MANOVA
Part II
Psych 516

Mike Strube

November 26, 2018

1 Preliminaries

The RStudio workspace and console panes are cleared of old output, variables, and other miscellaneous debris. Then some packages are loaded and the required data files are input.

1.1 Clear the Console Panes and Load Packages

```
options(replace.assign = TRUE, width = 65, digits = 4, scipen = 4,
    fig.width = 4, fig.height = 4)
# Clear the workspace and console.
rm(list = ls(all = TRUE))
cat("\f")
```

```
# Turn off showing of significance asterisks.
options(show.signif.stars = F)
# Set the contrast option; important for ANOVAs.
options(contrasts = c("contr.sum", "contr.poly"))
how_long <- Sys.time()
set.seed(123)
library(knitr)
```

```
library(psych)

## Warning:  package 'psych' was built under R version 3.5.1

library(ggplot2)

## Warning:  package 'ggplot2' was built under R version 3.5.1
##
## Attaching package:  'ggplot2'
## The following objects are masked from 'package:psych':
##
##    %+%, alpha

library(MASS)
library(sciplot)
library(dplyr)
```

```
## Warning:  package 'dplyr' was built under R version 3.5.1
##
## Attaching package:  'dplyr'
## The following object is masked from 'package:MASS':
##
##     select
## The following objects are masked from 'package:stats':
##
##     filter, lag
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(aod)
library(MVN)
```

```
## sROC 0.1-2 loaded
```

```
library(boot)
```

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

```
library(car)
```

```
## Warning:  package 'car' was built under R version 3.5.1
## Loading required package:  carData
##
## Attaching package:  'car'
## The following object is masked from 'package:boot':
##
##     logit
## The following object is masked from 'package:dplyr':
##
##     recode
## The following object is masked from 'package:psych':
##
##     logit
```

```
library(LogisticDx)
library(biotools)
```

```
## Loading required package:  rpanel
## Loading required package:  tcltk
## Package `rpanel', version 1.1-4:  type help(rpanel) for summary information
##
## Attaching package:  'rpanel'
## The following object is masked from 'package:boot':
##
##     poisons
## Loading required package:  tkrplot
## Loading required package:  lattice
##
## Attaching package:  'lattice'
```

```
## The following object is masked from 'package:boot':
##
##     melanoma
## Loading required package:  SpatialEpi
## Loading required package:  sp
## ---
## biotools version 3.1
##
library(multcomp)
## Loading required package:  mvtnorm
## Loading required package:  survival
##
## Attaching package:  'survival'
## The following object is masked from 'package:boot':
##
##     aml
## The following object is masked from 'package:aod':
##
##     rats
## Loading required package:  TH.data
##
## Attaching package:  'TH.data'
## The following object is masked from 'package:MASS':
##
##     geyser
library(candisc)
## Loading required package:  heplots
##
## Attaching package:  'heplots'
## The following object is masked from 'package:biotools':
##
##     boxM
##
## Attaching package:  'candisc'
## The following object is masked from 'package:stats':
##
##     cancor
library(ez)
library(GGally)
##
## Attaching package:  'GGally'
## The following object is masked from 'package:dplyr':
##
##     nasa
library(qqplotr)
##
## Attaching package:  'qqplotr'
## The following objects are masked from 'package:ggplot2':
##
##     stat_qq_line, StatQqLine
```

```
library(gridExtra)

##
## Attaching package:  'gridExtra'
## The following object is masked from 'package:dplyr':
##
##    combine

library(reshape)

##
## Attaching package:  'reshape'
## The following object is masked from 'package:dplyr':
##
##    rename

library(emmeans)

## Warning:  package 'emmeans' was built under R version 3.5.1
## NOTE: As of emmeans versions > 1.2.3,
##      The 'cld' function will be deprecated in favor of 'CLD'.
##      You may use 'cld' only if you have package:multcomp attached.
##
## Attaching package:  'emmeans'
## The following object is masked from 'package:GGally':
##
##    pigs
## The following object is masked from 'package:multcomp':
##
##    cld
```

1.2 Data

```
setwd("C:\\Courses\\Psychology 516\\PowerPoint\\2018")

Skills <- read.table("manova.csv", sep = ",", header = TRUE)
Skills <- as.data.frame(Skills)
```

1.3 Data Modifications

Residualized versions of continuous predictors are created so that preliminary analyses are not contaminated by outcome differences. Labeled variables are created to assist in creation of some tables and graphs. Dummy codes and linear combinations are created for specialized analyses (not all used here).

```
# Residuals
Skills$P_Verbal_R <- lm(P_Verbal ~ as.factor(Group), data = Skills)$residuals
Skills$P_Quant_R <- lm(P_Quant ~ as.factor(Group), data = Skills)$residuals
Skills$C_Verbal_R <- lm(C_Verbal ~ as.factor(Group), data = Skills)$residuals
Skills$C_Quant_R <- lm(C_Quant ~ as.factor(Group), data = Skills)$residuals

# Labels
Skills$Tx_P2[Skills$Tx_P == "1"] <- "No Paper Tx"
```

```
Skills$Tx_P2[Skills$Tx_P == "2"] <- "Paper Tx"

Skills$Tx_C2[Skills$Tx_C == "1"] <- "No Computer Tx"
Skills$Tx_C2[Skills$Tx_C == "2"] <- "Computer Tx"

Skills$Group2[Skills$Group == "1"] <- "No Paper Tx and No Computer Tx"
Skills$Group2[Skills$Group == "2"] <- "Paper Tx and No Computer Tx"
Skills$Group2[Skills$Group == "3"] <- "No Paper Tx and Computer Tx"
Skills$Group2[Skills$Group == "4"] <- "Paper Tx and Computer Tx"

Skills$Group3[Skills$Group == "1"] <- "No P, No C"
Skills$Group3[Skills$Group == "2"] <- "P, No C"
Skills$Group3[Skills$Group == "3"] <- "No P, C"
Skills$Group3[Skills$Group == "4"] <- "P, C"

# Dummy variables to be used in between-groups analyses.
Skills$D1[Skills$Group == 1] <- 1
Skills$D2[Skills$Group == 1] <- 0
Skills$D3[Skills$Group == 1] <- 0
Skills$D4[Skills$Group == 1] <- 0
Skills$D1[Skills$Group == 2] <- 0
Skills$D2[Skills$Group == 2] <- 1
Skills$D3[Skills$Group == 2] <- 0
Skills$D4[Skills$Group == 2] <- 0
Skills$D1[Skills$Group == 3] <- 0
Skills$D2[Skills$Group == 3] <- 0
Skills$D3[Skills$Group == 3] <- 1
Skills$D4[Skills$Group == 3] <- 0
Skills$D1[Skills$Group == 4] <- 0
Skills$D2[Skills$Group == 4] <- 0
Skills$D3[Skills$Group == 4] <- 0
Skills$D4[Skills$Group == 4] <- 1

# Add contrast codes to reflect main effects and interactions.
Skills$C1[Skills$Group == 1] <- -1
Skills$C2[Skills$Group == 1] <- -1
Skills$C3[Skills$Group == 1] <- 1
Skills$C1[Skills$Group == 2] <- 1
Skills$C2[Skills$Group == 2] <- -1
Skills$C3[Skills$Group == 2] <- -1
Skills$C1[Skills$Group == 3] <- -1
Skills$C2[Skills$Group == 3] <- 1
Skills$C3[Skills$Group == 3] <- -1
Skills$C1[Skills$Group == 4] <- 1
Skills$C2[Skills$Group == 4] <- 1
Skills$C3[Skills$Group == 4] <- 1

# Add contrast codes to reflect specialized comparisons.
Skills$S1[Skills$Group == 1] <- 3
Skills$S2[Skills$Group == 1] <- 0
Skills$S3[Skills$Group == 1] <- 0
Skills$S1[Skills$Group == 2] <- -1
Skills$S2[Skills$Group == 2] <- 2
```

```r
Skills$S3[Skills$Group == 2] <- 0
Skills$S1[Skills$Group == 3] <- -1
Skills$S2[Skills$Group == 3] <- -1
Skills$S3[Skills$Group == 3] <- 1
Skills$S1[Skills$Group == 4] <- -1
Skills$S2[Skills$Group == 4] <- -1
Skills$S3[Skills$Group == 4] <- -1

# Outcome linear combinations to be used in repeated measures
# analyses.
Skills$Sum <- Skills$P_Verbal + Skills$P_Quant + Skills$C_Verbal +
    Skills$C_Quant
Skills$Domain <- Skills$P_Verbal - Skills$P_Quant + Skills$C_Verbal -
    Skills$C_Quant
Skills$Mode <- Skills$P_Verbal + Skills$P_Quant - Skills$C_Verbal -
    Skills$C_Quant
Skills$DxM <- Skills$P_Verbal - Skills$P_Quant - Skills$C_Verbal +
    Skills$C_Quant

# Create a non-factor version of the condition variables before
# converting them to factors.
Skills$Tx_P_NF <- Skills$Tx_P
Skills$Tx_C_NF <- Skills$Tx_C

# Convert to factors
Skills$Tx_P = factor(Skills$Tx_P, levels = c(1, 2), labels = c("No Tx(P)",
    "Tx(P)"))
Skills$Tx_C = factor(Skills$Tx_C, levels = c(1, 2), labels = c("No Tx(C)",
    "Tx(C)"))

# Sort file by Group
Skills <- Skills[order(Skills$Group), ]
```

2 Data Characteristics

These hypothetical data simulate a training study in which students are given training to take tests of verbal and quantitative ability. The training is conducted either with paper-and-pencil (standard) tests or with computer-administered tests (or both) and the tests are administered in both formats. The basic nature of these data is explored here.

2.1 Some Descriptive Statistics

Some basic descriptive statistics give an initial glimpse of the data.

```
describeBy(Skills[, 2:5], group = Skills$Group)

##
##  Descriptive statistics by group
## group: 1
##           vars  n  mean    sd median trimmed   mad   min   max
## P_Verbal     1 25 47.86 10.59  48.94   48.13  9.42 26.20 64.23
## P_Quant      2 25 47.52  9.99  46.27   47.74 10.40 23.98 66.84
## C_Verbal     3 25 45.72 10.84  45.45   45.65 10.75 21.06 71.22
## C_Quant      4 25 46.28 10.70  48.09   46.66 12.55 22.20 62.60
##          range  skew kurtosis   se
## P_Verbal 38.04 -0.13    -0.86 2.12
## P_Quant  42.85 -0.12    -0.29 2.00
## C_Verbal 50.16  0.05    -0.17 2.17
## C_Quant  40.41 -0.31    -0.87 2.14
## ------------------------------------------------
## group: 2
##           vars  n  mean    sd median trimmed   mad   min   max
## P_Verbal     1 25 61.86 12.84  59.35   61.91  7.72 27.31 87.95
## P_Quant      2 25 71.83 10.87  70.09   72.39 10.67 40.38 94.83
## C_Verbal     3 25 48.77 10.28  50.08   48.51  9.93 30.46 71.87
## C_Quant      4 25 49.65 10.97  47.15   49.77  9.42 23.54 70.41
##          range  skew kurtosis   se
## P_Verbal 60.64 -0.05     0.78 2.57
## P_Quant  54.46 -0.62     1.22 2.17
## C_Verbal 41.41  0.23    -0.64 2.06
## C_Quant  46.87 -0.01    -0.44 2.19
## ------------------------------------------------
## group: 3
##           vars  n  mean    sd median trimmed   mad   min   max
## P_Verbal     1 25 24.17 11.09  23.60   24.03 12.56  6.98 42.79
## P_Quant      2 25 32.78  9.35  31.30   32.33  9.65 15.05 56.80
## C_Verbal     3 25 53.36 10.30  55.94   53.78  6.77 33.88 70.76
## C_Quant      4 25 60.61  9.01  61.38   60.39  9.86 46.46 77.37
##          range  skew kurtosis   se
## P_Verbal 35.80  0.07    -1.28 2.22
## P_Quant  41.74  0.51     0.01 1.87
## C_Verbal 36.88 -0.59    -0.66 2.06
## C_Quant  30.91  0.14    -1.09 1.80
## ------------------------------------------------
## group: 4
##           vars  n  mean   sd median trimmed  mad   min    max
## P_Verbal     1 25 92.45 5.77  92.33   92.70 7.01 82.09 100.00
```

```
## P_Quant     2 25 81.93 8.76  82.15   82.41 7.82 62.46  97.12
## C_Verbal    3 25 82.43 8.78  82.06   82.81 9.74 65.30  95.87
## C_Quant     4 25 91.51 6.26  89.99   91.67 7.10 79.47 100.00
##          range  skew kurtosis   se
## P_Verbal 17.91 -0.09    -1.29 1.15
## P_Quant  34.66 -0.52    -0.42 1.75
## C_Verbal 30.57 -0.43    -0.70 1.76
## C_Quant  20.53 -0.04    -1.27 1.25
```

```
with(Skills, tapply(P_Verbal, list(Tx_P, Tx_C), mean))
```

```
##          No Tx(C) Tx(C)
## No Tx(P)    47.86 24.17
## Tx(P)       61.86 92.45
```

```
with(Skills, tapply(P_Quant, list(Tx_P, Tx_C), mean))
```

```
##          No Tx(C) Tx(C)
## No Tx(P)    47.52 32.78
## Tx(P)       71.83 81.93
```

```
with(Skills, tapply(C_Verbal, list(Tx_P, Tx_C), mean))
```

```
##          No Tx(C) Tx(C)
## No Tx(P)    45.72 53.36
## Tx(P)       48.77 82.43
```

```
with(Skills, tapply(C_Quant, list(Tx_P, Tx_C), mean))
```

```
##          No Tx(C) Tx(C)
## No Tx(P)    46.28 60.61
## Tx(P)       49.65 91.51
```

```
with(Skills, tapply(P_Verbal, list(Tx_P, Tx_C), sd))
```

```
##          No Tx(C)  Tx(C)
## No Tx(P)    10.59 11.089
## Tx(P)       12.84  5.766
```

```
with(Skills, tapply(P_Quant, list(Tx_P, Tx_C), sd))
```

```
##          No Tx(C) Tx(C)
## No Tx(P)    9.985 9.353
## Tx(P)      10.873 8.764
```

```
with(Skills, tapply(C_Verbal, list(Tx_P, Tx_C), sd))
```

```
##          No Tx(C)  Tx(C)
## No Tx(P)    10.84 10.302
## Tx(P)       10.28  8.784
```

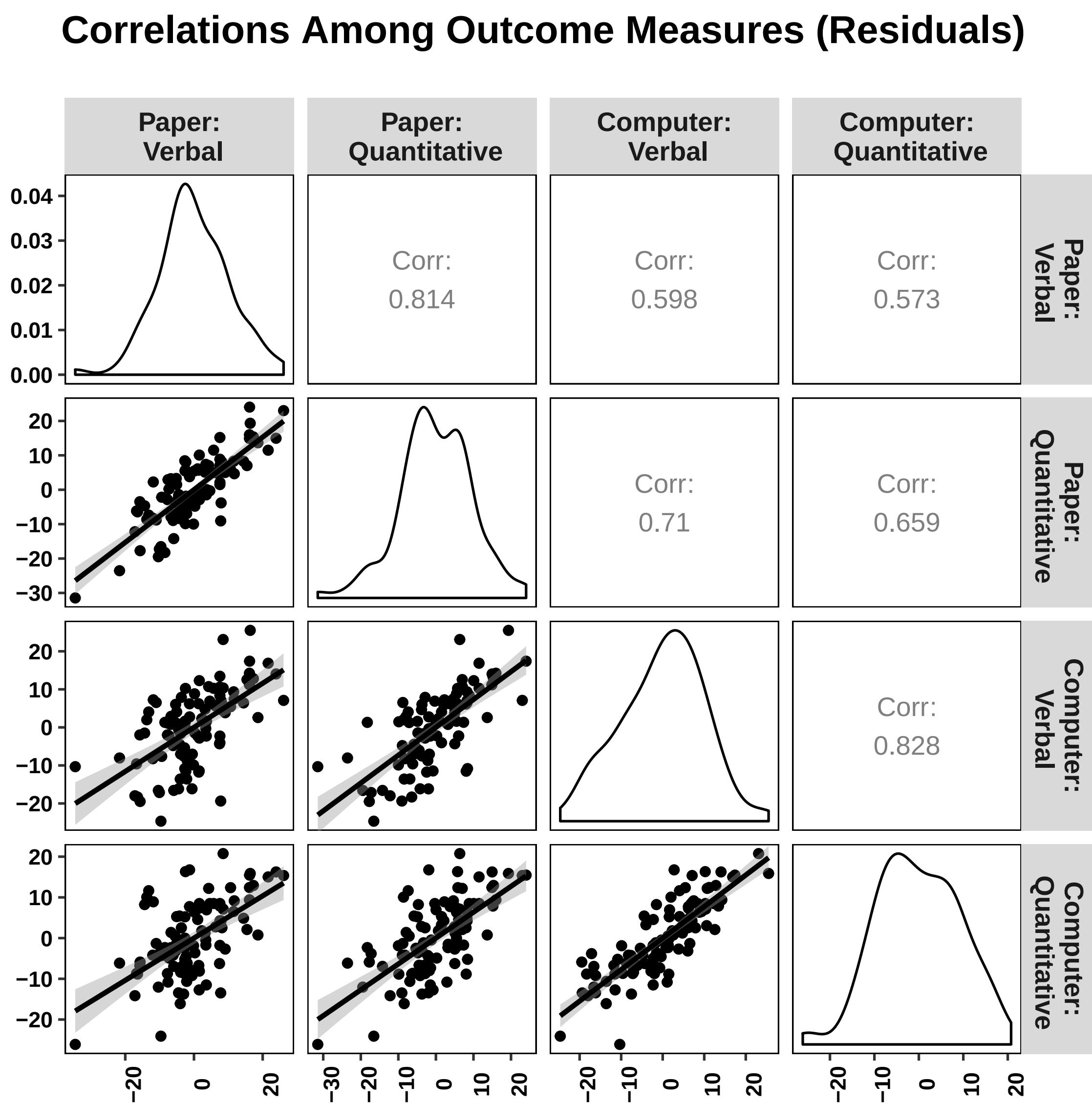
```
with(Skills, tapply(C_Quant, list(Tx_P, Tx_C), sd))
```

```
##          No Tx(C) Tx(C)
## No Tx(P)    10.70 9.005
## Tx(P)       10.97 6.262
```

2.2 Basic Visualization

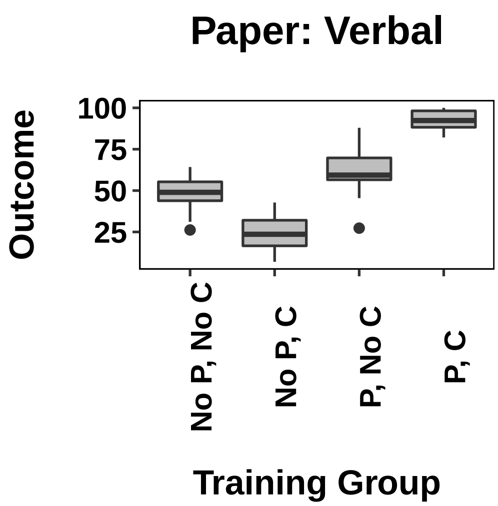
The basic nature of the data is easily viewed with some simple graphics.

