.05
##           power = 0.3051071
##     alternative = two.sided
##
## NOTE: n is number in *each* group
```

The meaningfulness of this outcome is we cannot determine whether the difference between the pretest and posttest is not significant because in order for a small effect size to be significant, 783 participants (per group in a means analysis) is needed. Even though the Shapiro-Wilk test allows us to assume that the data is normally distributed and we know how the test was implemented, more participants are required to detect the small effect size.

9

A paired samples t-test revealed that the pretest group and posttest group did not have much difference between score. This difference was significant, t=2.29, p = 0.03, 95% CI [0.12, 1.88]. This effect was of a small size, r2/d = 0.32.

Part B: Instruction and writing scores

1 and 2

The dependent variable is the writing score that is a interval measure, the independent variables include (no instruction, lectures, and guided writing) the instruction type that is a nominal measure.

3

In the case of independent variables, the independent variable has 3 levels.

4

```
Student <- seq(1, 30)
Type <- c(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2,
2, 2, 2, 2, 2, 2)
Score <- c(34, 58, 56, 47, 35, 31, 55, 65, 61, 27, 65, 54, 43, 57, 65, 49, 74, 79,
54, 65, 68, 87, 94, 69, 81, 75, 94, 78, 63, 78)
Practical6b = data.frame(Student, Type, Score)
Practical6b$Student <- as.factor(Practical6b$Student)
Practical6b$Type <- as.factor(Practical6b$Type)
```

5

The null hypothesis is that there is no difference between the three instruction types and their writing scores. The alternative hypothesis is that there is a difference between the three instruction types and their

writing scores.

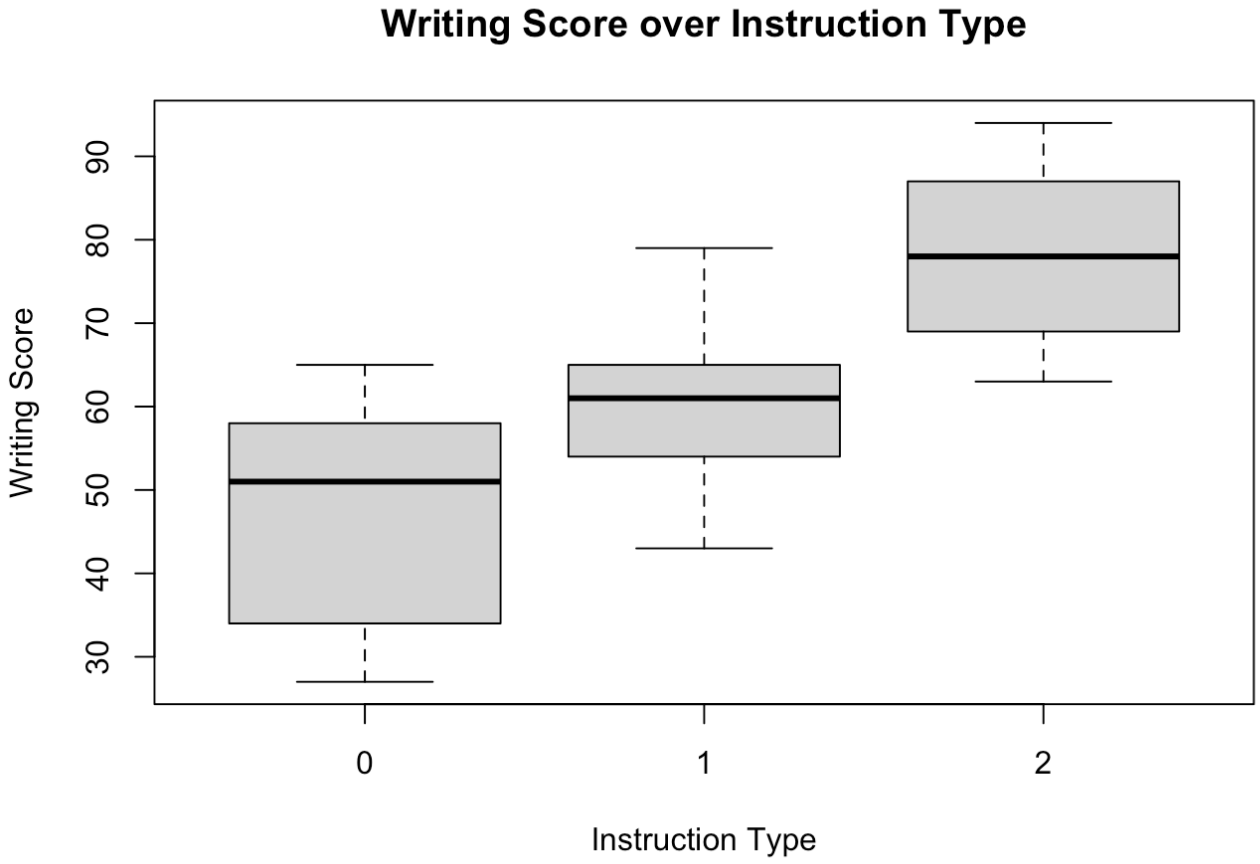
6

A statistical test that could be used is a one way ANOVA.

7

Figure 1: Boxplot of Writing Score over Instruction Type

```
boxplot(Score~Type, data=Practical6b, main="Writing Score over Instruction Type",
xlab="Instruction Type", ylab="Writing Score")
```



According to the boxplot, the guided writing instruction group seems to have performed the best in writing score out of the three instruction type groups.

8

```
by(Score, Type, describe)
```

```
## Type: 0
##    vars  n mean    sd median trimmed   mad min max range  skew kurtosis   se
## X1    1 10 46.9 13.96     51   47.12 17.79  27  65    38 -0.16    -1.84 4.42
## ------------------------------------------------------------
## Type: 1
##    vars  n mean    sd median trimmed   mad min max range skew kurtosis   se
## X1    1 10 60.5 11.16     61   60.38 10.38  43  79    36 0.11    -1.28 3.53
## ------------------------------------------------------------
## Type: 2
##    vars  n mean   sd median trimmed   mad min max range skew kurtosis   se
## X1    1 10 78.7 10.6     78   78.75 13.34  63  94    31 0.14    -1.44 3.35
```

```
describe(Practical6b$Score)
```

```
##    vars  n  mean   sd median trimmed   mad min max range  skew kurtosis   se
## X1    1 30 62.03 17.6     64   62.25 15.57  27  94    67 -0.14    -0.68 3.21
```

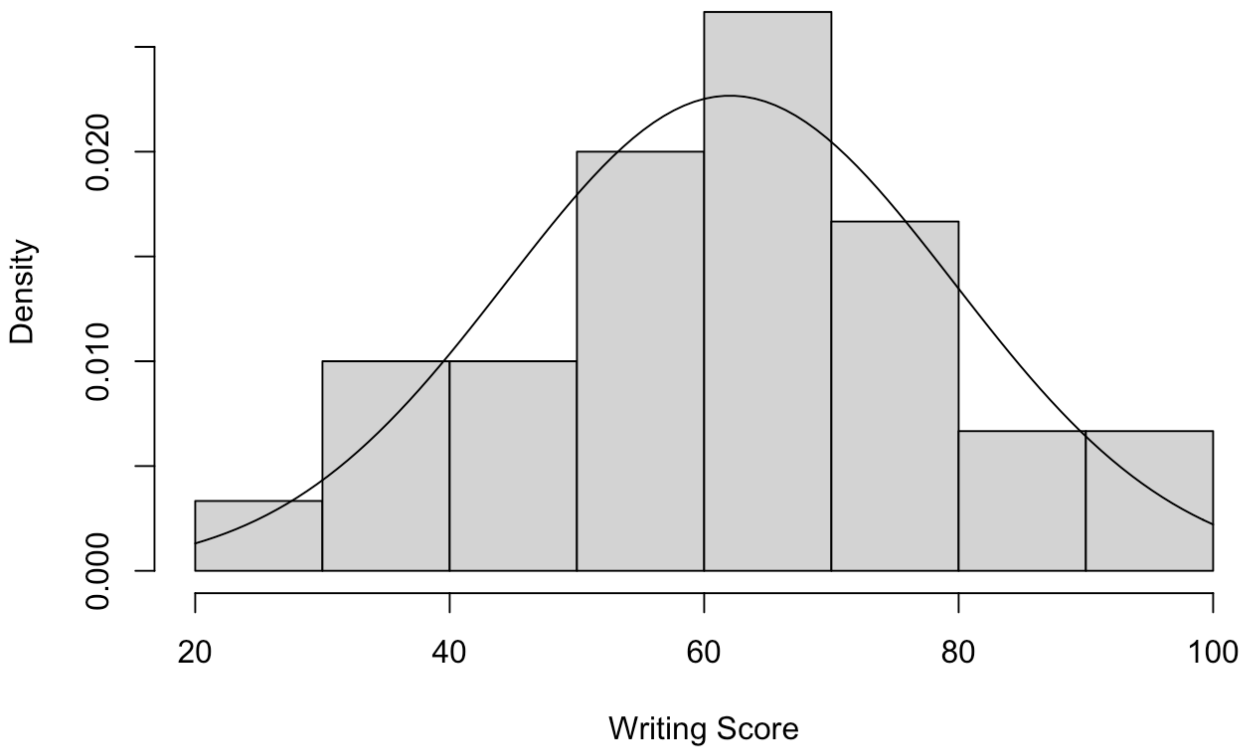
Table Y Descriptives for score as a function of type of instruction

value	Overall	No Instruction	Lectures
mean	62.03	46.9	60.5
max	94	65	79
min	27	27	43
sd	17.6	13.96	11.16

9

Figure 2: Histogram of Density over Vocabulary Score