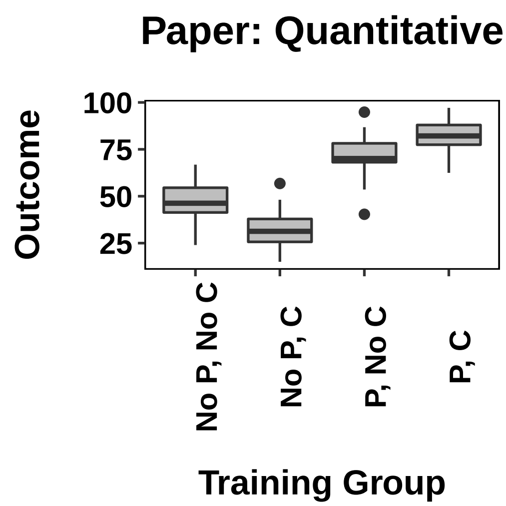
```
ggpairs(Skills[9:12], lower = list(continuous = "smooth"), upper = list(continuous = "cor"),
    columnLabels = c("Paper:\n Verbal", "Paper:\n Quantitative", "Computer:\n Verbal",
        "Computer:\n Quantitative")) + theme(text = element_text(size = 14,
    family = "sans", color = "black", face = "bold"), axis.text.y = element_text(colour = "black",
    size = 9, face = "bold"), axis.text.x = element_text(colour = "black",
    size = 9, face = "bold", angle = 90), axis.title.x = element_text(margin = margin(15,
    0, 0, 0), size = 16), axis.title.y = element_text(margin = margin(0,
    15, 0, 0), size = 16), axis.line.x = element_blank(), axis.line.y = element_blank(),
    plot.title = element_text(size = 16, face = "bold", margin = margin(0,
        0, 20, 0), hjust = 0.5), panel.background = element_rect(fill = "white",
        linetype = 1, color = "black"), panel.grid.major = element_blank(),
    panel.grid.minor = element_blank(), plot.background = element_rect(fill = "white"),
    plot.margin = unit(c(1, 1, 1, 1), "cm"), legend.position = "bottom",
    legend.title = element_blank()) + ggtitle("Correlations Among Outcome Measures (Residuals)")
```

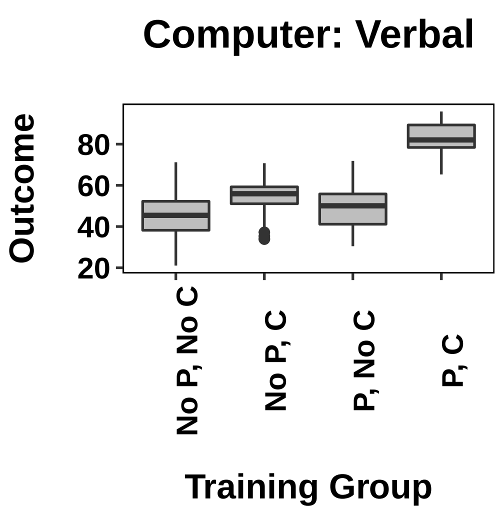


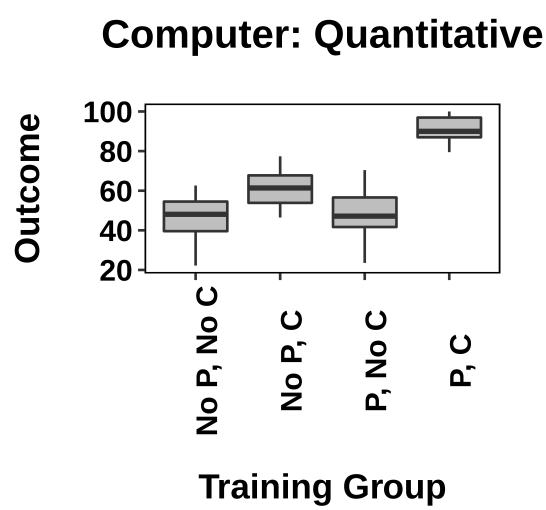

```
Skills$Group4 <- factor(Skills$Group3, levels = c("No P, No C", "No P, C",
    "P, No C", "P, C"), labels = c("No P, No C", "No P, C", "P, No C",
    "P, C"))

p1 <- ggplot(Skills, aes(x = as.factor(Group4), y = P_Verbal)) + geom_boxplot(fill = "gray") +
    ylab("Outcome") + xlab("Training Group") + theme(text = element_text(size = 14,
    family = "sans", color = "black", face = "bold"), axis.text.y = element_text(colour = "black",
    size = 12, face = "bold"), axis.text.x = element_text(colour = "black",
    size = 12, face = "bold", angle = 90), axis.title.x = element_text(margin = margin(15,
    0, 0, 0), size = 14), axis.title.y = element_text(margin = margin(0,
    15, 0, 0), size = 14), axis.line.x = element_blank(), axis.line.y = element_blank(),
    plot.title = element_text(size = 16, face = "bold", margin = margin(0,
        0, 20, 0), hjust = 0.5), panel.background = element_rect(fill = "white",
        linetype = 1, color = "black"), panel.grid.major = element_blank(),
```

```r
    panel.grid.minor = element_blank(), plot.background = element_rect(fill = "white"),
    plot.margin = unit(c(1, 1, 1, 1), "cm"), legend.position = "bottom",
    legend.title = element_blank()) + ggtitle("Paper: Verbal")

p2 <- ggplot(Skills, aes(x = as.factor(Group4), y = P_Quant)) + geom_boxplot(fill = "gray") +
    ylab("Outcome") + xlab("Training Group") + theme(text = element_text(size = 14,
    family = "sans", color = "black", face = "bold"), axis.text.y = element_text(colour = "black",
    size = 12, face = "bold"), axis.text.x = element_text(colour = "black",
    size = 12, face = "bold", angle = 90), axis.title.x = element_text(margin = margin(15,
    0, 0, 0), size = 14), axis.title.y = element_text(margin = margin(0,
    15, 0, 0), size = 14), axis.line.x = element_blank(), axis.line.y = element_blank(),
    plot.title = element_text(size = 16, face = "bold", margin = margin(0,
        0, 20, 0), hjust = 0.5), panel.background = element_rect(fill = "white",
        linetype = 1, color = "black"), panel.grid.major = element_blank(),
    panel.grid.minor = element_blank(), plot.background = element_rect(fill = "white"),
    plot.margin = unit(c(1, 1, 1, 1), "cm"), legend.position = "bottom",
    legend.title = element_blank()) + ggtitle("Paper: Quantitative")

p3 <- ggplot(Skills, aes(x = as.factor(Group4), y = C_Verbal)) + geom_boxplot(fill = "gray") +
    ylab("Outcome") + xlab("Training Group") + theme(text = element_text(size = 14,
    family = "sans", color = "black", face = "bold"), axis.text.y = element_text(colour = "black",
    size = 12, face = "bold"), axis.text.x = element_text(colour = "black",
    size = 12, face = "bold", angle = 90), axis.title.x = element_text(margin = margin(15,
    0, 0, 0), size = 14), axis.title.y = element_text(margin = margin(0,
    15, 0, 0), size = 14), axis.line.x = element_blank(), axis.line.y = element_blank(),
    plot.title = element_text(size = 16, face = "bold", margin = margin(0,
        0, 20, 0), hjust = 0.5), panel.background = element_rect(fill = "white",
        linetype = 1, color = "black"), panel.grid.major = element_blank(),
    panel.grid.minor = element_blank(), plot.background = element_rect(fill = "white"),
    plot.margin = unit(c(1, 1, 1, 1), "cm"), legend.position = "bottom",
    legend.title = element_blank()) + ggtitle("Computer: Verbal")

p4 <- ggplot(Skills, aes(x = as.factor(Group4), y = C_Quant)) + geom_boxplot(fill = "gray") +
    ylab("Outcome") + xlab("Training Group") + theme(text = element_text(size = 14,
    family = "sans", color = "black", face = "bold"), axis.text.y = element_text(colour = "black",
    size = 12, face = "bold"), axis.text.x = element_text(colour = "black",
    size = 12, face = "bold", angle = 90), axis.title.x = element_text(margin = margin(15,
    0, 0, 0), size = 14), axis.title.y = element_text(margin = margin(0,
    15, 0, 0), size = 14), axis.line.x = element_blank(), axis.line.y = element_blank(),
    plot.title = element_text(size = 16, face = "bold", margin = margin(0,
        0, 20, 0), hjust = 0.5), panel.background = element_rect(fill = "white",
        linetype = 1, color = "black"), panel.grid.major = element_blank(),
    panel.grid.minor = element_blank(), plot.background = element_rect(fill = "white"),
    plot.margin = unit(c(1, 1, 1, 1), "cm"), legend.position = "bottom",
    legend.title = element_blank()) + ggtitle("Computer: Quantitative")
grid.arrange(p1, p2, p3, p4, nrow = 2)
```









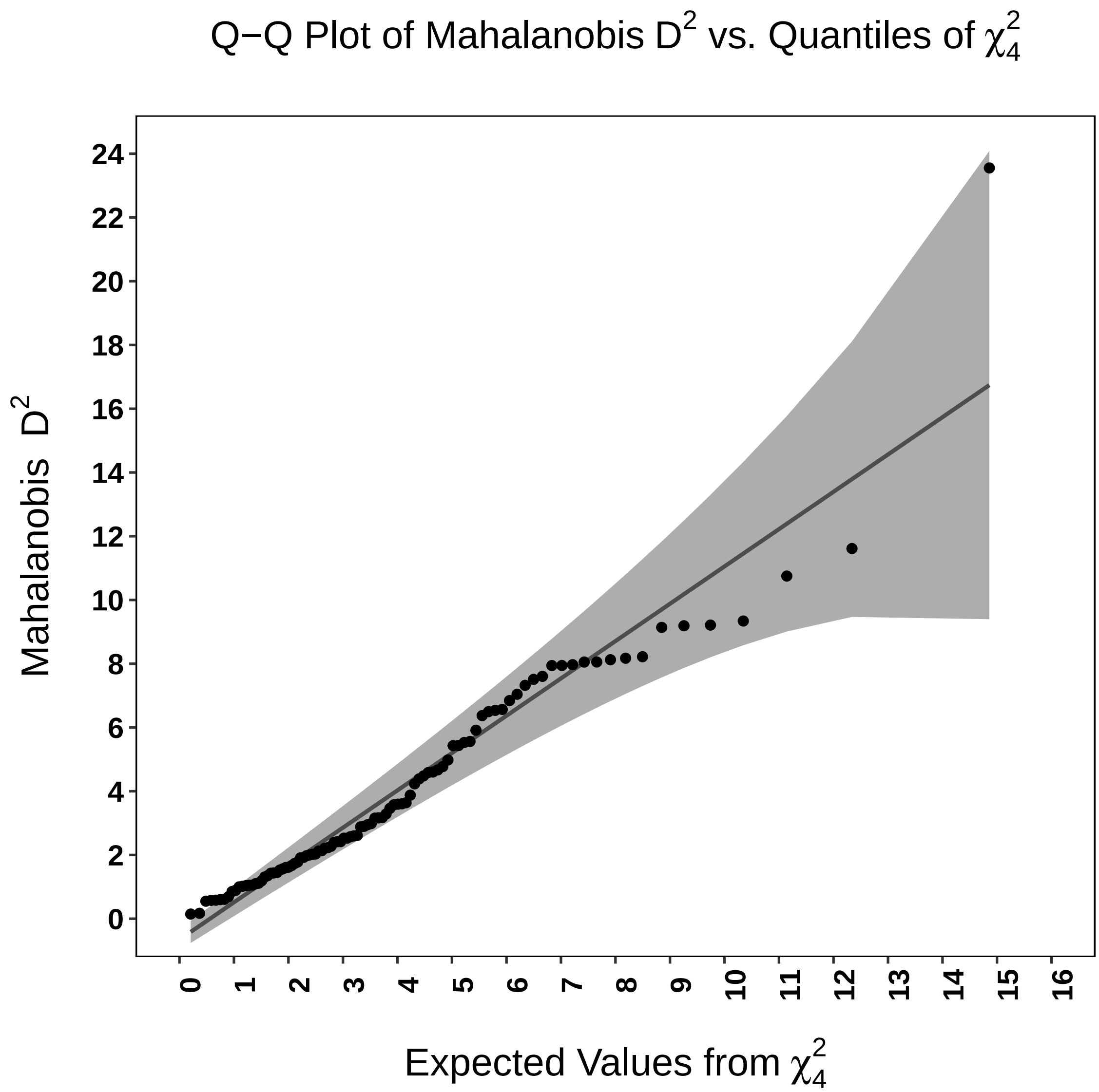
3 Multivariate Normality Assumption

The classification part of discriminant analysis (as well as any significance tests for the discriminant functions) rely on the multivariate normality assumption. Because MANOVA is inherently a discriminant analysis, we make the same assumption. The tests are performed on the residualized data so that group differences do not affect the results. Note that a violation of multivariate normality will also affect the test of homogeneity of covariance matrices.

3.1 Full Sample

```r
mvn(Skills[, 9:12], mvnTest = "mardia")

## $multivariateNormality
##              Test        Statistic            p value Result
## 1 Mardia Skewness  33.714208460761 0.0281242141329968     NO
## 2 Mardia Kurtosis 2.67851447156136 0.0073949536550868     NO
## 3             MVN             <NA>               <NA>     NO
##
## $univariateNormality
##           Test   Variable Statistic   p value Normality
## 1 Shapiro-Wilk P_Verbal_R    0.9857    0.3545    YES
## 2 Shapiro-Wilk P_Quant_R     0.9843    0.2825    YES
## 3 Shapiro-Wilk C_Verbal_R    0.9881    0.5174    YES
## 4 Shapiro-Wilk C_Quant_R     0.9867    0.4203    YES
##
## $Descriptives
##              n       Mean Std.Dev  Median    Min   Max   25th
## P_Verbal_R 100 -2.821e-16  10.248 -1.2840 -34.55 26.09 -5.576
## P_Quant_R  100 -6.008e-17   9.626 -1.0052 -31.45 24.02 -5.440
## C_Verbal_R 100  2.696e-16   9.927  1.2883 -24.66 25.50 -7.084
## C_Quant_R  100 -1.776e-16   9.279 -0.8129 -26.11 20.75 -6.727
##             75th     Skew Kurtosis
## P_Verbal_R 7.550 -0.04563   0.5699
## P_Quant_R  6.135 -0.23775   0.6170
## C_Verbal_R 6.623 -0.15956  -0.2856
## C_Quant_R  7.167 -0.09505  -0.2942
```

```r
CV <- cov(Skills[, 9:12])
D2_1 <- mahalanobis(Skills[, 9:12], center = colMeans(Skills[, 9:12]),
    cov = CV)
D2_1 <- as.data.frame(D2_1)
ggplot(D2_1, aes(sample = D2_1)) + stat_qq_band(distribution = "chisq",
    dparams = list(df = 4)) + stat_qq_line(distribution = "chisq",
    dparams = list(df = 4)) + stat_qq(distribution = "qchisq", dparams = list(df = 4)) +
    scale_y_continuous(breaks = seq(0, 24, 2)) + scale_x_continuous(breaks = seq(0,
    16, 1)) + coord_cartesian(xlim = c(0, 16), ylim = c(0, 24)) +
    xlab(expression("Expected Values from" * ~chi[4]^2)) + ylab(expression("Mahalanobis " *
    ~D^2)) + theme(text = element_text(size = 14, family = "sans",
    color = "black", face = "bold"), axis.text.y = element_text(colour = "black",
    size = 12, face = "bold"), axis.text.x = element_text(colour = "black",
    size = 12, face = "bold", angle = 90), axis.title.x = element_text(margin = margin(15,
    0, 0, 0), size = 16), axis.title.y = element_text(margin = margin(0,
    15, 0, 0), size = 16), axis.line.x = element_blank(), axis.line.y = element_blank(),
    plot.title = element_text(size = 16, face = "bold", margin = margin(0,
        0, 20, 0), hjust = 0.5), panel.background = element_rect(fill = "white",
        linetype = 1, color = "black"), panel.grid.major = element_blank(),
    panel.grid.minor = element_blank(), plot.background = element_rect(fill = "white"),
    plot.margin = unit(c(1, 1, 1, 1), "cm"), legend.position = "bottom",
    legend.title = element_blank()) + ggtitle(expression("Q-Q Plot of Mahalanobis" *
    ~D^2 * " vs. Quantiles of" * ~chi[4]^2))
```

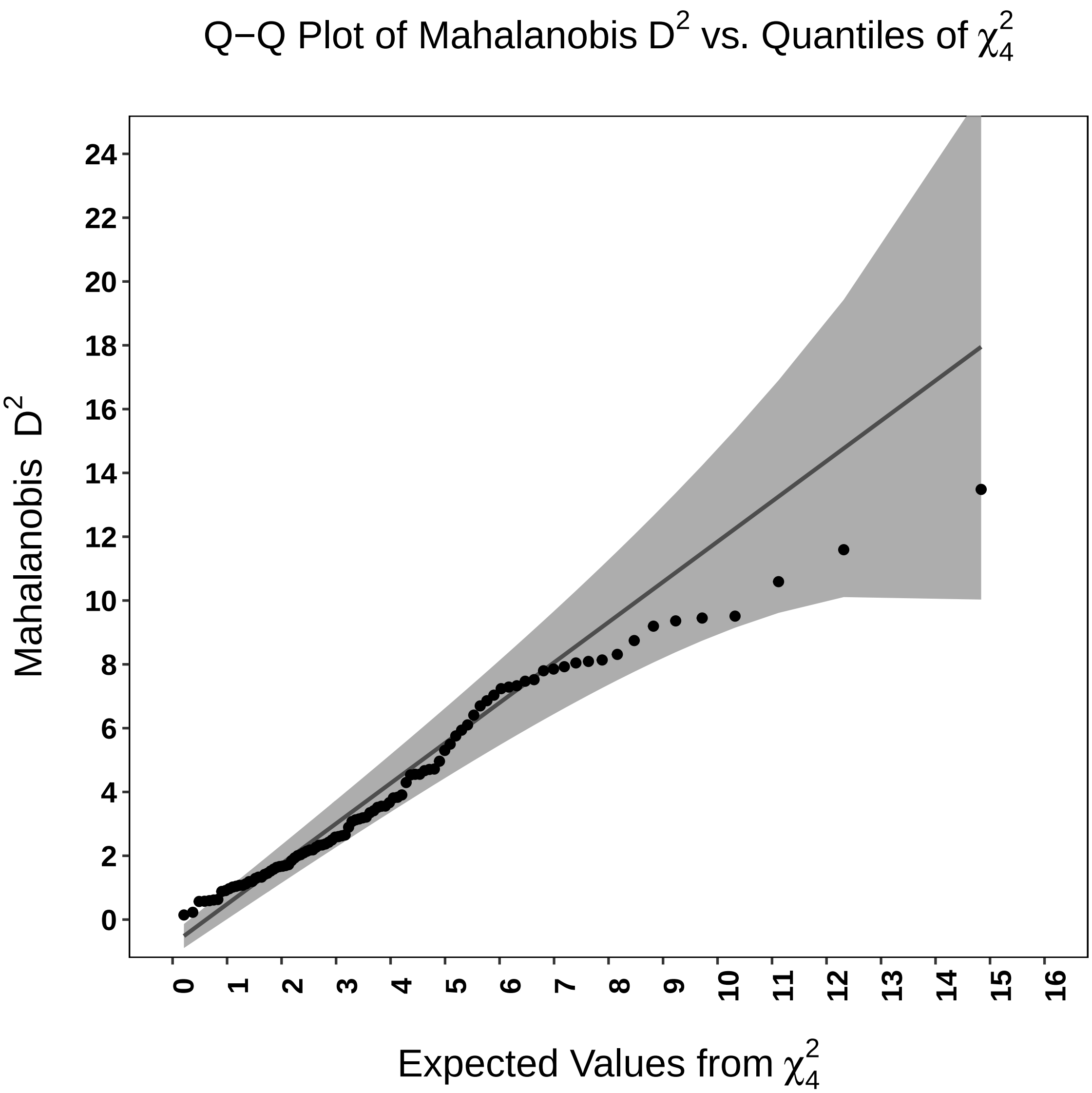


3.2 Outlier Excluded

```
Skills$D2_1 <- D2_1
Skills_Trimmed <- Skills[which(Skills$D2_1 != max(Skills$D2_1)), ]

mvn(Skills_Trimmed[, 9:12], mvnTest = "mardia")

## $multivariateNormality
##              Test          Statistic           p value Result
## 1 Mardia Skewness  20.2783280472259 0.440644455966184    YES
## 2 Mardia Kurtosis 0.600058653103908 0.548467146873458    YES
## 3             MVN              <NA>              <NA>    YES
##
```