```
hist(Practical6b$Score, prob=TRUE, xlab="Writing Score")
curve(dnorm(x, mean=mean(Practical6b$Score),sd=sd(Practical6b$Score)), add=TRUE)
```




```
library('car')
```

```
## Loading required package: carData
```

```
##
## Attaching package: 'car'
```

```
## The following object is masked from 'package:psych':
##
##     logit
```

```
leveneTest(Score~Type, data=Practical6b)
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##       Df F value Pr(>F)
## group  2  1.0496 0.3639
##       27
```

```
Test = aov(Score~Type, data=Practical6b)
summary(Test)
```

```
##             Df Sum Sq Mean Sq F value   Pr(>F)
## Type         2   5091    2546   17.68 1.24e-05 ***
## Residuals   27   3887     144
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
TukeyHSD(Test)
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = Score ~ Type, data = Practical6b)
##
## $Type
##     diff       lwr     upr     p adj
## 1-0 13.6  0.294904 26.9051 0.0443937
## 2-0 31.8 18.494904 45.1051 0.0000075
## 2-1 18.2  4.894904 31.5051 0.0059135
```

Since the p-value is 1.24e-05 and is less than the significance value of 0.05, we can reject the null hypothesis. The difference in means between the two groups for no instruction is 13.6, lectures is 31.8, and Guided Writing is 18.2.

10

```
r2 <- summary.lm(Test)$r.squared
sqrt(r2)
```

```
## [1] 0.7530231
```

```
library('sjstats')
```

```
## Registered S3 methods overwritten by 'lme4':
##   method                          from
##   cooks.distance.influence.merMod car
##   influence.merMod                car
##   dfbeta.influence.merMod         car
##   dfbetas.influence.merMod        car
```

```
##
## Attaching package: 'sjstats'
```

```
## The following object is masked from 'package:psych':
##
##     phi
```

```
library('pwr')
anova_stats(Test)
```

```
##           term df    sumsq   meansq statistic p.value etasq partial.etasq
## Type      Type  2 5091.467 2545.733    17.681       0 0.567         0.567
## ...2 Residuals 27 3887.500  143.981        NA      NA    NA            NA
##      omegasq partial.omegasq epsilonsq cohens.f power
## Type   0.527           0.527     0.535    1.144     1
## ...2      NA              NA        NA       NA    NA
```

```
pwr.anova.test(k=3, n=10, f=1.144 , sig.level=1.24e-05)
```

```
##
##      Balanced one-way analysis of variance power calculation
##
##               k = 3
##               n = 10
##               f = 1.144
##       sig.level = 1.24e-05
##           power = 0.6409735
##
## NOTE: n is number in each group
```

The effect size is 0.75, which is a large effect size.

11

The meaningfulness of this outcome is that there is a significant difference of a large magnitude between the instruction types and the writing scores. Specifically that there is a positive effect on writing scores in the order of no instruction, lecture, to guided writing.

12

There was a significant effect of instruction type on writing score, F(2, 27) = 1.0496, p = 0.3639. This effect was of a large size, r2/η2=0.7530231.

A Tukey post hoc analysis [or Bonferroni] revealed that the GW (guided writing) group (M=78.7, SD=10.6) did score significantly higher than the lecture group (M=60.5, SD=11.2), p < .001], 95% CI [18.494904,

45.1051] and the no instruction group (M=46.9, SD=13.96), p = 0.0443937, 95% CI [0.294904, 26.9051]. No significant differences were found between lecture group and no instruction group.

Part C: Subtitles and vocabulary learning

1

```
Practical6c <- readRDS(file="Practical6C.rds")
str(Practical6c)
```

```
## 'data.frame':    60 obs. of  3 variables:
##  $ Proficiency: Factor w/ 2 levels "High","Low": 2 2 2 2 2 2 2 2 2 2 ...
##  $ Subtitles  : Factor w/ 2 levels "L1","L2": 1 1 1 1 1 1 1 1 1 1 ...
##  $ Score      : int  10 5 9 6 6 11 12 13 8 10 ...
```

2

The dependent variable is score which has a interval measure, the independent variables are proficiency and subtitles which are both nominal measures.

3

The null hypothesis is that there is no difference between low proficient and high proficient Dutch-English learners and there is no effect of subtitles in the L2 (English subtitles in English shows) to subtitles in the L1 (Dutch subtitles in English shows) on vocabulary or proficiency score. The alternative hypothesis is that there is a difference between low proficient and high proficient Dutch-English learners and there is an effect of subtitles in the L2 to subtitles in the L1 on vocabulary or proficiency score.

4

A statistical test that could be used is factorial ANOVA or multiple linear regression.

5

```
library('psych')
by(Practical6c$Score, list(Practical6c$Proficiency, Practical6c$Subtitles), descri
be)
```

```
## : High
## : L1
##    vars  n mean   sd median trimmed  mad min max range  skew kurtosis   se
## X1    1 15 4.53 3.98      4    4.62 4.45  -2  10    12 -0.24     -1.4 1.03
## ------------------------------------------------------------
## : Low
## : L1
##    vars  n mean   sd median trimmed  mad min max range  skew kurtosis   se
## X1    1 15 9.73 2.52     10    9.85 2.97   5  13     8 -0.49    -1.04 0.65
## ------------------------------------------------------------
## : High
## : L2
##    vars  n  mean  sd median trimmed  mad min max range skew kurtosis  se
## X1    1 15 13.53 3.5     13   13.62 4.45   8  18    10 0.05    -1.67 0.9
## ------------------------------------------------------------
## : Low
## : L2
##    vars  n mean   sd median trimmed  mad min max range skew kurtosis   se
## X1    1 15  4.2 3.78      4    4.15 2.97  -3  12    15 0.19    -0.27 0.98
```