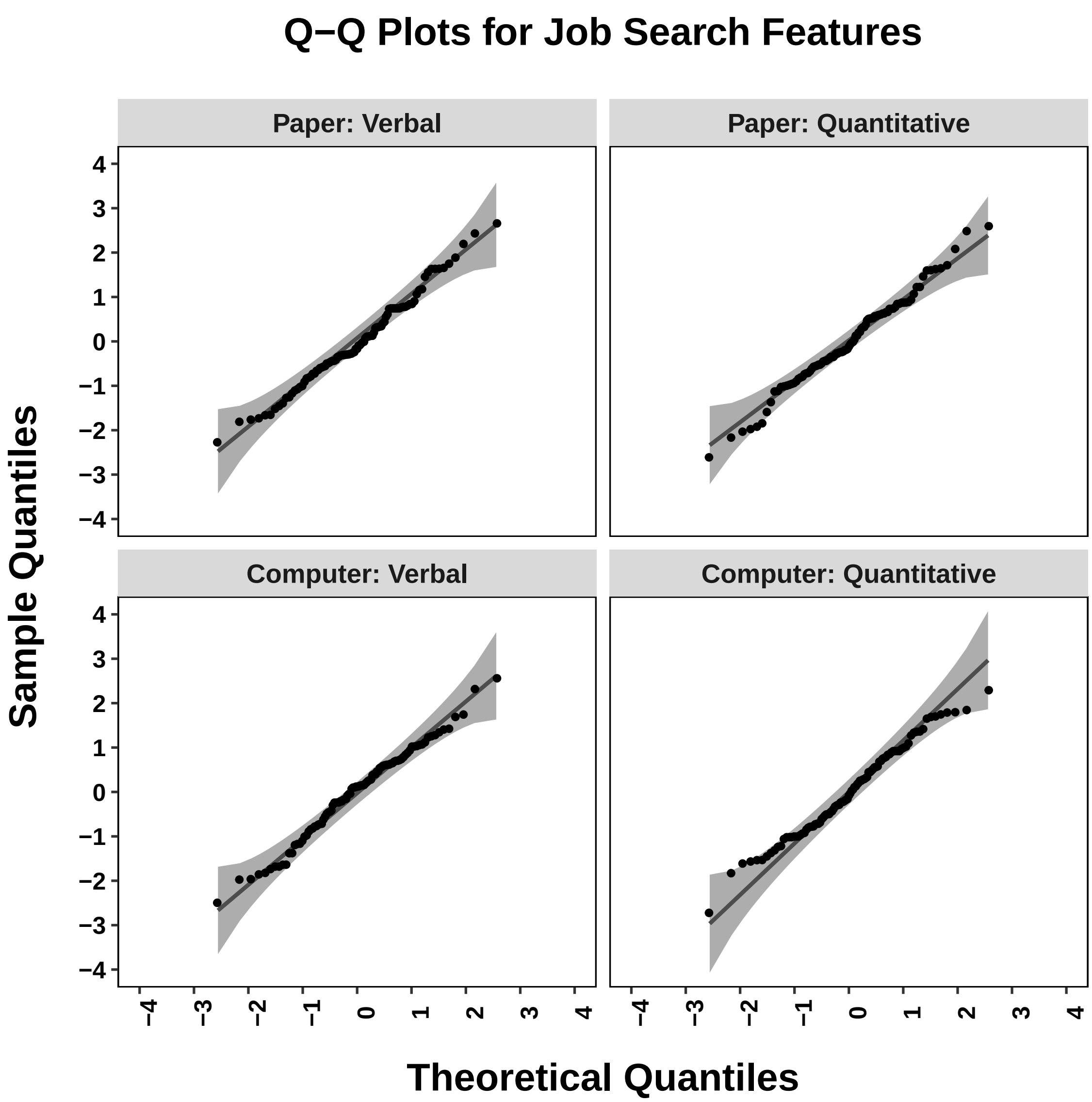
```
## $univariateNormality
##           Test   Variable Statistic   p value Normality
## 1 Shapiro-Wilk P_Verbal_R    0.9857    0.3630    YES
## 2 Shapiro-Wilk P_Quant_R     0.9889    0.5820    YES
## 3 Shapiro-Wilk C_Verbal_R    0.9872    0.4598    YES
## 4 Shapiro-Wilk C_Quant_R     0.9862    0.3923    YES
##
## $Descriptives
##             n   Mean Std.Dev  Median    Min   Max   25th  75th
## P_Verbal_R 99 0.3490   9.684 -1.2790 -21.66 26.09 -5.405 7.550
## P_Quant_R  99 0.3177   9.133 -0.7615 -23.53 24.02 -5.114 6.200
## C_Verbal_R 99 0.1044   9.922  1.3085 -24.66 25.50 -7.033 6.710
## C_Quant_R  99 0.2638   8.941 -0.4738 -24.09 20.75 -6.692 7.215
##                Skew Kurtosis
## P_Verbal_R  0.29712 -0.16599
## P_Quant_R   0.02811  0.07662
## C_Verbal_R -0.18123 -0.25780
## C_Quant_R   0.05667 -0.57682
```

```
CV <- cov(Skills_Trimmed[, 9:12])
D2_1 <- mahalanobis(Skills_Trimmed[, 9:12], center = colMeans(Skills_Trimmed[,
    9:12]), cov = CV)
D2_1 <- as.data.frame(D2_1)
ggplot(D2_1, aes(sample = D2_1)) + stat_qq_band(distribution = "chisq",
    dparams = list(df = 4)) + stat_qq_line(distribution = "chisq",
    dparams = list(df = 4)) + stat_qq(distribution = "qchisq", dparams = list(df = 4)) +
    scale_y_continuous(breaks = seq(0, 24, 2)) + scale_x_continuous(breaks = seq(0,
    16, 1)) + coord_cartesian(xlim = c(0, 16), ylim = c(0, 24)) +
    xlab(expression("Expected Values from" * ~chi[4]^2)) + ylab(expression("Mahalanobis " *
    ~D^2)) + theme(text = element_text(size = 14, family = "sans",
    color = "black", face = "bold"), axis.text.y = element_text(colour = "black",
    size = 12, face = "bold"), axis.text.x = element_text(colour = "black",
    size = 12, face = "bold", angle = 90), axis.title.x = element_text(margin = margin(15,
    0, 0, 0), size = 16), axis.title.y = element_text(margin = margin(0,
    15, 0, 0), size = 16), axis.line.x = element_blank(), axis.line.y = element_blank(),
    plot.title = element_text(size = 16, face = "bold", margin = margin(0,
        0, 20, 0), hjust = 0.5), panel.background = element_rect(fill = "white",
        linetype = 1, color = "black"), panel.grid.major = element_blank(),
    panel.grid.minor = element_blank(), plot.background = element_rect(fill = "white"),
    plot.margin = unit(c(1, 1, 1, 1), "cm"), legend.position = "bottom",
    legend.title = element_blank()) + ggtitle(expression("Q-Q Plot of Mahalanobis" *
    ~D^2 * " vs. Quantiles of" * ~chi[4]^2))
```




```
Skills_Trimmed_QQ <- scale(Skills_Trimmed[, 9:12])
Data_long <- melt(Skills_Trimmed_QQ)
Data_long <- as.data.frame(Data_long)
names(Data_long) <- c("Index", "feature", "value")
Data_long$feature_F <- factor(Data_long$feature, levels = c("P_Verbal_R",
    "P_Quant_R", "C_Verbal_R", "C_Quant_R"), labels = c("Paper: Verbal",
    "Paper: Quantitative", "Computer: Verbal", "Computer: Quantitative"))
p <- ggplot(Data_long, aes(sample = value)) + stat_qq_band() + stat_qq_line() +
    stat_qq(distribution = qnorm, size = 1) + scale_y_continuous(breaks = seq(-4,
    4, 1)) + scale_x_continuous(breaks = seq(-4, 4, 1)) + coord_cartesian(xlim = c(-4,
    4), ylim = c(-4, 4)) + xlab("Theoretical Quantiles") + ylab("Sample Quantiles") +
    theme(text = element_text(size = 14, family = "sans", color = "black",
        face = "bold"), axis.text.y = element_text(colour = "black",
        size = 10, face = "bold"), axis.text.x = element_text(colour = "black",
```

```
        size = 10, face = "bold", angle = 90), axis.title.x = element_text(margin = margin(15,
        0, 0, 0), size = 16), axis.title.y = element_text(margin = margin(0,
        15, 0, 0), size = 16), axis.line.x = element_blank(), axis.line.y = element_blank(),
        plot.title = element_text(size = 16, face = "bold", margin = margin(0,
            0, 20, 0), hjust = 0.5), panel.background = element_rect(fill = "white",
            linetype = 1, color = "black"), panel.grid.major = element_blank(),
        panel.grid.minor = element_blank(), plot.background = element_rect(fill = "white"),
        plot.margin = unit(c(1, 1, 1, 1), "cm"), legend.position = "bottom",
        legend.title = element_blank()) + ggtitle("Q-Q Plots for Job Search Features")
p + facet_wrap(~feature_F)
```



4 Homogeneity Assumption

We assume in discriminant analysis that the separate group variance-covariance matrices are homogeneous. Box's test can be used to test this assumption. Note, however, that it is also sensitive to violations of multivariate normality.

```
boxM(Skills[, 2:5], Skills$Group)

##
##  Box's M-test for Homogeneity of Covariance Matrices
##
## data:  Skills[, 2:5]
## Chi-Sq (approx.) = 93, df = 30, p-value = 2e-08

boxM(Skills[, 2:5], Skills$Group)$cov

## $`1`
##          P_Verbal P_Quant C_Verbal C_Quant
## P_Verbal   112.10   94.54    67.22   46.25
## P_Quant     94.54   99.71    82.84   70.41
## C_Verbal    67.22   82.84   117.58  104.82
## C_Quant     46.25   70.41   104.82  114.46
##
## $`2`
##          P_Verbal P_Quant C_Verbal C_Quant
## P_Verbal   164.89  121.83    90.29  121.84
## P_Quant    121.83  118.23    52.23   89.24
## C_Verbal    90.29   52.23   105.61   95.50
## C_Quant    121.84   89.24    95.50  120.37
##
## $`3`
##          P_Verbal P_Quant C_Verbal C_Quant
## P_Verbal   122.97   77.14    60.02   34.16
## P_Quant     77.14   87.49    76.05   46.13
## C_Verbal    60.02   76.05   106.12   72.31
## C_Quant     34.16   46.13    72.31   81.09
##
## $`4`
##          P_Verbal P_Quant C_Verbal C_Quant
## P_Verbal    33.24   37.85    33.51   22.40
## P_Quant     37.85   76.80    68.76   37.15
## C_Verbal    33.51   68.76    77.15   42.04
## C_Quant     22.40   37.15    42.04   39.21

boxM(Skills[, 2:5], Skills$Group)$pooled

##          P_Verbal P_Quant C_Verbal C_Quant
## P_Verbal   108.30   82.84    62.76   56.16
## P_Quant     82.84   95.56    69.97   60.74
## C_Verbal    62.76   69.97   101.62   78.67
## C_Quant     56.16   60.74    78.67   88.79

boxM(Skills_Trimmed[, 2:5], Skills_Trimmed$Group)

##
##  Box's M-test for Homogeneity of Covariance Matrices
```

```
##
## data:  Skills_Trimmed[, 2:5]
## Chi-Sq (approx.) = 81, df = 30, p-value = 0.000002
```

```r
boxM(Skills_Trimmed[, 2:5], Skills_Trimmed$Group)$cov
```

```
## $`1`
##          P_Verbal P_Quant C_Verbal C_Quant
## P_Verbal   112.10   94.54    67.22   46.25
## P_Quant     94.54   99.71    82.84   70.41
## C_Verbal    67.22   82.84   117.58  104.82
## C_Quant     46.25   70.41   104.82  114.46
##
## $`2`
##          P_Verbal P_Quant C_Verbal C_Quant
## P_Verbal   117.98   77.91    78.04   86.27
## P_Quant     77.91   78.56    39.78   55.92
## C_Verbal    78.04   39.78   105.37   87.43
## C_Quant     86.27   55.92    87.43   94.73
##
## $`3`
##          P_Verbal P_Quant C_Verbal C_Quant
## P_Verbal   122.97   77.14    60.02   34.16
## P_Quant     77.14   87.49    76.05   46.13
## C_Verbal    60.02   76.05   106.12   72.31
## C_Quant     34.16   46.13    72.31   81.09
##
## $`4`
##          P_Verbal P_Quant C_Verbal C_Quant
## P_Verbal    33.24   37.85    33.51   22.40
## P_Quant     37.85   76.80    68.76   37.15
## C_Verbal    33.51   68.76    77.15   42.04
## C_Quant     22.40   37.15    42.04   39.21
```

```r
boxM(Skills_Trimmed[, 2:5], Skills_Trimmed$Group)$pooled
```

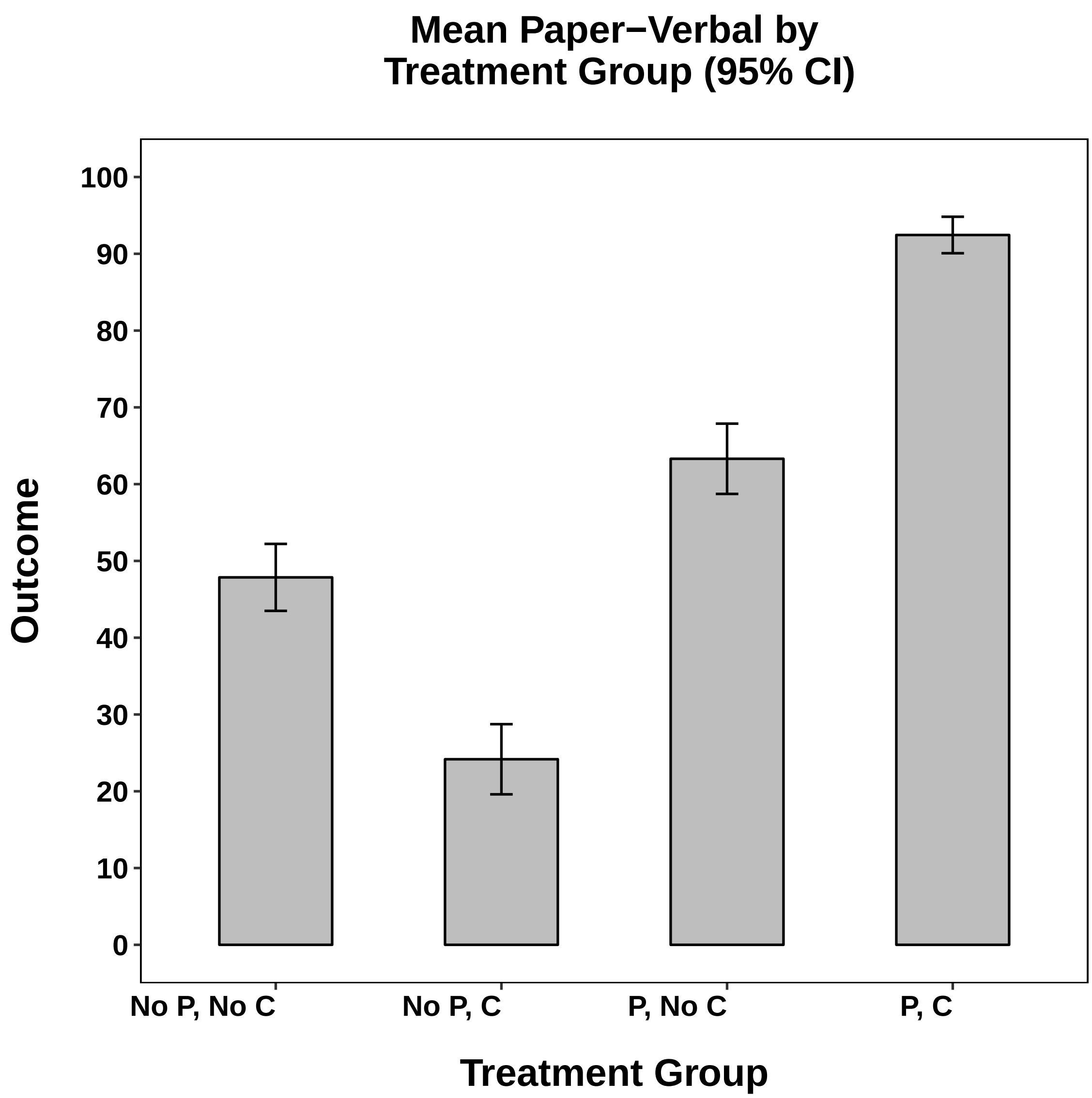
```
##          P_Verbal P_Quant C_Verbal C_Quant
## P_Verbal    96.35   71.80    59.51   46.86
## P_Quant     71.80   85.71    67.14   52.37
## C_Verbal    59.51   67.14   101.52   76.54
## C_Quant     46.86   52.37    76.54   82.24
```

5 Means and Confidence Intervals

Displayed here are bar graphs of the condition means with 95% confidence intervals.

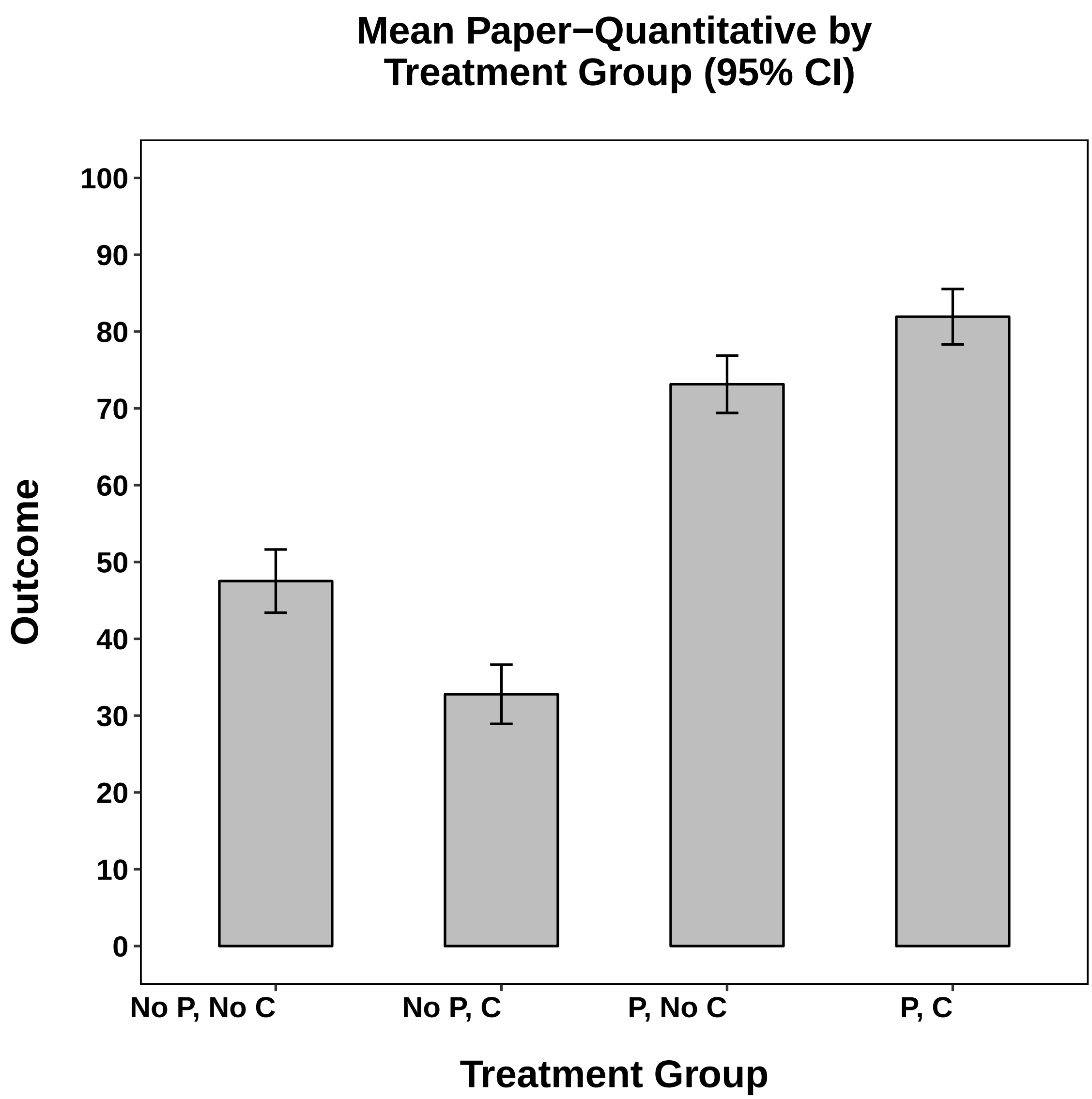
```
D <- describeBy(Skills_Trimmed[, 2:5], group = Skills_Trimmed$Group4)

plot_data <- matrix(NA, nrow = 4, ncol = 8)

for (i in 1:4) {
    for (j in 1:4) {
        plot_data[i, j] <- D[[i]]$mean[j]
        plot_data[i, j + 4] <- qt(0.975, D[[i]]$n[j]) * D[[i]]$sd[j]/sqrt(D[[i]]$n[j])
    }
}
plot_data <- as.data.frame(plot_data)
names(plot_data) <- c("PV_mean", "PQ_mean", "CV_mean", "CQ_mean",
    "PV_CI", "PQ_CI", "CV_CI", "CQ_CI")
plot_data$Group3 <- factor(c("No P, No C", "No P, C", "P, No C", "P, C"))
plot_data$Group4 <- factor(plot_data$Group3, levels = c("No P, No C",
    "No P, C", "P, No C", "P, C"), labels = c("No P, No C", "No P, C",
    "P, No C", "P, C"))
```

```
p1 <- ggplot(plot_data, aes(x = as.factor(Group4), y = PV_mean)) +
    geom_bar(position = position_dodge(), stat = "identity", color = "black",
        width = 0.5, fill = "grey") + geom_errorbar(aes(ymin = PV_mean -
    PV_CI, ymax = PV_mean + PV_CI), width = 0.1, position = position_dodge(0.5)) +
    scale_y_continuous(breaks = c(seq(0, 100, 10))) + coord_cartesian(ylim = c(0,
    100)) + xlab("Treatment Group") + ylab("Outcome") + theme(text = element_text(size = 14,
    family = "sans", color = "black", face = "bold"), axis.text.y = element_text(colour = "black",
    size = 12, face = "bold"), axis.text.x = element_text(colour = "black",
    size = 12, face = "bold", angle = 0, hjust = 1), axis.title.x = element_text(margin = margin(15,
    0, 0, 0), size = 16), axis.title.y = element_text(margin = margin(0,
    15, 0, 0), size = 16), axis.line.x = element_blank(), axis.line.y = element_blank(),
    plot.title = element_text(size = 16, face = "bold", margin = margin(0,
        0, 20, 0), hjust = 0.5), panel.background = element_rect(fill = "white",
        linetype = 1, color = "black"), panel.grid.major = element_blank(),
    panel.grid.minor = element_blank(), panel.border = element_rect(fill = NA,
        size = 0.5), plot.background = element_rect(fill = "white"),
    plot.margin = unit(c(1, 1, 1, 1), "cm"), legend.position = "bottom",
    legend.title = element_blank()) + ggtitle("Mean Paper-Verbal by\n Treatment Group (95% CI)")
print(p1)
```