```
describe(Practical6c$Score)
```

```
##    vars  n mean   sd median trimmed  mad min max range  skew kurtosis   se
## X1    1 60    8 5.18    8.5    7.96 5.19  -3  18    21 -0.03    -0.57 0.67
```

Table Z Descriptives for score as a function of proficiency and subtitles

value	Overall	High L1	Low L1	High L2	Low L2
mean	8	4.53	9.73	13.53	4.2
max	18	10	13	18	12
min	-3	-2	5	8	-3
sd	5.18	3.98	2.52	3.5	3.78

6

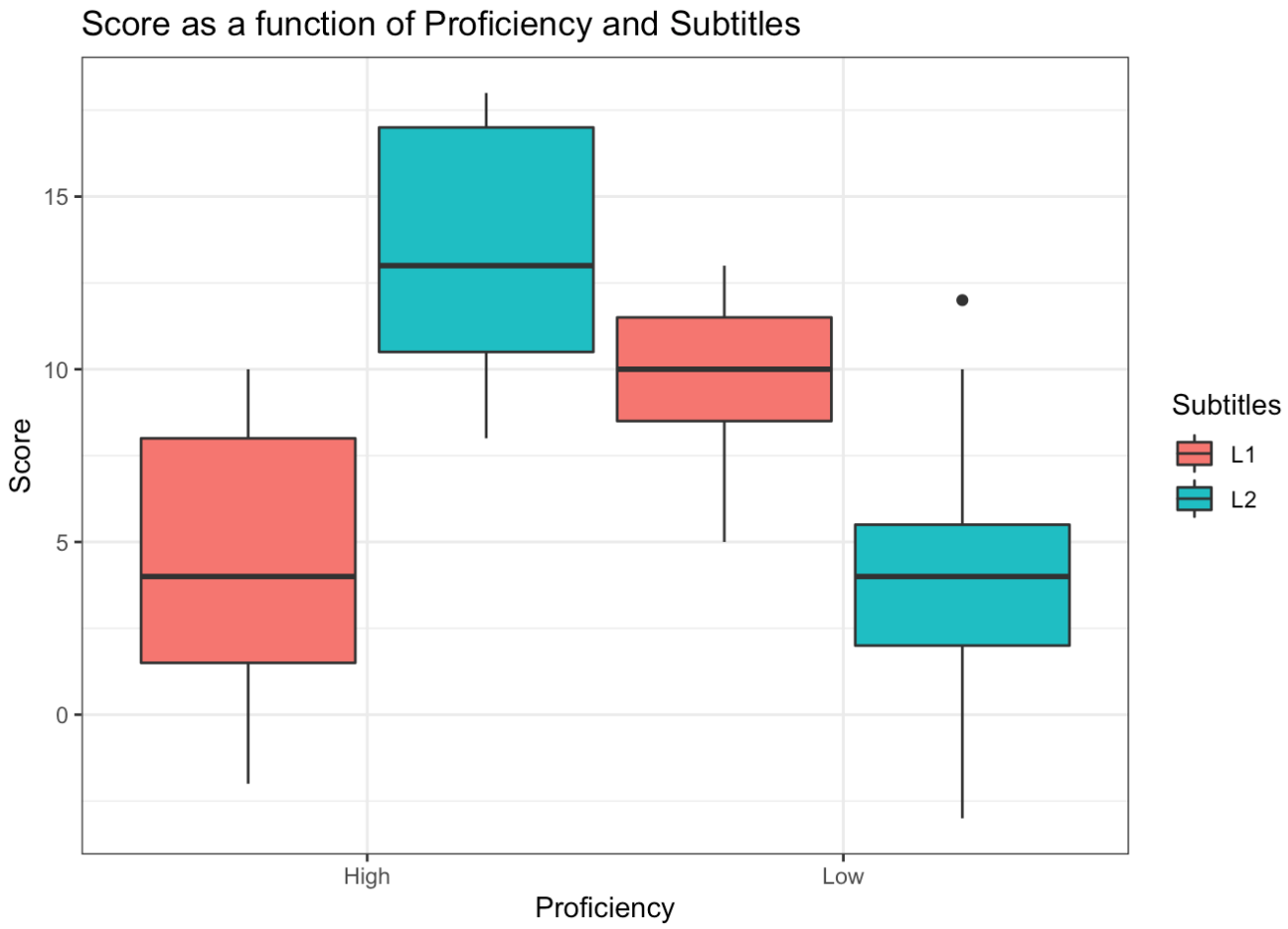
Figure 3 Plot of Score as a function of Proficiency and Subtitles

```
#boxplot(Practical6c$Score ~ Practical6c$Proficiency + Practical6c$Subtitles)
library('ggplot2')
```

```
##
## Attaching package: 'ggplot2'
```

```
## The following objects are masked from 'package:psych':
##
##     %+%, alpha
```

```
ggplot(data = Practical6c, aes(x = Proficiency, y = Score, fill = Subtitles)) +
geom_boxplot(aes(fill = Subtitles), width = 1) + theme_bw() + ggtitle("Score as a
function of Proficiency and Subtitles") + labs(x="Proficiency",y="Score",
fill="Subtitles")
```



7

```
library('car')
leveneTest(Practical6c$Score, interaction(Practical6c$Proficiency, Practical6c$Sub
titles))
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##       Df F value Pr(>F)
## group  3  1.3321 0.2731
##       56
```

```
library('pastecs')
by(Practical6c$Score,Practical6c$Proficiency, stat.desc,
basic=FALSE, norm=TRUE)
```

```
## Practical6c$Proficiency: High
##       median         mean      SE.mean CI.mean.0.95          var      std.dev
##    9.5000000    9.0333333    1.0726306    2.1937759   34.5160920    5.8750397
##     coef.var     skewness     skew.2SE     kurtosis     kurt.2SE   normtest.W
##    0.6503734   -0.1425327   -0.1669421   -0.9376538   -0.5629893    0.9555810
##   normtest.p
##    0.2378735
## ------------------------------------------------------------
## Practical6c$Proficiency: Low
##       median         mean      SE.mean CI.mean.0.95          var      std.dev
##    6.5000000    6.9666667    0.7723428    1.5796184   17.8954023    4.2302958
##     coef.var     skewness     skew.2SE     kurtosis     kurt.2SE   normtest.W
##    0.6072195   -0.3923684   -0.4595636   -0.7491506   -0.4498076    0.9470286
##   normtest.p
##    0.1406835
```

```
by(Practical6c$Score,Practical6c$Subtitles, stat.desc,
basic=FALSE, norm=TRUE)
```