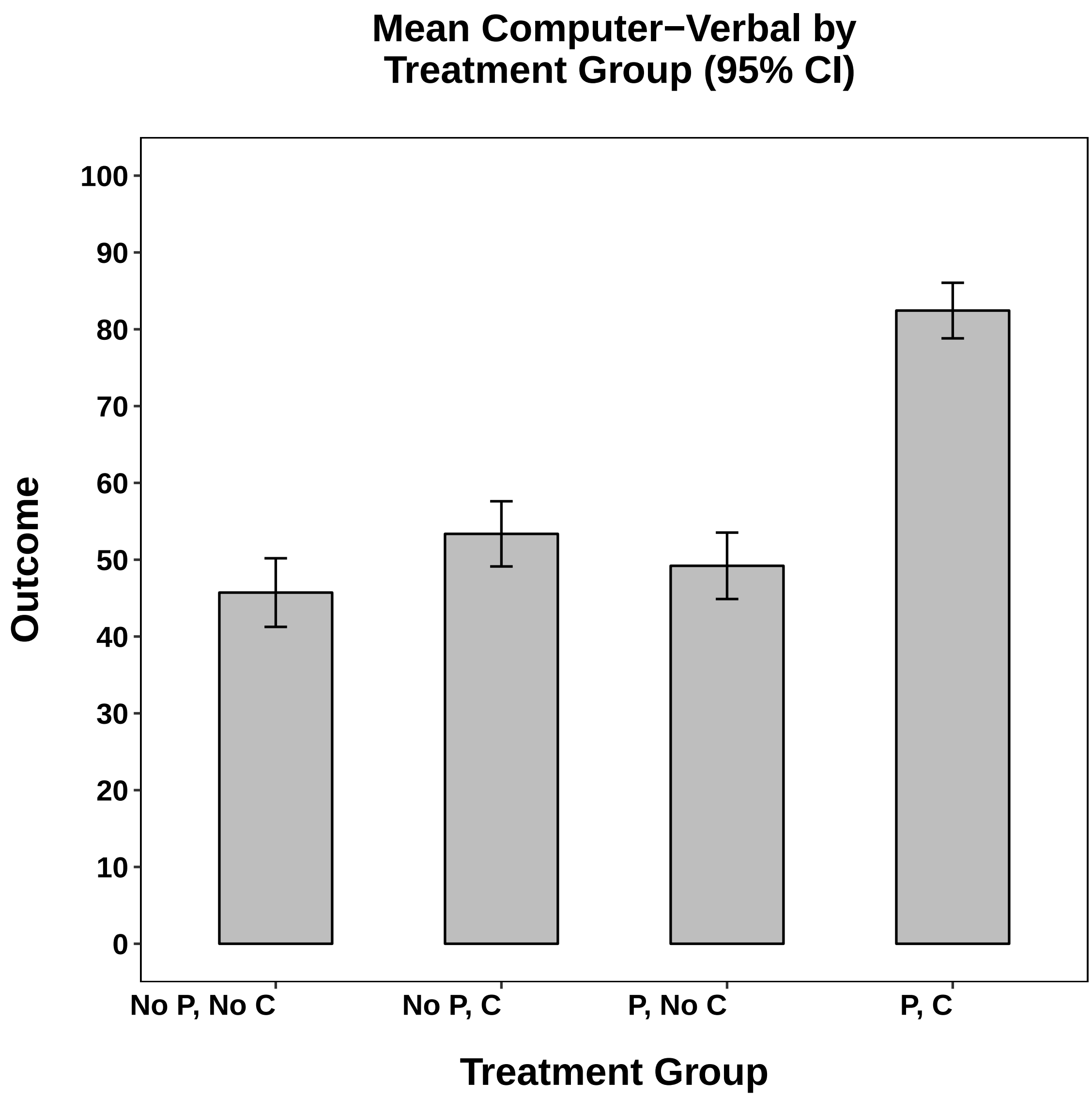

```
p2 <- ggplot(plot_data, aes(x = as.factor(Group4), y = PQ_mean)) +
    geom_bar(position = position_dodge(), stat = "identity", color = "black",
        width = 0.5, fill = "grey") + geom_errorbar(aes(ymin = PQ_mean -
    PQ_CI, ymax = PQ_mean + PQ_CI), width = 0.1, position = position_dodge(0.5)) +
    scale_y_continuous(breaks = c(seq(0, 100, 10))) + coord_cartesian(ylim = c(0,
    100)) + xlab("Treatment Group") + ylab("Outcome") + theme(text = element_text(size = 14,
    family = "sans", color = "black", face = "bold"), axis.text.y = element_text(colour = "black",
    size = 12, face = "bold"), axis.text.x = element_text(colour = "black",
    size = 12, face = "bold", angle = 0, hjust = 1), axis.title.x = element_text(margin = margin(15,
    0, 0, 0), size = 16), axis.title.y = element_text(margin = margin(0,
    15, 0, 0), size = 16), axis.line.x = element_blank(), axis.line.y = element_blank(),
    plot.title = element_text(size = 16, face = "bold", margin = margin(0,
        0, 20, 0), hjust = 0.5), panel.background = element_rect(fill = "white",
        linetype = 1, color = "black"), panel.grid.major = element_blank(),
```

```
    panel.grid.minor = element_blank(), panel.border = element_rect(fill = NA,
        size = 0.5), plot.background = element_rect(fill = "white"),
    plot.margin = unit(c(1, 1, 1, 1), "cm"), legend.position = "bottom",
    legend.title = element_blank()) + ggtitle("Mean Paper-Quantitative by\n Treatment Group (95% CI)")
print(p2)
```

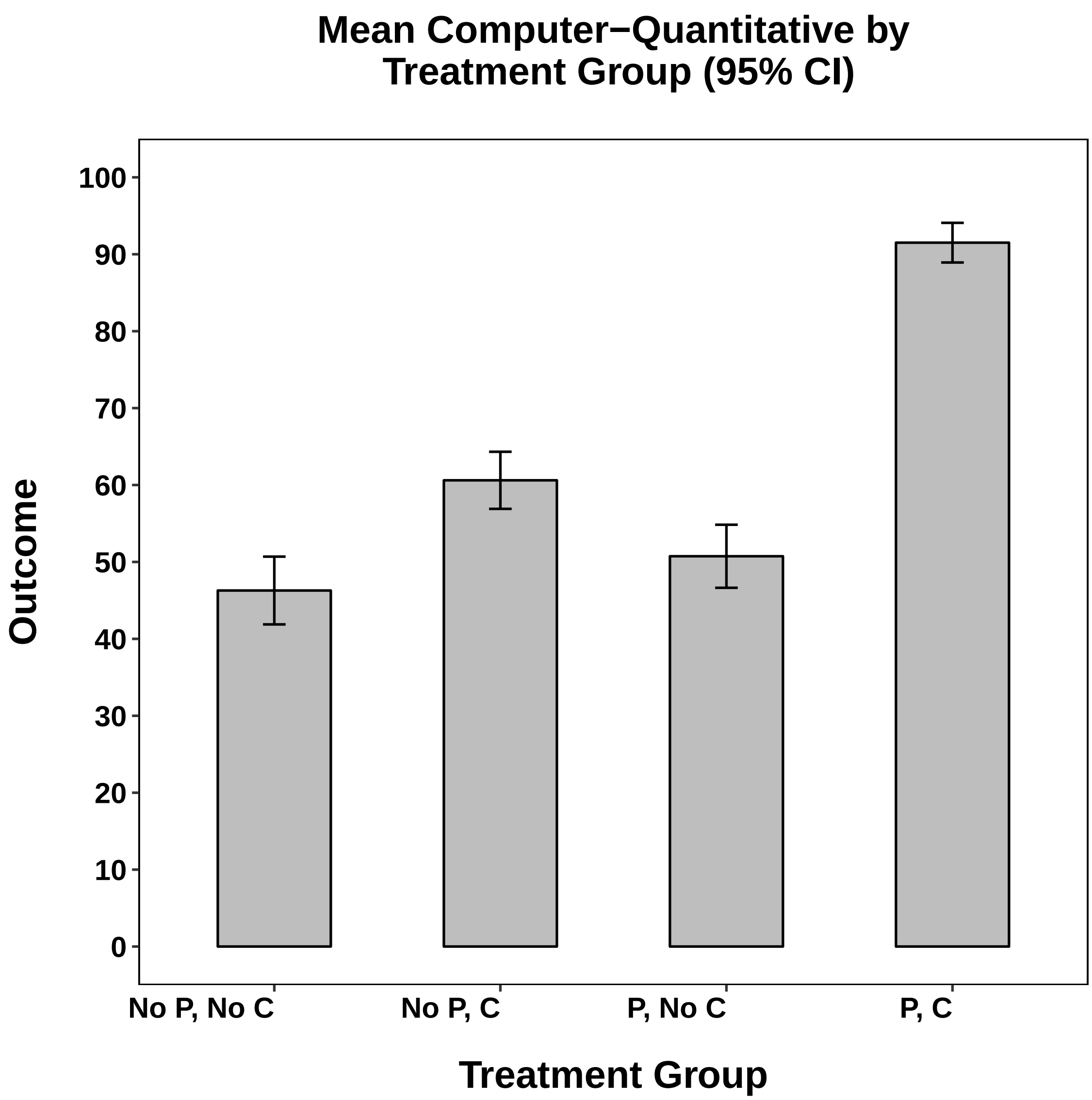



```
p3 <- ggplot(plot_data, aes(x = as.factor(Group4), y = CV_mean)) +
    geom_bar(position = position_dodge(), stat = "identity", color = "black",
        width = 0.5, fill = "grey") + geom_errorbar(aes(ymin = CV_mean -
    CV_CI, ymax = CV_mean + CV_CI), width = 0.1, position = position_dodge(0.5)) +
    scale_y_continuous(breaks = c(seq(0, 100, 10))) + coord_cartesian(ylim = c(0,
    100)) + xlab("Treatment Group") + ylab("Outcome") + theme(text = element_text(size = 14,
    family = "sans", color = "black", face = "bold"), axis.text.y = element_text(colour = "black",
```

```
    size = 12, face = "bold"), axis.text.x = element_text(colour = "black",
    size = 12, face = "bold", angle = 0, hjust = 1), axis.title.x = element_text(margin = margin(15,
    0, 0, 0), size = 16), axis.title.y = element_text(margin = margin(0,
    15, 0, 0), size = 16), axis.line.x = element_blank(), axis.line.y = element_blank(),
    plot.title = element_text(size = 16, face = "bold", margin = margin(0,
        0, 20, 0), hjust = 0.5), panel.background = element_rect(fill = "white",
        linetype = 1, color = "black"), panel.grid.major = element_blank(),
    panel.grid.minor = element_blank(), panel.border = element_rect(fill = NA,
        size = 0.5), plot.background = element_rect(fill = "white"),
    plot.margin = unit(c(1, 1, 1, 1), "cm"), legend.position = "bottom",
    legend.title = element_blank()) + ggtitle("Mean Computer-Verbal by\n Treatment Group (95% CI)")
print(p3)
```



```r
p4 <- ggplot(plot_data, aes(x = as.factor(Group4), y = CQ_mean)) +
    geom_bar(position = position_dodge(), stat = "identity", color = "black",
        width = 0.5, fill = "grey") + geom_errorbar(aes(ymin = CQ_mean -
    CQ_CI, ymax = CQ_mean + CQ_CI), width = 0.1, position = position_dodge(0.5)) +
    scale_y_continuous(breaks = c(seq(0, 100, 10))) + coord_cartesian(ylim = c(0,
    100)) + xlab("Treatment Group") + ylab("Outcome") + theme(text = element_text(size = 14,
    family = "sans", color = "black", face = "bold"), axis.text.y = element_text(colour = "black",
    size = 12, face = "bold"), axis.text.x = element_text(colour = "black",
    size = 12, face = "bold", angle = 0, hjust = 1), axis.title.x = element_text(margin = margin(15,
    0, 0, 0), size = 16), axis.title.y = element_text(margin = margin(0,
    15, 0, 0), size = 16), axis.line.x = element_blank(), axis.line.y = element_blank(),
    plot.title = element_text(size = 16, face = "bold", margin = margin(0,
        0, 20, 0), hjust = 0.5), panel.background = element_rect(fill = "white",
        linetype = 1, color = "black"), panel.grid.major = element_blank(),
    panel.grid.minor = element_blank(), panel.border = element_rect(fill = NA,
        size = 0.5), plot.background = element_rect(fill = "white"),
    plot.margin = unit(c(1, 1, 1, 1), "cm"), legend.position = "bottom",
    legend.title = element_blank()) + ggtitle("Mean Computer-Quantitative by\n Treatment Group (95% CI)'
print(p4)
```




```
p1 <- ggplot(plot_data, aes(x = as.factor(Group4), y = PV_mean)) +
    geom_bar(position = position_dodge(), stat = "identity", color = "black",
        width = 0.5, fill = "grey") + geom_errorbar(aes(ymin = PV_mean -
    PV_CI, ymax = PV_mean + PV_CI), width = 0.1, position = position_dodge(0.5)) +
    scale_y_continuous(breaks = c(seq(0, 100, 20))) + coord_cartesian(ylim = c(0,
    100)) + xlab("Treatment Group") + ylab("Outcome") + theme(text = element_text(size = 14,
    family = "sans", color = "black", face = "bold"), axis.text.y = element_text(colour = "black",
    size = 10, face = "bold"), axis.text.x = element_text(colour = "black",
    size = 10, face = "bold", angle = 45, hjust = 1), axis.title.x = element_text(margin = margin(5,
    0, 0, 0), size = 12), axis.title.y = element_text(margin = margin(0,
    15, 0, 0), size = 12), axis.line.x = element_blank(), axis.line.y = element_blank(),
    plot.title = element_text(size = 14, face = "bold", margin = margin(0,
        0, 5, 0), hjust = 0.5), panel.background = element_rect(fill = "white",
        linetype = 1, color = "black"), panel.grid.major = element_blank(),
```

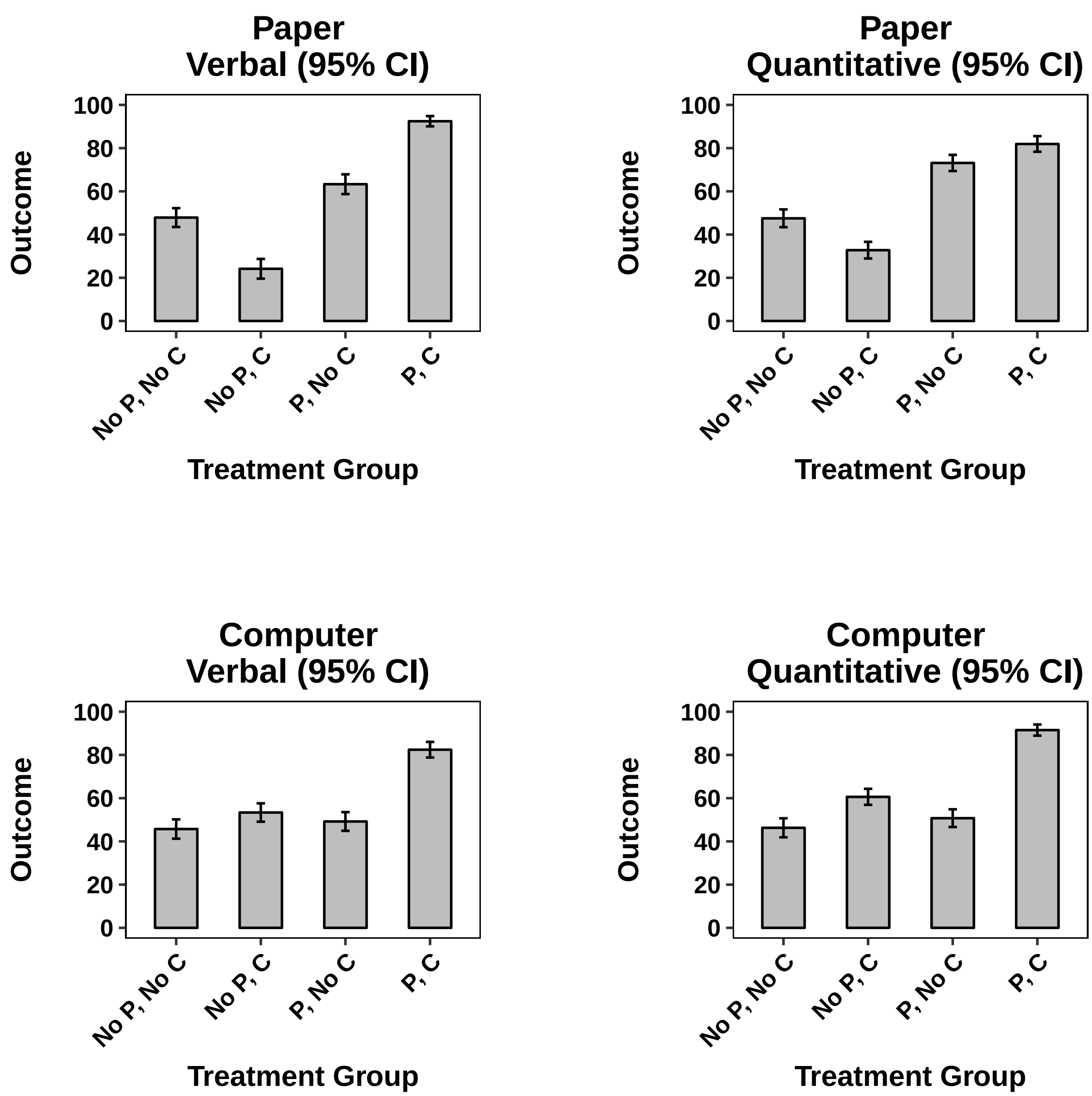
```
    panel.grid.minor = element_blank(), panel.border = element_rect(fill = NA,
        size = 0.5), plot.background = element_rect(fill = "white"),
    plot.margin = unit(c(1, 1, 1, 1), "cm"), legend.position = "bottom",
    legend.title = element_blank()) + ggtitle("Paper \n Verbal (95% CI)")
```

```
p2 <- ggplot(plot_data, aes(x = as.factor(Group4), y = PQ_mean)) +
    geom_bar(position = position_dodge(), stat = "identity", color = "black",
        width = 0.5, fill = "grey") + geom_errorbar(aes(ymin = PQ_mean -
    PQ_CI, ymax = PQ_mean + PQ_CI), width = 0.1, position = position_dodge(0.5)) +
    scale_y_continuous(breaks = c(seq(0, 100, 20))) + coord_cartesian(ylim = c(0,
    100)) + xlab("Treatment Group") + ylab("Outcome") + theme(text = element_text(size = 14,
    family = "sans", color = "black", face = "bold"), axis.text.y = element_text(colour = "black",
    size = 10, face = "bold"), axis.text.x = element_text(colour = "black",
    size = 10, face = "bold", angle = 45, hjust = 1), axis.title.x = element_text(margin = margin(5,
    0, 0, 0), size = 12), axis.title.y = element_text(margin = margin(0,
    15, 0, 0), size = 12), axis.line.x = element_blank(), axis.line.y = element_blank(),
    plot.title = element_text(size = 14, face = "bold", margin = margin(0,
        0, 5, 0), hjust = 0.5), panel.background = element_rect(fill = "white",
        linetype = 1, color = "black"), panel.grid.major = element_blank(),
    panel.grid.minor = element_blank(), panel.border = element_rect(fill = NA,
        size = 0.5), plot.background = element_rect(fill = "white"),
    plot.margin = unit(c(1, 1, 1, 1), "cm"), legend.position = "bottom",
    legend.title = element_blank()) + ggtitle("Paper \n Quantitative (95% CI)")
```

```
p3 <- ggplot(plot_data, aes(x = as.factor(Group4), y = CV_mean)) +
    geom_bar(position = position_dodge(), stat = "identity", color = "black",
        width = 0.5, fill = "grey") + geom_errorbar(aes(ymin = CV_mean -
    CV_CI, ymax = CV_mean + CV_CI), width = 0.1, position = position_dodge(0.5)) +
    scale_y_continuous(breaks = c(seq(0, 100, 20))) + coord_cartesian(ylim = c(0,
    100)) + xlab("Treatment Group") + ylab("Outcome") + theme(text = element_text(size = 14,
    family = "sans", color = "black", face = "bold"), axis.text.y = element_text(colour = "black",
    size = 10, face = "bold"), axis.text.x = element_text(colour = "black",
    size = 10, face = "bold", angle = 45, hjust = 1), axis.title.x = element_text(margin = margin(5,
    0, 0, 0), size = 12), axis.title.y = element_text(margin = margin(0,
    15, 0, 0), size = 12), axis.line.x = element_blank(), axis.line.y = element_blank(),
    plot.title = element_text(size = 14, face = "bold", margin = margin(0,
        0, 5, 0), hjust = 0.5), panel.background = element_rect(fill = "white",
        linetype = 1, color = "black"), panel.grid.major = element_blank(),
    panel.grid.minor = element_blank(), panel.border = element_rect(fill = NA,
        size = 0.5), plot.background = element_rect(fill = "white"),
    plot.margin = unit(c(1, 1, 1, 1), "cm"), legend.position = "bottom",
    legend.title = element_blank()) + ggtitle("Computer \n Verbal (95% CI)")
```

```
p4 <- ggplot(plot_data, aes(x = as.factor(Group4), y = CQ_mean)) +
    geom_bar(position = position_dodge(), stat = "identity", color = "black",
        width = 0.5, fill = "grey") + geom_errorbar(aes(ymin = CQ_mean -
    CQ_CI, ymax = CQ_mean + CQ_CI), width = 0.1, position = position_dodge(0.5)) +
    scale_y_continuous(breaks = c(seq(0, 100, 20))) + coord_cartesian(ylim = c(0,
    100)) + xlab("Treatment Group") + ylab("Outcome") + theme(text = element_text(size = 14,
    family = "sans", color = "black", face = "bold"), axis.text.y = element_text(colour = "black",
    size = 10, face = "bold"), axis.text.x = element_text(colour = "black",
```

```
    size = 10, face = "bold", angle = 45, hjust = 1), axis.title.x = element_text(margin = margin(5,
    0, 0, 0), size = 12), axis.title.y = element_text(margin = margin(0,
    15, 0, 0), size = 12), axis.line.x = element_blank(), axis.line.y = element_blank(),
    plot.title = element_text(size = 14, face = "bold", margin = margin(0,
        0, 5, 0), hjust = 0.5), panel.background = element_rect(fill = "white",
        linetype = 1, color = "black"), panel.grid.major = element_blank(),
    panel.grid.minor = element_blank(), panel.border = element_rect(fill = NA,
        size = 0.5), plot.background = element_rect(fill = "white"),
    plot.margin = unit(c(1, 1, 1, 1), "cm"), legend.position = "bottom",
    legend.title = element_blank()) + ggtitle("Computer \n Quantitative (95% CI)")
grid.arrange(p1, p2, p3, p4, nrow = 2)
```



6 Within-Groups Correlations

Coorelations among the residuals indicate the amount of redundancy among the variables.

```
cor(Skills_Trimmed[, 9:12])

##            P_Verbal_R P_Quant_R C_Verbal_R C_Quant_R
## P_Verbal_R     1.0000    0.7909     0.6016    0.5279
## P_Quant_R      0.7909    1.0000     0.7195    0.6249
## C_Verbal_R     0.6016    0.7195     1.0000    0.8373
## C_Quant_R      0.5279    0.6249     0.8373    1.0000
```

7 ANOVA of Each Outcome
No Group Structure

Simple one-way ANOVAs can determine if each measure can distinguish the groups.

```
AOV_1 <- aov(P_Verbal ~ as.factor(Group), data = Skills_Trimmed)
summary(AOV_1)
```

```
##                  Df Sum Sq Mean Sq F value Pr(>F)
## as.factor(Group)  3  61407   20469     212 <2e-16
## Residuals        95   9153      96
```

```
TukeyHSD(AOV_1)
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = P_Verbal ~ as.factor(Group), data = Skills_Trimmed)
##
## $`as.factor(Group)`
##       diff     lwr    upr p adj
## 2-1  15.45   8.112  22.78     0
## 3-1 -23.69 -30.947 -16.43     0
## 4-1  44.59  37.334  51.85     0
## 3-2 -39.13 -46.469 -31.80     0
## 4-2  29.15  21.812  36.48     0
## 4-3  68.28  61.021  75.54     0
```

```
AOV_2 <- aov(P_Quant ~ as.factor(Group), data = Skills_Trimmed)
summary(AOV_2)
```

```
##                  Df Sum Sq Mean Sq F value Pr(>F)
## as.factor(Group)  3  38419   12806     149 <2e-16
## Residuals        95   8143      86
```

```
TukeyHSD(AOV_2)
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = P_Quant ~ as.factor(Group), data = Skills_Trimmed)
##
## $`as.factor(Group)`
##        diff     lwr     upr  p adj
## 2-1  25.625  18.706  32.544 0.0000
## 3-1 -14.735 -21.583  -7.888 0.0000
## 4-1  34.415  27.567  41.263 0.0000
## 3-2 -40.361 -47.280 -33.442 0.0000
## 4-2   8.789   1.871  15.708 0.0068
## 4-3  49.150  42.302  55.998 0.0000
```

```
AOV_3 <- aov(C_Verbal ~ as.factor(Group), data = Skills_Trimmed)
summary(AOV_3)
```

```
##                  Df Sum Sq Mean Sq F value Pr(>F)
## as.factor(Group)  3  21084    7028    69.2 <2e-16
## Residuals        95   9644     102
```

```r
TukeyHSD(AOV_3)

##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = C_Verbal ~ as.factor(Group), data = Skills_Trimmed)
##
## $`as.factor(Group)`
##       diff     lwr   upr  p adj
## 2-1  3.484 -4.0452 11.01 0.6220
## 3-1  7.643  0.1904 15.10 0.0422
## 4-1 36.714 29.2612 44.17 0.0000
## 3-2  4.158 -3.3712 11.69 0.4752
## 4-2 33.229 25.6995 40.76 0.0000
## 4-3 29.071 21.6183 36.52 0.0000

AOV_4 <- aov(C_Quant ~ as.factor(Group), data = Skills_Trimmed)
summary(AOV_4)

##                  Df Sum Sq Mean Sq F value Pr(>F)
## as.factor(Group)  3  31016   10339     126 <2e-16
## Residuals        95   7813      82

TukeyHSD(AOV_4)

##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = C_Quant ~ as.factor(Group), data = Skills_Trimmed)
##
## $`as.factor(Group)`
##       diff    lwr   upr  p adj
## 2-1  4.456 -2.322 11.23 0.3195
## 3-1 14.328  7.621 21.04 0.0000
## 4-1 45.223 38.515 51.93 0.0000
## 3-2  9.873  3.096 16.65 0.0014
## 4-2 40.767 33.990 47.54 0.0000
## 4-3 30.894 24.186 37.60 0.0000
```

8 Discriminant Analysis

A discriminant analysis with no imposed structure on the groups. This is the most exploratory approach that we can take with these data, aimed at discovering how the groups can best be separated with weighted linear combinations of the measures.

8.1 No Group Structure

The candisc( ) function provides a flexible way to conduct the discriminant analysis. It provides the most important information (coefficients, significance tests, etc.), including the ability to plot the group locations (and individual data points) on the discriminant functions.

```
# This function takes as input the data frame used for a
# discriminant analysis along with the object into which the
# discriminant analysis results are saved. The candisc( ) function
# is assumed to be used for the discriminant analysis. The
# function return a chi-square test of the hypothesis that the
# current discriminant function and all subsequent discriminant
# functions provide no significant group separation. The test
# parallels the F ratio version reported by candisc( ) function.
DA_Chi_Square <- function(data_frame, candisc_object) {
    n <- length(data_frame[, 1])
    q <- length(candisc_object$coeffs.std[, 1])
    g <- length(unique(candisc_object$factors)[, 1])
    W <- Wilks(candisc_object)
    for (i in seq(1, candisc_object$ndim, 1)) {
        k <- i - 1
        chi_test <- -(n - (q + g)/2 - 1) * log(W$`LR test stat`[i])
        chi_df <- (q - k) * (g - k - 1)
        chi_p <- pchisq(chi_test, chi_df, lower.tail = FALSE)
        if (i == 1) {
            results <- c(chi_test, chi_df, chi_p)
        } else {
            results <- rbind(results, c(chi_test, chi_df, chi_p))
        }
    }
    colnames(results) <- c("Chi_Sq", "df", "p")
    return(results)
}
```

```
LM_1 <- lm(cbind(P_Verbal, P_Quant, C_Verbal, C_Quant) ~ as.factor(Group),
    data = Skills_Trimmed)
LDA_1 <- candisc(LM_1, data = Skills_Trimmed)
LDA_1

##
## Canonical Discriminant Analysis for as.factor(Group):
##
##   CanRsq Eigenvalue Difference Percent Cumulative
## 1  0.885      7.689       2.57    57.3       57.3
## 2  0.837      5.118       2.57    38.1       95.4
## 3  0.382      0.617       2.57     4.6      100.0
```

```
##
## Test of H0: The canonical correlations in the
## current row and all that follow are zero
##
##   LR test stat approx F numDF denDF Pr(> F)
## 1        0.012     89.0    12   244  <2e-16
## 2        0.101     66.5     6   186  <2e-16
## 3        0.618              2
```

```
summary(LDA_1)
```

```
##
## Canonical Discriminant Analysis for as.factor(Group):
##
##   CanRsq Eigenvalue Difference Percent Cumulative
## 1 0.8849     7.6888      2.571  57.276      57.28
## 2 0.8366     5.1181      2.571  38.126      95.40
## 3 0.3817     0.6172      2.571   4.598     100.00
##
## Class means:
##
##     Can1   Can2     Can3
## 1  0.903 -1.137 -1.23762
## 2 -1.655 -3.040  0.71859
## 3  3.916  1.301  0.56375
## 4 -3.230  2.755 -0.01598
##
##  std coefficients:
##             Can1    Can2    Can3
## P_Verbal -0.8401  0.5345 -1.2666
## P_Quant  -0.4977 -1.2016  1.3543
## C_Verbal  0.7376  0.1601 -0.7209
## C_Quant  -0.2314  1.0724  0.9161
```

```
LDA_1$dfh
```

```
## [1] 3
```

```
LDA_1$dfe
```

```
## [1] 95
```

```
LDA_1$pct
```

```
## [1]  5.728e+01  3.813e+01  4.598e+00 -2.955e-15
```

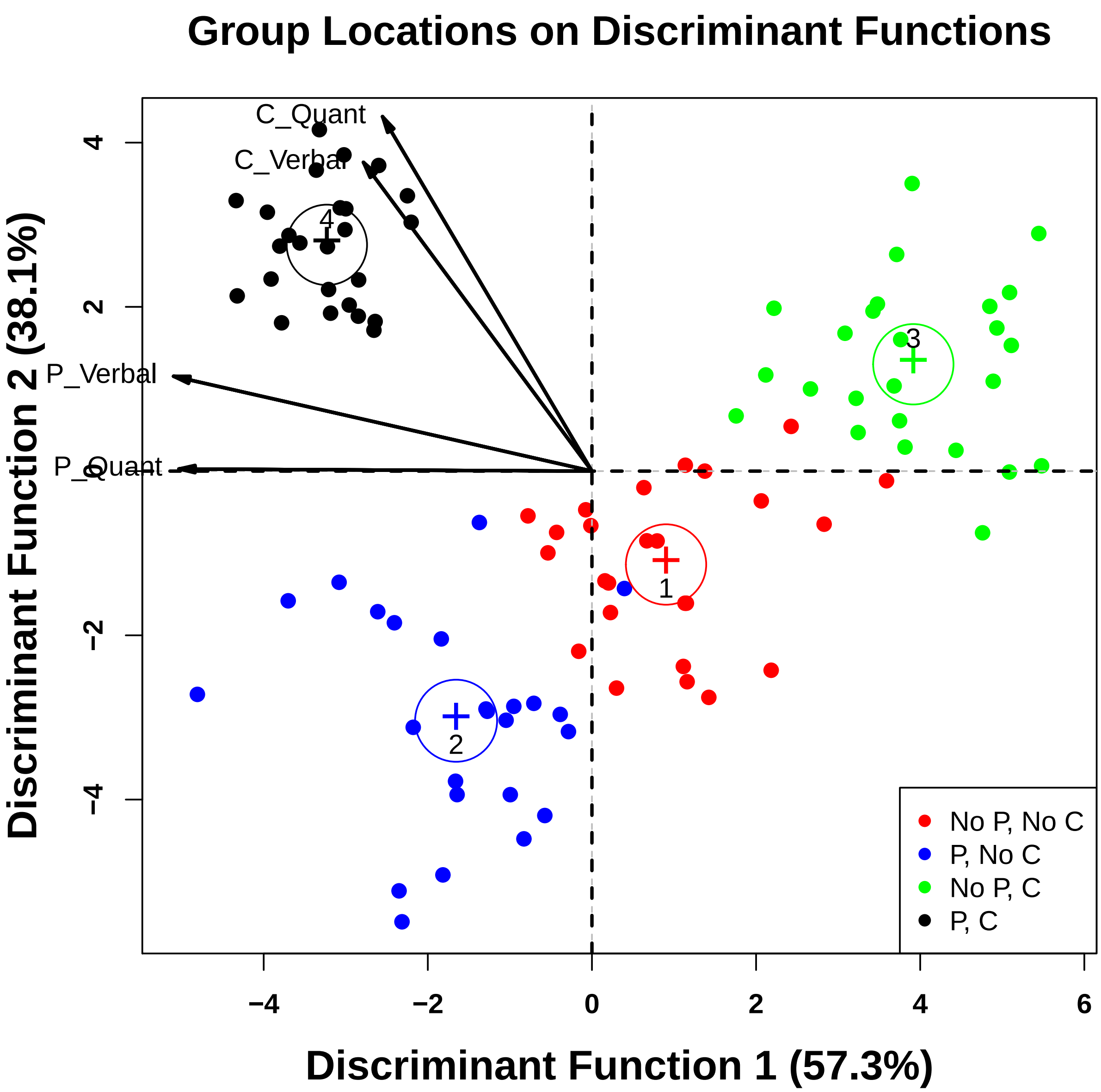
```
LDA_1$ndim
```

```
## [1] 3
```

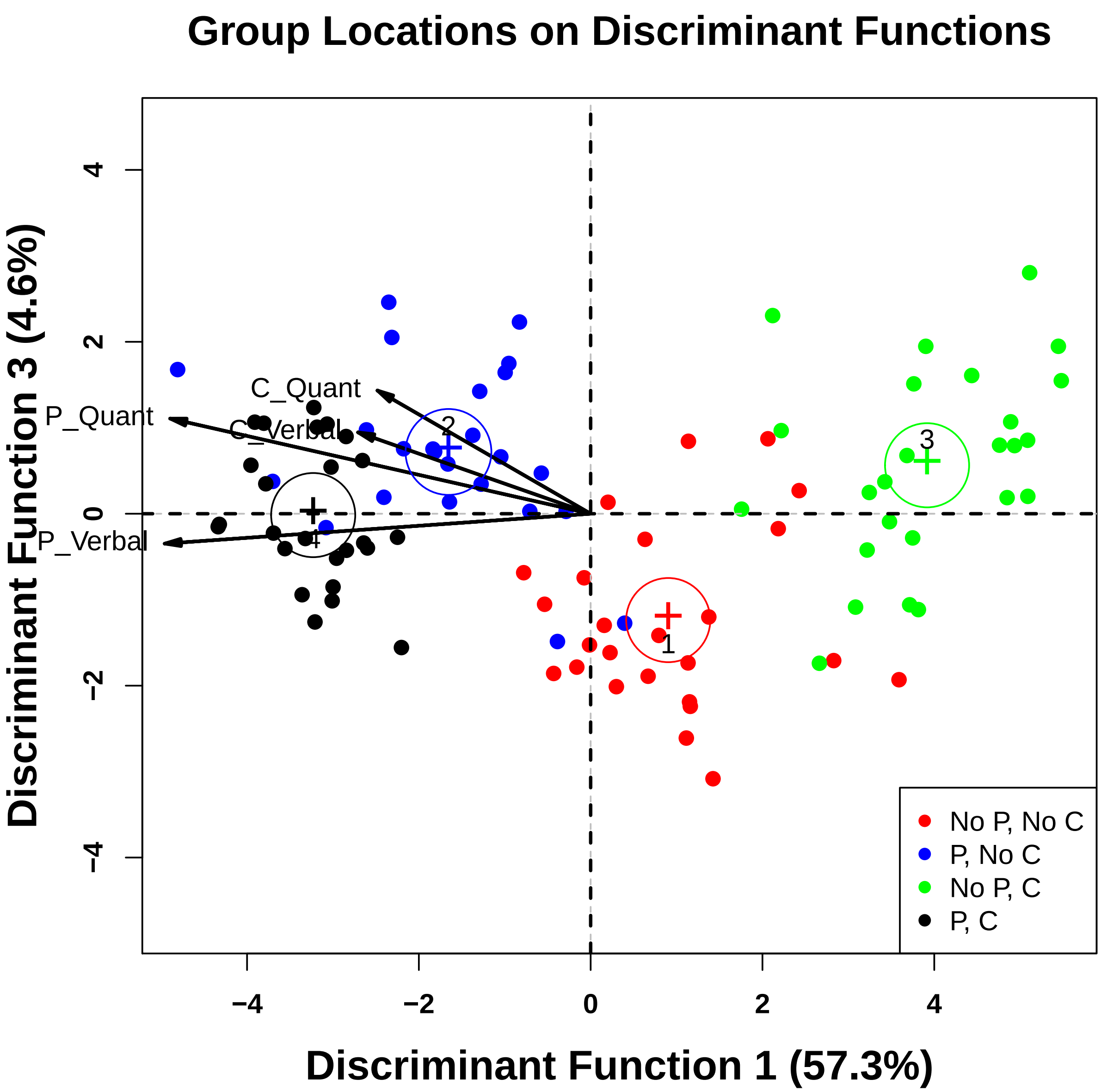
```
LDA_1$coeffs.raw
```

```
##              Can1     Can2     Can3
## P_Verbal -0.08559  0.05446 -0.12903
## P_Quant  -0.05376 -0.12979  0.14629
## C_Verbal  0.07321  0.01589 -0.07155
## C_Quant  -0.02552  0.11825  0.10102
```

```r
LDA_1$coeffs.std

##             Can1    Can2    Can3
## P_Verbal -0.8401  0.5345 -1.2666
## P_Quant  -0.4977 -1.2016  1.3543
## C_Verbal  0.7376  0.1601 -0.7209
## C_Quant  -0.2314  1.0724  0.9161

LDA_1$structure

##             Can1     Can2     Can3
## P_Verbal -0.9675 0.219137 -0.06784
## P_Quant  -0.9551 0.004908  0.21602
## C_Verbal -0.5284 0.713651  0.18468
## C_Quant  -0.4843 0.819206  0.27969

DA_Chi_Square(Skills_Trimmed, LDA_1)

##         Chi_Sq df         p
## results 418.68 12 4.180e-82
##         215.44  6 9.737e-44
##          45.19  2 1.541e-10

plot(LDA_1, main = list("Group Locations on Discriminant Functions",
    cex = 1.5), cex = 1.25, font.axis = 2, col = c("red", "blue",
    "green", "black"), pch = c(16, 16, 16, 16), font.lab = 2, cex.lab = 1.5,
    prefix = "Discriminant Function ", var.col = "black", var.lwd = 2,
    which = c(1, 2))

## Vector scale factor set to 5.269

abline(v = 0, lty = 2, lwd = 2, col = "black")
abline(h = 0, lty = 2, lwd = 2, col = "black")
legend("bottomright", c("No P, No C", "P, No C", "No P, C", "P, C"),
    col = c("red", "blue", "green", "black"), pch = 16)
```