```
## Practical6c$Subtitles: L1
##       median         mean      SE.mean CI.mean.0.95          var      std.dev
##   8.00000000   7.13333333   0.76823927   1.57122572  17.70574713   4.20781976
##     coef.var     skewness     skew.2SE     kurtosis     kurt.2SE   normtest.W
##   0.58988127  -0.61629076  -0.72183384  -0.63250102  -0.37976845   0.93523309
##   normtest.p
##   0.06772492
## ------------------------------------------------------------
## Practical6c$Subtitles: L2
##       median         mean      SE.mean CI.mean.0.95          var      std.dev
##  10.00000000   8.86666667   1.08567828   2.22046139  35.36091954   5.94650482
##     coef.var     skewness     skew.2SE     kurtosis     kurt.2SE   normtest.W
##   0.67065844  -0.03451724  -0.04042850  -1.08101974  -0.64906961   0.95595842
##   normtest.p
##   0.24336609
```

```
contrasts(Practical6c$Proficiency) <- c(-1, 1)
contrasts(Practical6c$Subtitles) <- c(-1, 1)
model = aov(Score ~ Proficiency*Subtitles, data=Practical6c)
model
```

```
## Call:
##    aov(formula = Score ~ Proficiency * Subtitles, data = Practical6c)
##
## Terms:
##                 Proficiency Subtitles Proficiency:Subtitles Residuals
## Sum of Squares      64.0667   45.0667              792.0667  682.8000
## Deg. of Freedom           1         1                     1        56
##
## Residual standard error: 3.491827
## Estimated effects are balanced
```

```
library('car')
anovamodel <- Anova(model, type="III")
anovamodel
```

```
## Anova Table (Type III tests)
##
## Response: Score
##                       Sum Sq Df  F value    Pr(>F)
## (Intercept)           3840.0  1 314.9385 < 2.2e-16 ***
## Proficiency             64.1  1   5.2544   0.02567 *
## Subtitles               45.1  1   3.6962   0.05963 .
## Proficiency:Subtitles  792.1  1  64.9615 6.163e-11 ***
## Residuals              682.8 56
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The different null hypothesis that I formulated before can be rejected since the p-values are all less than or equal to the significance value of 0.05.

8

```
library('sjstats')
anova_stats(model)
```

```
##                                     term df   sumsq  meansq statistic
## Proficiency                   Proficiency  1  64.067  64.067     5.254
## Subtitles                       Subtitles  1  45.067  45.067     3.696
## Proficiency:Subtitles Proficiency:Subtitles  1 792.067 792.067    64.962
## ...4                              Residuals 56 682.800  12.193        NA
##                       p.value etasq partial.etasq omegasq partial.omegasq
## Proficiency             0.026 0.040         0.086   0.032           0.066
## Subtitles               0.060 0.028         0.062   0.021           0.043
## Proficiency:Subtitles   0.000 0.500         0.537   0.489           0.516
## ...4                       NA    NA            NA      NA              NA
##                       epsilonsq cohens.f power
## Proficiency               0.033    0.306 0.630
## Subtitles                 0.021    0.257 0.485
## Proficiency:Subtitles     0.492    1.077 1.000
## ...4                         NA       NA    NA
```

The effect size for the two main effects are: Proficiency = 0.066 and Subtitles = 0.043. Proficiency is a medium effect size and Subtitles is a small effect size nearing on medium. The interaction is 0.516 which is a large effect size.

10

There was a significant main effect of Proficiency on Writing Score, F(1, 1) = 5.2544, p <.001. This effect was medium, ωp2 = 0.066. There was also a significant main effect of Subtitles on Writing Score, F(1, 1) = 3.6962, p = 0.05963. This effect was small, ωp2 = 0.043.

There was a significant interaction effect between Proficiency and Subtitles, F(1, 1) = 64.9615, p <.001}. This effect was large, ωp2 = 0.516. This interaction showed that it is the most significant effect since partial ωp2 assesses the effect sizes of each effect while partialling out the other effects. It also shows that a combination of proficiency and subtitles shows more significant correlations of effect on writing score. Specifically, the writing score was higher for High Proficiency (M = 9.03, SD = 5.88) than for Low Proficiency (M = 6.97, SD = 4.23) in L1 Subtitles, but they differed in L2 Subtitles with higher scores for L2 Subtitles (M = 8.87, SD = 5.95) than for L1 Subtitles (M = 7.13, SD = 4.21) in L1 Subtitles.