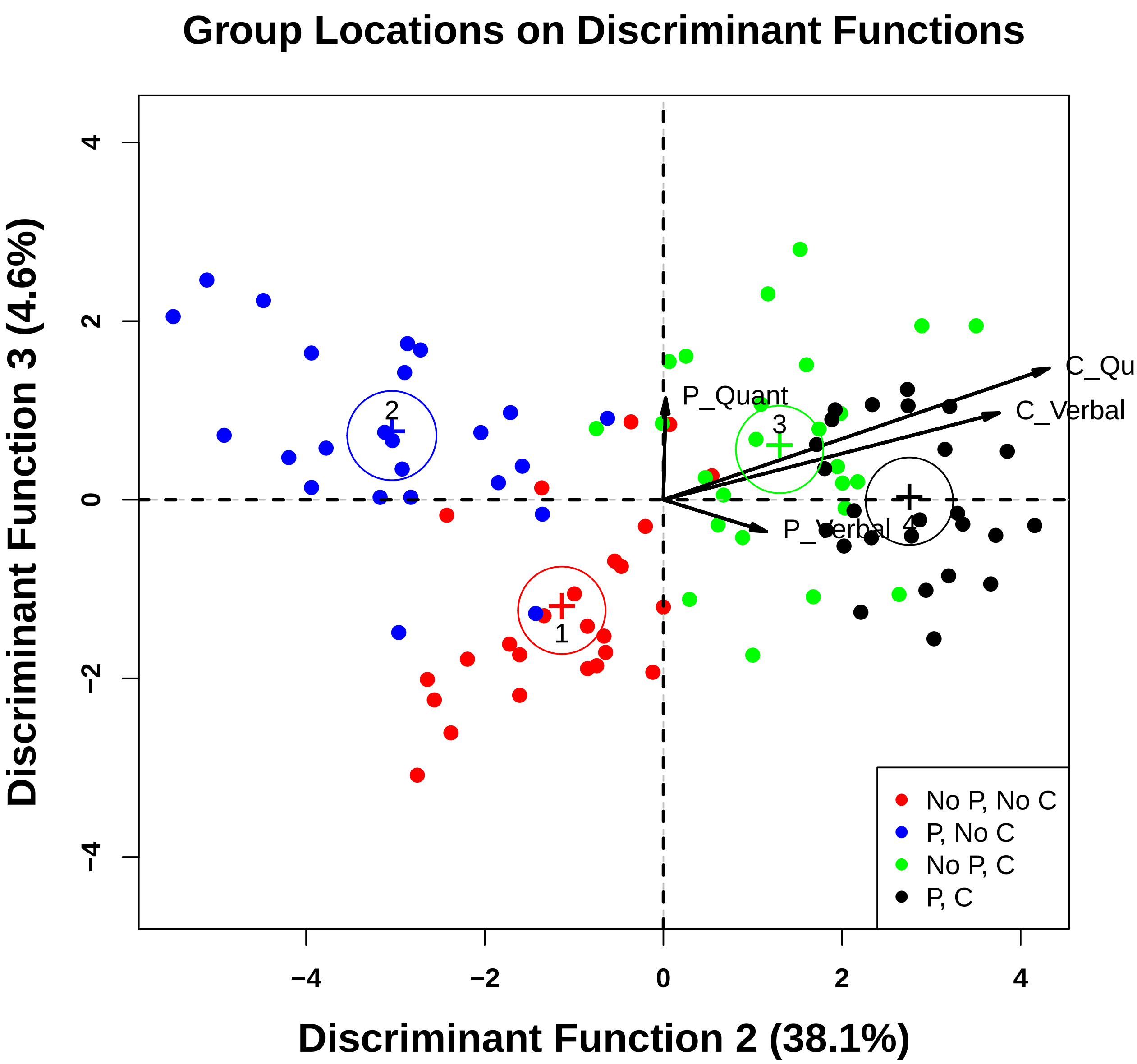

```r
plot(LDA_1, main = list("Group Locations on Discriminant Functions",
    cex = 1.5), cex = 1.25, font.axis = 2, col = c("red", "blue",
    "green", "black"), pch = c(16, 16, 16, 16), font.lab = 2, cex.lab = 1.5,
    prefix = "Discriminant Function ", var.col = "black", var.lwd = 2,
    which = c(1, 3))

## Vector scale factor set to 5.125

abline(v = 0, lty = 2, lwd = 2, col = "black")
abline(h = 0, lty = 2, lwd = 2, col = "black")
legend("bottomright", c("No P, No C", "P, No C", "No P, C", "P, C"),
    col = c("red", "blue", "green", "black"), pch = 16)
```




```r
plot(LDA_1, main = list("Group Locations on Discriminant Functions",
    cex = 1.5), cex = 1.25, font.axis = 2, col = c("red", "blue",
    "green", "black"), pch = c(16, 16, 16, 16), font.lab = 2, cex.lab = 1.5,
    prefix = "Discriminant Function ", var.col = "black", var.lwd = 2,
    which = c(2, 3))
```

```
## Vector scale factor set to 5.269
```

```r
abline(v = 0, lty = 2, lwd = 2, col = "black")
abline(h = 0, lty = 2, lwd = 2, col = "black")
legend("bottomright", c("No P, No C", "P, No C", "No P, C", "P, C"),
    col = c("red", "blue", "green", "black"), pch = 16)
```



The manova( ) function is also useful. It provides the full set of significance tests as well as the sums of squares and cross-products matrices on which the tests are based.

```
Outcomes <- as.matrix(Skills_Trimmed[, 2:5])
MANOVA_1 <- manova(Outcomes ~ as.factor(Group), data = Skills_Trimmed)
summary(MANOVA_1, test = "Wilks")

##                   Df  Wilks approx F num Df den Df Pr(>F)
## as.factor(Group)   3 0.0116       89     12    244 <2e-16
## Residuals         95

summary(MANOVA_1, test = "Hotelling")

##                   Df Hotelling-Lawley approx F num Df den Df
## as.factor(Group)   3             13.4      101     12    272
```

```
## Residuals        95
##                  Pr(>F)
## as.factor(Group) <2e-16
## Residuals
```

```r
summary(MANOVA_1, test = "Pillai")
```

```
##                  Df Pillai approx F num Df den Df Pr(>F)
## as.factor(Group)  3    2.1     55.1     12    282 <2e-16
## Residuals        95
```

```r
summary(MANOVA_1, test = "Roy")
```

```
##                  Df  Roy approx F num Df den Df Pr(>F)
## as.factor(Group)  3 7.69      181      4     94 <2e-16
## Residuals        95
```

```r
summary(MANOVA_1)$SS
```

```
## $`as.factor(Group)`
##          P_Verbal P_Quant C_Verbal C_Quant
## P_Verbal    61407   46604    26935   29184
## P_Quant     46604   38419    17580   18527
## C_Verbal    26935   17580    21084   25396
## C_Quant     29184   18527    25396   31016
##
## $Residuals
##          P_Verbal P_Quant C_Verbal C_Quant
## P_Verbal     9153    6821     5653    4452
## P_Quant      6821    8143     6379    4975
## C_Verbal     5653    6379     9644    7271
## C_Quant      4452    4975     7271    7813
```

```r
summary(MANOVA_1)$stats
```

```
##                  Df Pillai approx F num Df den Df    Pr(>F)
## as.factor(Group)  3  2.103    55.11     12    282 1.022e-66
## Residuals        95     NA       NA     NA     NA        NA
```

```r
summary(MANOVA_1)$Eigenvalues
```

```
##                   [,1]  [,2]   [,3]      [,4]
## as.factor(Group) 7.689 5.118 0.6172 1.145e-15
```

```r
summary.aov(MANOVA_1)
```

```
##  Response P_Verbal :
##                  Df Sum Sq Mean Sq F value Pr(>F)
## as.factor(Group)  3  61407   20469     212 <2e-16
## Residuals        95   9153      96
##
##  Response P_Quant :
##                  Df Sum Sq Mean Sq F value Pr(>F)
## as.factor(Group)  3  38419   12806     149 <2e-16
## Residuals        95   8143      86
##
```

```
##  Response C_Verbal :
##                   Df Sum Sq Mean Sq F value Pr(>F)
## as.factor(Group)  3  21084    7028    69.2 <2e-16
## Residuals        95   9644     102
##
##  Response C_Quant :
##                   Df Sum Sq Mean Sq F value Pr(>F)
## as.factor(Group)  3  31016   10339     126 <2e-16
## Residuals        95   7813      82
```

The Manova( ) function (note capitalization) allows specifying Type II or Type III sums of squares. That is not useful here with group unstructured, but could be important in unbalanced designs. This function also provides the sums of squares and cross-products matrices.

```
LM_4 <- lm(cbind(P_Verbal, P_Quant, C_Verbal, C_Quant) ~ as.factor(Group),
    data = Skills_Trimmed)
MANOVA_3 <- Manova(LM_4, type = "III")
summary(MANOVA_3, multivariate = TRUE)

##
## Type III MANOVA Tests:
##
## Sum of squares and products for error:
##          P_Verbal P_Quant C_Verbal C_Quant
## P_Verbal     9153    6821     5653    4452
## P_Quant      6821    8143     6379    4975
## C_Verbal     5653    6379     9644    7271
## C_Quant      4452    4975     7271    7813
##
## ------------------------------------------
##
## Term: (Intercept)
##
## Sum of squares and products for the hypothesis:
##          P_Verbal P_Quant C_Verbal C_Quant
## P_Verbal   320920  331620   325071  351025
## P_Quant    331620  342677   335910  362729
## C_Verbal   325071  335910   329276  355565
## C_Quant    351025  362729   355565  383954
##
## Multivariate Tests: (Intercept)
##                   Df test stat approx F num Df den Df Pr(>F)
## Pillai             1      0.98     1376      4     92 <2e-16
## Wilks              1      0.02     1376      4     92 <2e-16
## Hotelling-Lawley  1     59.82     1376      4     92 <2e-16
## Roy                1     59.82     1376      4     92 <2e-16
##
## ------------------------------------------
##
## Term: as.factor(Group)
##
## Sum of squares and products for the hypothesis:
##          P_Verbal P_Quant C_Verbal C_Quant
```

```
## P_Verbal    61407   46604    26935   29184
## P_Quant     46604   38419    17580   18527
## C_Verbal    26935   17580    21084   25396
## C_Quant     29184   18527    25396   31016
##
## Multivariate Tests: as.factor(Group)
##                  Df test stat approx F num Df den Df Pr(>F)
## Pillai            3     2.103    55.11     12  282.0 <2e-16
## Wilks             3     0.012    89.03     12  243.7 <2e-16
## Hotelling-Lawley  3    13.424   101.43     12  272.0 <2e-16
## Roy               3     7.689   180.69      4   94.0 <2e-16
```

```
MANOVA_3 <- Manova(LM_4, type = "II")
summary(MANOVA_3, multivariate = TRUE)
```

```
##
## Type II MANOVA Tests:
##
## Sum of squares and products for error:
##          P_Verbal P_Quant C_Verbal C_Quant
## P_Verbal     9153    6821     5653    4452
## P_Quant      6821    8143     6379    4975
## C_Verbal     5653    6379     9644    7271
## C_Quant      4452    4975     7271    7813
##
## ------------------------------------------
##
## Term: as.factor(Group)
##
## Sum of squares and products for the hypothesis:
##          P_Verbal P_Quant C_Verbal C_Quant
## P_Verbal    61407   46604    26935   29184
## P_Quant     46604   38419    17580   18527
## C_Verbal    26935   17580    21084   25396
## C_Quant     29184   18527    25396   31016
##
## Multivariate Tests: as.factor(Group)
##                  Df test stat approx F num Df den Df Pr(>F)
## Pillai            3     2.103    55.11     12  282.0 <2e-16
## Wilks             3     0.012    89.03     12  243.7 <2e-16
## Hotelling-Lawley  3    13.424   101.43     12  272.0 <2e-16
## Roy               3     7.689   180.69      4   94.0 <2e-16
```

8.2 Group Structure

The discriminant analysis can also be performed on "groups" defined by contrasts. These might represent a factorial structure or other comparisons of interest.

8.2.1 Factorial Structure

```
LM_2 <- lm(cbind(P_Verbal, P_Quant, C_Verbal, C_Quant) ~ Tx_P + Tx_C +
    Tx_P:Tx_C, data = Skills_Trimmed)
```

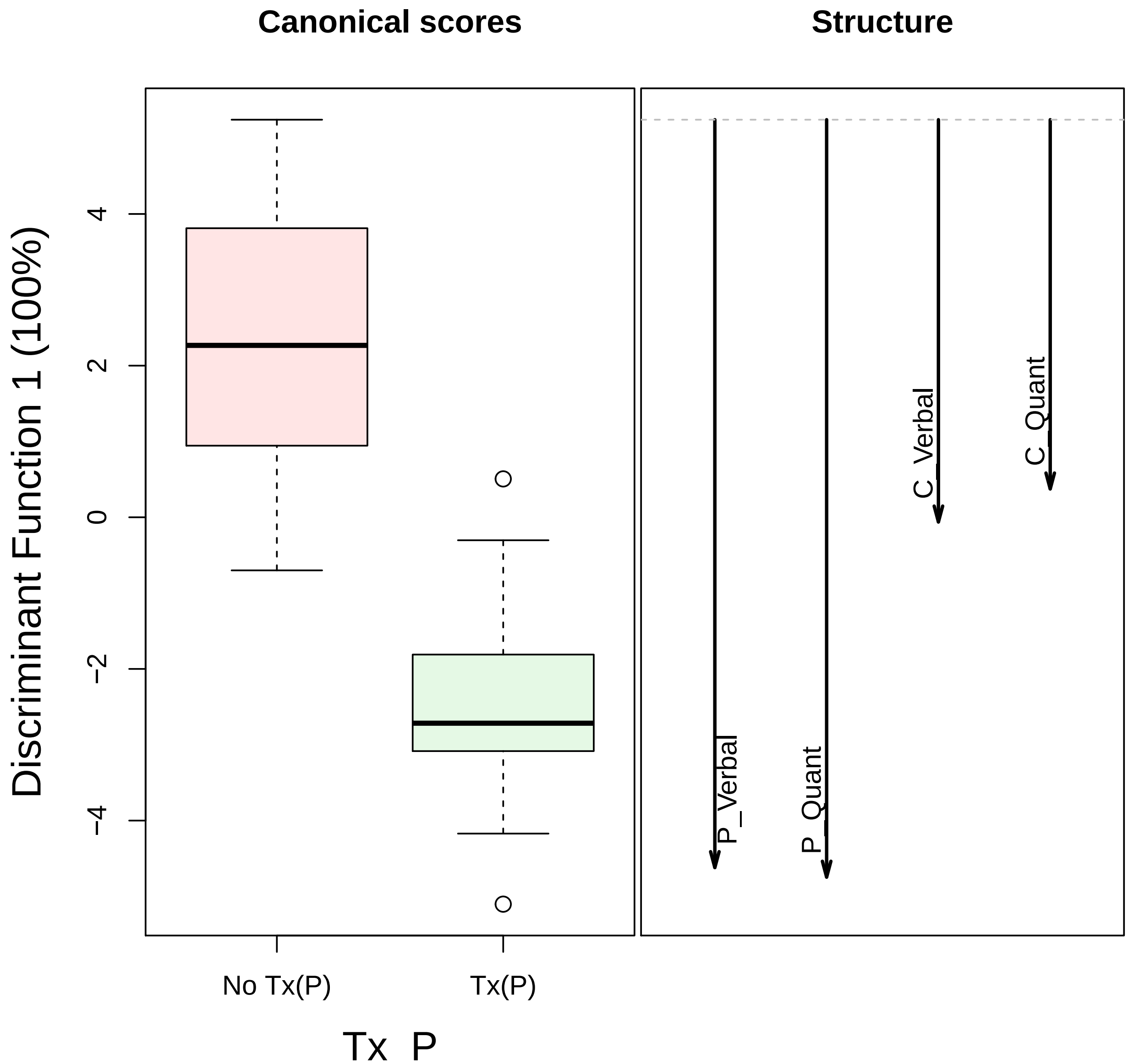
```
LDA_2 <- candisc(LM_2, term = "Tx_P", data = Skills_Trimmed, type = "2")
LDA_2

##
## Canonical Discriminant Analysis for Tx_P:
##
##   CanRsq Eigenvalue Difference Percent Cumulative
## 1  0.863       6.32                100        100
##
## Test of H0: The canonical correlations in the
## current row and all that follow are zero
##
##   LR test stat approx F numDF denDF Pr(> F)
## 1        0.137      149     4    94  <2e-16

summary(LDA_2)

##
## Canonical Discriminant Analysis for Tx_P:
##
##   CanRsq Eigenvalue Difference Percent Cumulative
## 1 0.8634      6.323                100        100
##
## Class means:
##
## [1]  2.434 -2.484
##
##  std coefficients:
## P_Verbal  P_Quant C_Verbal  C_Quant
##  -0.6354  -0.7287   0.8355  -0.3095

LDA_2$dfh

## [1] 1

LDA_2$dfe

## [1] 95

LDA_2$pct

## [1]  1.000e+02  5.619e-14 -1.237e-14 -1.290e-14

LDA_2$ndim

## [1] 1

LDA_2$coeffs.raw

##              Can1
## P_Verbal -0.06473
## P_Quant  -0.07870
## C_Verbal  0.08292
## C_Quant  -0.03413

LDA_2$coeffs.std
```

```
##             Can1
## P_Verbal -0.6354
## P_Quant  -0.7287
## C_Verbal  0.8355
## C_Quant  -0.3095
```

```
LDA_2$structure
```

```
##             Can1
## P_Verbal -0.9534
## P_Quant  -0.9655
## C_Verbal -0.5128
## C_Quant  -0.4707
```

```
plot(LDA_2, cex = 1.25, font.axis = 2, font.lab = 2, cex.lab = 1.5,
    prefix = "Discriminant Function ", var.col = "black", var.lwd = 2)
```




```
LM_3 <- lm(cbind(P_Verbal, P_Quant, C_Verbal, C_Quant) ~ Tx_P + Tx_C +
    Tx_P:Tx_C, data = Skills_Trimmed)
LDA_3 <- candisc(LM_3, term = "Tx_C", data = Skills_Trimmed, type = "2")
LDA_3

##
## Canonical Discriminant Analysis for Tx_C:
##
##   CanRsq Eigenvalue Difference Percent Cumulative
## 1  0.821        4.6                 100        100
##
## Test of H0: The canonical correlations in the
## current row and all that follow are zero
##
```

```
##   LR test stat approx F numDF denDF Pr(> F)
## 1        0.179      108     4    94  <2e-16
```

```r
summary(LDA_3)
```

```
##
## Canonical Discriminant Analysis for Tx_C:
##
##   CanRsq Eigenvalue Difference Percent Cumulative
## 1 0.8214      4.599                100        100
##
## Class means:
##
## [1] -2.117  2.075
##
##  std coefficients:
## P_Verbal  P_Quant C_Verbal  C_Quant
##   0.2082  -1.0842   0.1928   1.1245
```

```r
LDA_3$dfh
```

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

```r
LDA_3$dfe
```

```
## [1] 95
```

```r
LDA_3$pct
```

```
## [1]  1.000e+02  1.931e-14  0.000e+00 -8.562e-15
```

```r
LDA_3$ndim
```

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

```r
LDA_3$coeffs.raw
```

```
##               Can1
## P_Verbal  0.02121
## P_Quant  -0.11711
## C_Verbal  0.01914
## C_Quant   0.12399
```

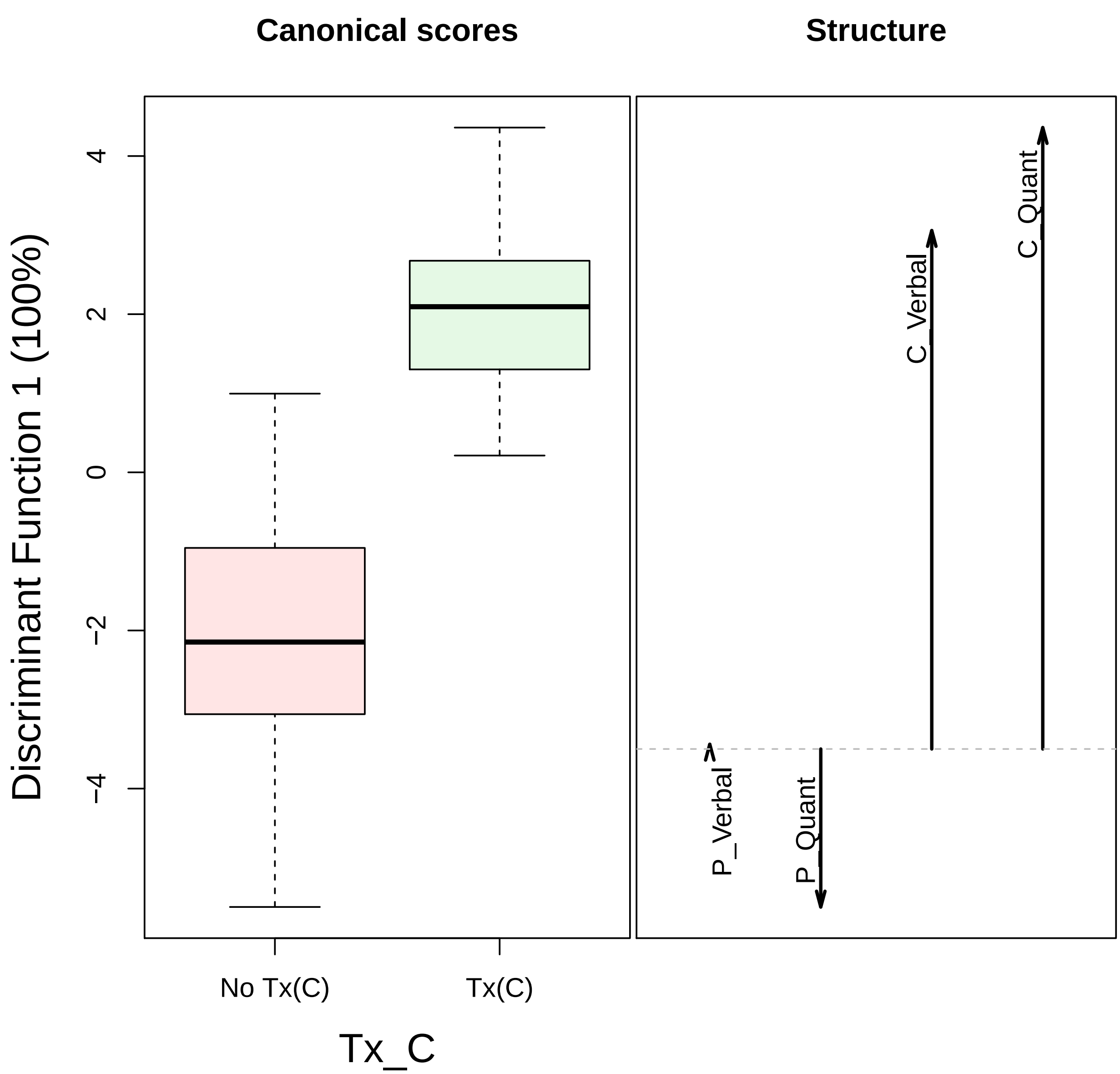
```r
LDA_3$coeffs.std
```

```
##              Can1
## P_Verbal  0.2082
## P_Quant  -1.0842
## C_Verbal  0.1928
## C_Quant   1.1245
```

```r
LDA_3$structure
```

```
##              Can1
## P_Verbal  0.0054
## P_Quant  -0.1820
## C_Verbal  0.5971
## C_Quant   0.7156
```

```r
plot(LDA_3, cex = 1.25, font.axis = 2, font.lab = 2, cex.lab = 1.5,
    prefix = "Discriminant Function ", var.col = "black", var.lwd = 2)
```




```
LM_4 <- lm(cbind(P_Verbal, P_Quant, C_Verbal, C_Quant) ~ Tx_P + Tx_C +
    Tx_P:Tx_C, data = Skills_Trimmed)
LDA_4 <- candisc(LM_4, term = "Tx_P:Tx_C", data = Skills_Trimmed,
    type = "2")
LDA_4

##
## Canonical Discriminant Analysis for Tx_P:Tx_C:
##
##    CanRsq Eigenvalue Difference Percent Cumulative
## 1  0.716       2.52                 100        100
##
## Test of H0: The canonical correlations in the
## current row and all that follow are zero
```

```
##
##   LR test stat approx F numDF denDF Pr(> F)
## 1        0.284     59.3     4    94  <2e-16
```

```r
summary(LDA_4)
```

```
##
## Canonical Discriminant Analysis for Tx_P:Tx_C:
##
##   CanRsq Eigenvalue Difference Percent Cumulative
## 1 0.7162      2.523                100        100
##
## Class means:
##
## [1]  0.7747  0.7119  2.4148 -3.8729
##
##  std coefficients:
## P_Verbal  P_Quant C_Verbal  C_Quant
##  -1.4234   0.8328   0.1636  -0.3756
```

```r
LDA_4$dfh
```

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

```r
LDA_4$dfe
```

```
## [1] 95
```

```r
LDA_4$pct
```

```
## [1]  1.000e+02  2.568e-14  8.801e-15 -1.276e-14
```

```r
LDA_4$ndim
```

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

```r
LDA_4$coeffs.raw
```

```
##              Can1
## P_Verbal -0.14501
## P_Quant   0.08996
## C_Verbal  0.01624
## C_Quant  -0.04142
```

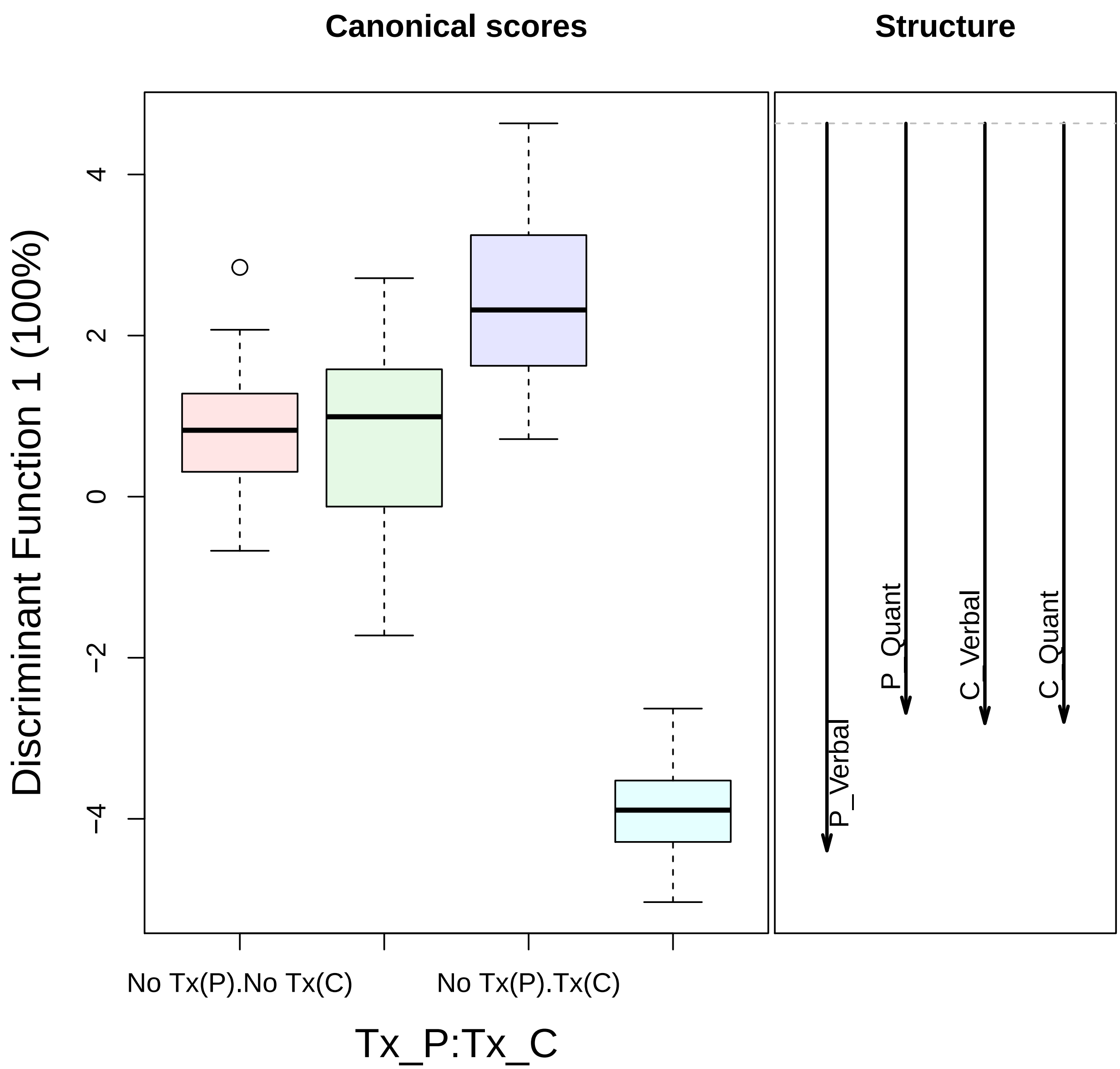
```r
LDA_4$coeffs.std
```

```
##             Can1
## P_Verbal -1.4234
## P_Quant   0.8328
## C_Verbal  0.1636
## C_Quant  -0.3756
```

```r
LDA_4$structure
```

```
##             Can1
## P_Verbal -0.9339
## P_Quant  -0.7572
## C_Verbal -0.7704
## C_Quant  -0.7687
```

```r
plot(LDA_4, cex = 1.25, font.axis = 2, font.lab = 2, cex.lab = 1.5,
    prefix = "Discriminant Function ", var.col = "black", var.lwd = 2)
```



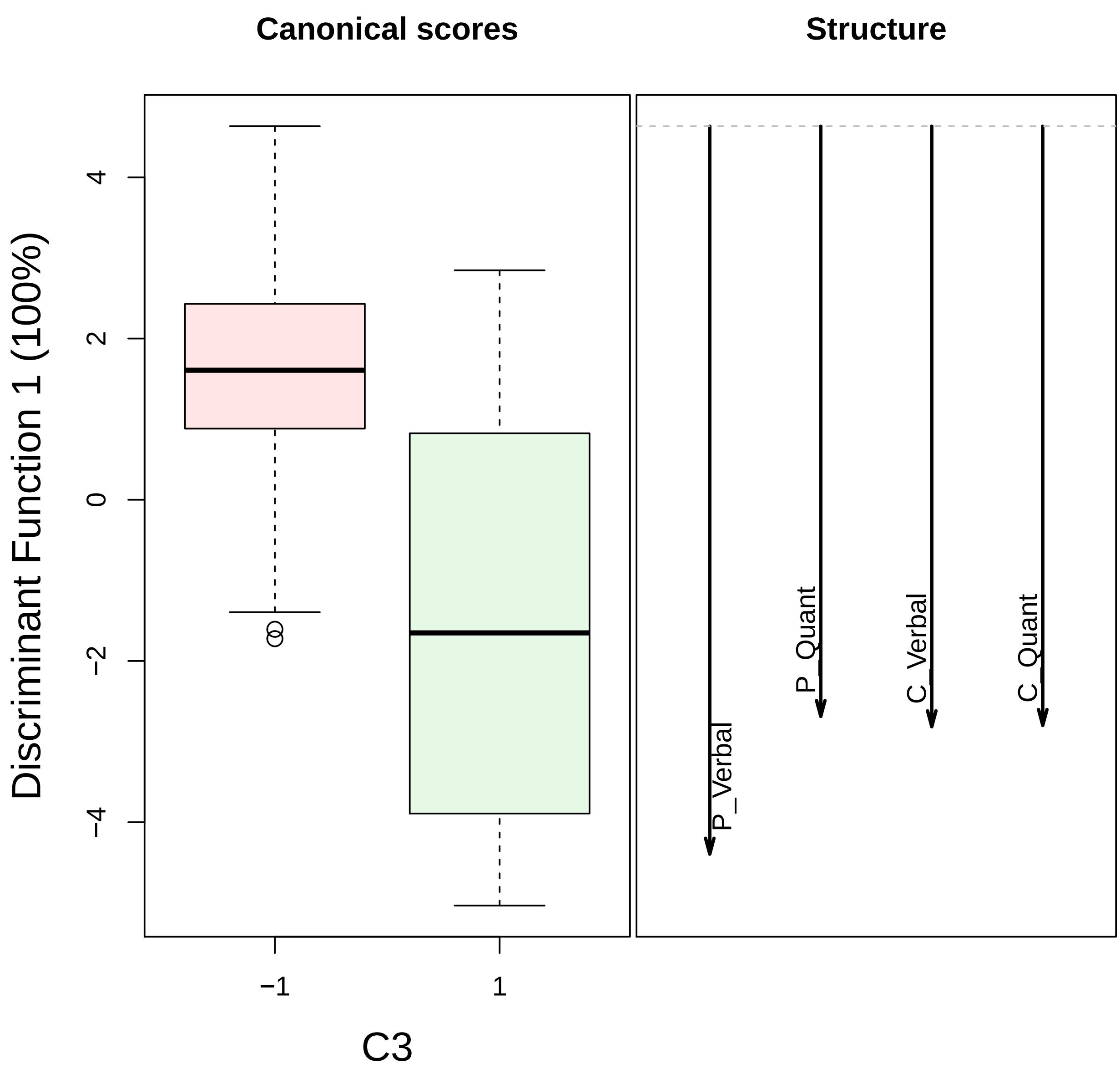
8.2.2 Factorial Structure by Contrast Codes

```
LM_5 <- lm(cbind(P_Verbal, P_Quant, C_Verbal, C_Quant) ~ C1 + C2 +
    C3, data = Skills_Trimmed)
LDA_5 <- candisc(LM_5, term = "C3", data = Skills_Trimmed, type = "2")
LDA_5

##
## Canonical Discriminant Analysis for C3:
##
##   CanRsq Eigenvalue Difference Percent Cumulative
## 1  0.716       2.52                 100        100
##
```

```
## Test of H0: The canonical correlations in the
## current row and all that follow are zero
##
##   LR test stat approx F numDF denDF Pr(> F)
## 1        0.284     59.3     4    94  <2e-16

summary(LDA_5)

##
## Canonical Discriminant Analysis for C3:
##
##   CanRsq Eigenvalue Difference Percent Cumulative
## 1 0.7162      2.523                100        100
##
## Class means:
##
## [1]  1.581 -1.549
##
##  std coefficients:
## P_Verbal  P_Quant C_Verbal  C_Quant
##  -1.4234   0.8328   0.1636  -0.3756

LDA_5$dfh

## [1] 1

LDA_5$dfe

## [1] 95

LDA_5$pct

## [1]  1.000e+02  2.568e-14  8.801e-15 -1.276e-14

LDA_5$ndim

## [1] 1

LDA_5$coeffs.raw

##              Can1
## P_Verbal -0.14501
## P_Quant   0.08996
## C_Verbal  0.01624
## C_Quant  -0.04142

LDA_5$coeffs.std

##             Can1
## P_Verbal -1.4234
## P_Quant   0.8328
## C_Verbal  0.1636
## C_Quant  -0.3756

LDA_5$structure
```

```
##              Can1
## P_Verbal -0.9339
## P_Quant  -0.7572
## C_Verbal -0.7704
## C_Quant  -0.7687
```

```
plot(LDA_5, cex = 1.25, font.axis = 2, font.lab = 2, cex.lab = 1.5,
    prefix = "Discriminant Function ", var.col = "black", var.lwd = 2)
```




```
LM_6 <- lm(cbind(P_Verbal, P_Quant, C_Verbal, C_Quant) ~ C1 + C2,
    data = Skills_Trimmed)
LDA_6 <- candisc(LM_6, term = "C2", data = Skills_Trimmed, type = "2")
LDA_6
```

```
##
## Canonical Discriminant Analysis for C2:
##
##   CanRsq Eigenvalue Difference Percent Cumulative
## 1  0.808       4.21                100        100
##
## Test of H0: The canonical correlations in the
## current row and all that follow are zero
##
##   LR test stat approx F numDF denDF Pr(> F)
## 1        0.192       99     4    94  <2e-16
```

```
summary(LDA_6)
```

```
##
## Canonical Discriminant Analysis for C2:
##
##   CanRsq Eigenvalue Difference Percent Cumulative
## 1 0.8081      4.212                100        100
##
## Class means:
##
## [1] -2.032  1.991
##
##  std coefficients:
## P_Verbal  P_Quant C_Verbal  C_Quant
##  -0.2506  -1.0956   0.2901   1.3444
```

```
LDA_6$dfh
```

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

```
LDA_6$dfe
```

```
## [1] 96
```

```
LDA_6$pct
```

```
## [1]  1.000e+02  4.217e-14  7.289e-16 -9.460e-15
```

```
LDA_6$ndim
```

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

```
LDA_6$coeffs.raw
```

```
##              Can1
## P_Verbal -0.01511
## P_Quant  -0.09982
## C_Verbal  0.02429
## C_Quant   0.11956
```

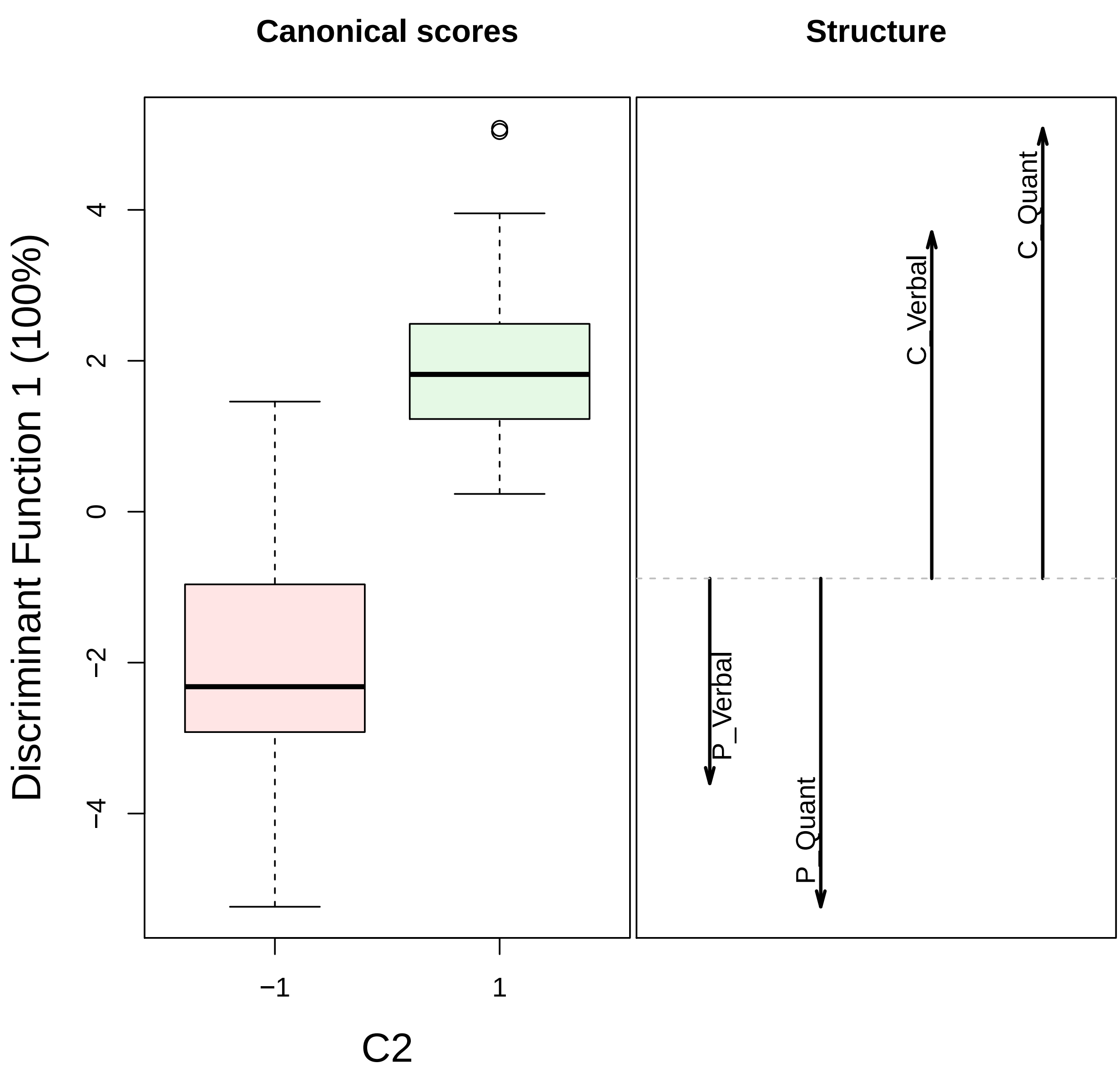
```
LDA_6$coeffs.std
```

```
##             Can1
## P_Verbal -0.2506
## P_Quant  -1.0956
## C_Verbal  0.2901
## C_Quant   1.3444
```

```r
LDA_6$structure
```

```
##              Can1
## P_Verbal -0.2475
## P_Quant  -0.3964
## C_Verbal  0.4183
## C_Quant   0.5434
```

```r
plot(LDA_6, cex = 1.25, font.axis = 2, font.lab = 2, cex.lab = 1.5,
    prefix = "Discriminant Function ", var.col = "black", var.lwd = 2)
```



8.2.3 Special Contrast Codes

```
LM_7 <- lm(cbind(P_Verbal, P_Quant, C_Verbal, C_Quant) ~ S1 + S2 +
    S3, data = Skills_Trimmed)
LDA_7 <- candisc(LM_7, term = "S1", data = Skills_Trimmed, type = "2")
LDA_7

##
## Canonical Discriminant Analysis for S1:
##
##   CanRsq Eigenvalue Difference Percent Cumulative
## 1  0.559       1.27                 100        100
##
## Test of H0: The canonical correlations in the
## current row and all that follow are zero
##
##   LR test stat approx F numDF denDF Pr(> F)
## 1        0.441     29.8     4    94 5.5e-16

summary(LDA_7)

##
## Canonical Discriminant Analysis for S1:
##
##   CanRsq Eigenvalue Difference Percent Cumulative
## 1 0.5587      1.266                 100        100
##
## Class means:
##
## [1]  0.6445 -1.9078
##
##  std coefficients:
## P_Verbal  P_Quant C_Verbal  C_Quant
##  -0.1116   0.4274  -0.7349   1.3351

LDA_7$dfh

## [1] 1

LDA_7$dfe

## [1] 95

LDA_7$pct

## [1]  1.000e+02  2.385e-14  1.839e-15 -2.245e-14

LDA_7$ndim

## [1] 1

LDA_7$coeffs.raw

##              Can1
## P_Verbal -0.01137
## P_Quant   0.04616
## C_Verbal -0.07294
## C_Quant   0.14722
```

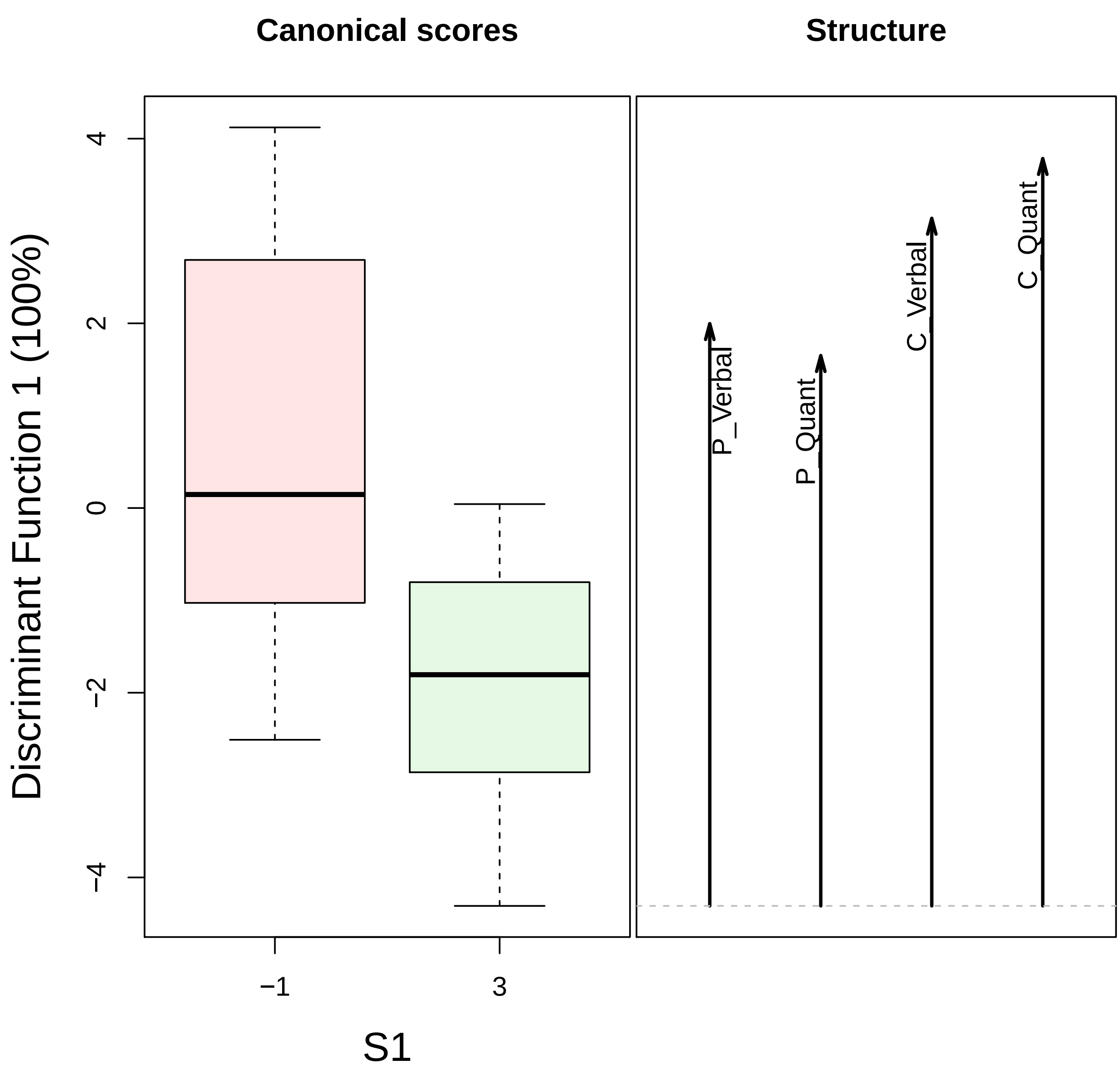
```
LDA_7$coeffs.std
```

```
##             Can1
## P_Verbal -0.1116
## P_Quant   0.4274
## C_Verbal -0.7349
## C_Quant   1.3351
```

```
LDA_7$structure
```

```
##            Can1
## P_Verbal 0.7478
## P_Quant  0.7068
## C_Verbal 0.8832
## C_Quant  0.9599
```

```
plot(LDA_7, cex = 1.25, font.axis = 2, font.lab = 2, cex.lab = 1.5,
    prefix = "Discriminant Function ", var.col = "black", var.lwd = 2)
```




```
LM_8 <- lm(cbind(P_Verbal, P_Quant, C_Verbal, C_Quant) ~ S1 + S2 +
    S3, data = Skills_Trimmed)
LDA_8 <- candisc(LM_8, term = "S2", data = Skills_Trimmed, type = "2")
LDA_8

##
## Canonical Discriminant Analysis for S2:
##
##   CanRsq Eigenvalue Difference Percent Cumulative
## 1  0.836        5.1                100        100
##
## Test of H0: The canonical correlations in the
## current row and all that follow are zero
##
```

```
##   LR test stat approx F numDF denDF Pr(> F)
## 1        0.164      120     4    94  <2e-16
```

```r
summary(LDA_8)
```

```
##
## Canonical Discriminant Analysis for S2:
##
##   CanRsq Eigenvalue Difference Percent Cumulative
## 1 0.8361        5.1                100        100
##
## Class means:
##
## [1]  1.9835 -0.6239 -3.4824
##
##  std coefficients:
## P_Verbal  P_Quant C_Verbal  C_Quant
##   0.2915  -1.4063   0.4767   0.8352
```

```r
LDA_8$dfh
```

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

```r
LDA_8$dfe
```

```
## [1] 95
```

```r
LDA_8$pct
```

```
## [1] 1.000e+02 1.742e-14 3.705e-15 9.210e-16
```

```r
LDA_8$ndim
```

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

```r
LDA_8$coeffs.raw
```

```
##              Can1
## P_Verbal  0.02970
## P_Quant  -0.15190
## C_Verbal  0.04732
## C_Quant   0.09209
```

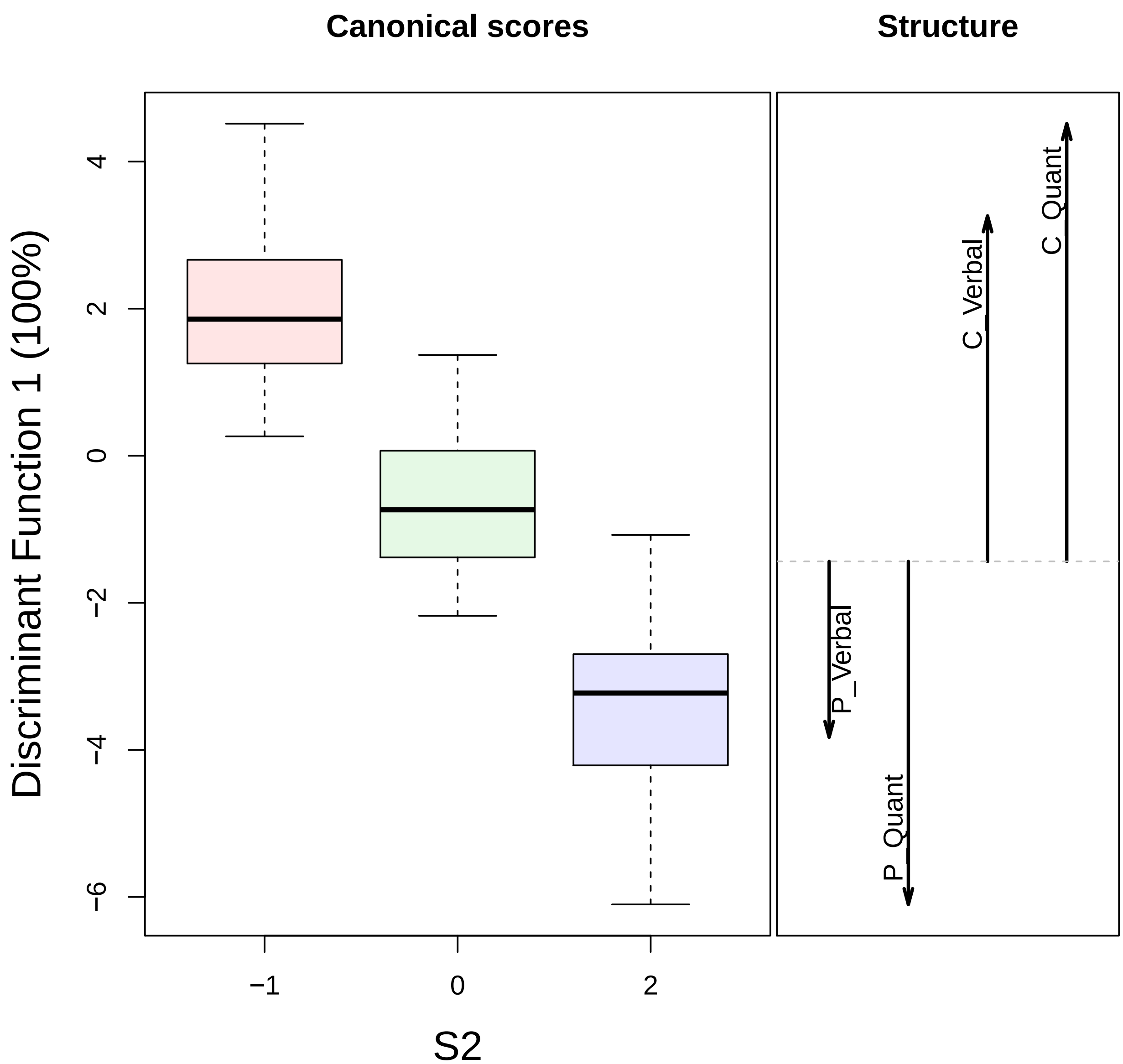
```r
LDA_8$coeffs.std
```

```
##             Can1
## P_Verbal  0.2915
## P_Quant  -1.4063
## C_Verbal  0.4767
## C_Quant   0.8352
```

```r
LDA_8$structure
```

```
##             Can1
## P_Verbal -0.2102
## P_Quant  -0.4103
## C_Verbal  0.4133
## C_Quant   0.5236
```

```r
plot(LDA_8, cex = 1.25, font.axis = 2, font.lab = 2, cex.lab = 1.5,
    prefix = "Discriminant Function ", var.col = "black", var.lwd = 2)
```




```
LM_9 <- lm(cbind(P_Verbal, P_Quant, C_Verbal, C_Quant) ~ S1 + S2 +
    S3, data = Skills_Trimmed)
LDA_9 <- candisc(LM_9, term = "S3", data = Skills_Trimmed, type = "2")
LDA_9

##
## Canonical Discriminant Analysis for S3:
##
##    CanRsq Eigenvalue Difference Percent Cumulative
## 1  0.876       7.04                 100        100
##
## Test of H0: The canonical correlations in the
## current row and all that follow are zero
##
```

```
##   LR test stat approx F numDF denDF Pr(> F)
## 1        0.124      166     4    94  <2e-16
```

```r
summary(LDA_9)
```

```
##
## Canonical Discriminant Analysis for S3:
##
##   CanRsq Eigenvalue Difference Percent Cumulative
## 1 0.8757      7.042                100        100
##
## Class means:
##
## [1]  3.70416 -0.04715 -3.61174
##
##  std coefficients:
## P_Verbal  P_Quant C_Verbal  C_Quant
##   1.0273   0.1400  -0.6316   0.3666
```

```r
LDA_9$dfh
```

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

```r
LDA_9$dfe
```

```
## [1] 95
```

```r
LDA_9$pct
```

```
## [1]  1.000e+02  7.567e-14  1.517e-14 -1.359e-14
```

```r
LDA_9$ndim
```

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

```r
LDA_9$coeffs.raw
```

```
##              Can1
## P_Verbal  0.10465
## P_Quant   0.01513
## C_Verbal -0.06268
## C_Quant   0.04042
```

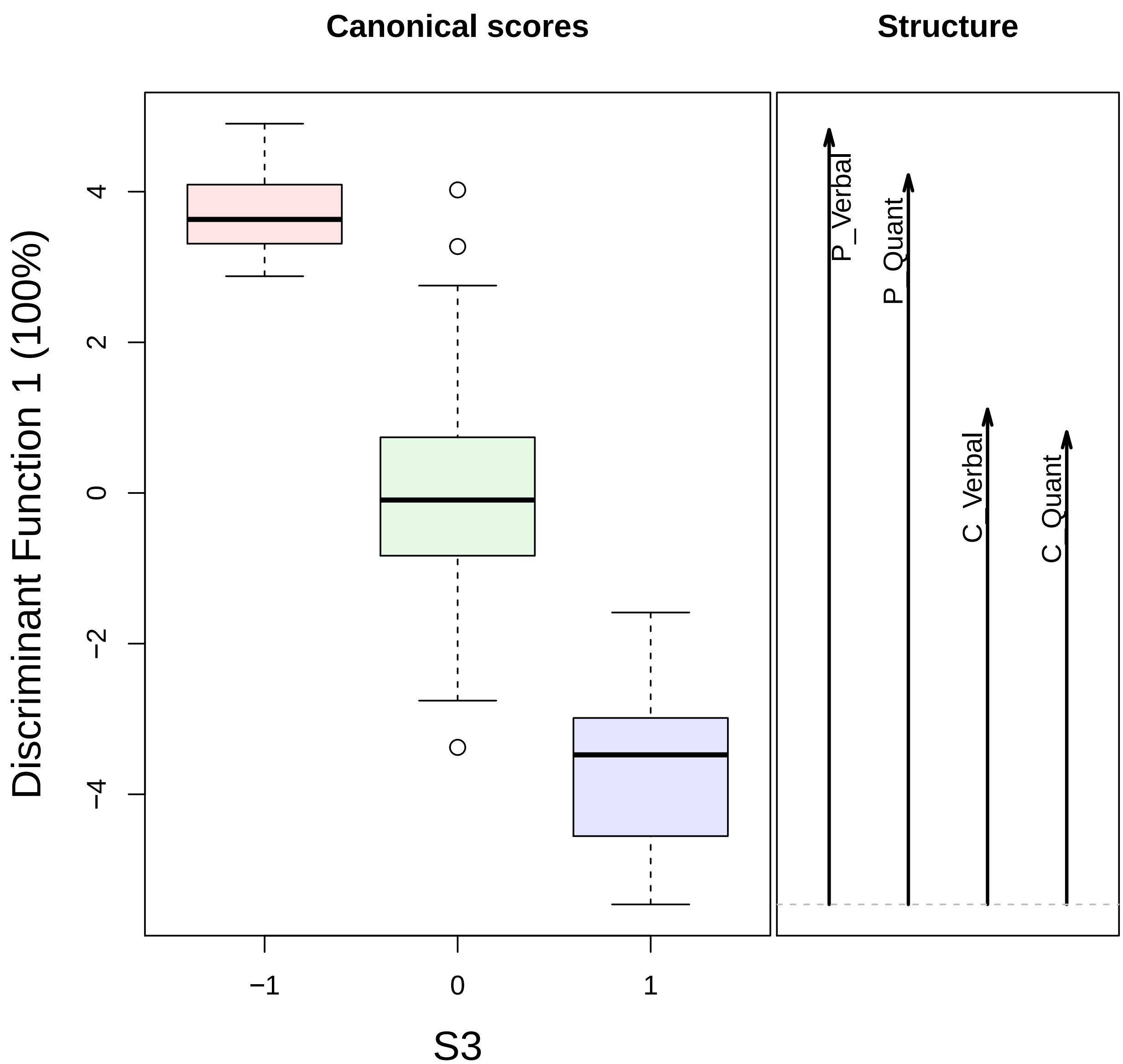
```r
LDA_9$coeffs.std
```

```
##             Can1
## P_Verbal  1.0273
## P_Quant   0.1400
## C_Verbal -0.6316
## C_Quant   0.3666
```

```r
LDA_9$structure
```

```
##            Can1
## P_Verbal 0.9923
## P_Quant  0.9343
## C_Verbal 0.6342
## C_Quant  0.6052
```

```r
plot(LDA_9, cex = 1.25, font.axis = 2, font.lab = 2, cex.lab = 1.5,
    prefix = "Discriminant Function ", var.col = "black", var.lwd = 2)
```



8.2.4 Dummy Codes

```
LM_10 <- lm(cbind(P_Verbal, P_Quant, C_Verbal, C_Quant) ~ D1 + D2 +
    D3, data = Skills_Trimmed)
LDA_10 <- candisc(LM_10, term = "D3", data = Skills_Trimmed, type = "2")
LDA_10

##
## Canonical Discriminant Analysis for D3:
##
##    CanRsq Eigenvalue Difference Percent Cumulative
## 1  0.876       7.04                 100        100
##
```

```
## Test of H0: The canonical correlations in the
## current row and all that follow are zero
##
##   LR test stat approx F numDF denDF Pr(> F)
## 1        0.124      166     4    94  <2e-16
```

```r
summary(LDA_10)
```

```
##
## Canonical Discriminant Analysis for D3:
##
##   CanRsq Eigenvalue Difference Percent Cumulative
## 1 0.8757      7.042                100        100
##
## Class means:
##
## [1] -1.220  3.612
##
##  std coefficients:
## P_Verbal  P_Quant C_Verbal  C_Quant
##  -1.0273  -0.1400   0.6316  -0.3666
```

```r
LDA_10$dfh
```

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

```r
LDA_10$dfe
```

```
## [1] 95
```

```r
LDA_10$pct
```

```
## [1]  1.000e+02  7.567e-14  9.459e-15 -2.049e-14
```

```r
LDA_10$ndim
```

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

```r
LDA_10$coeffs.raw
```

```
##              Can1
## P_Verbal -0.10465
## P_Quant  -0.01513
## C_Verbal  0.06268
## C_Quant  -0.04042
```

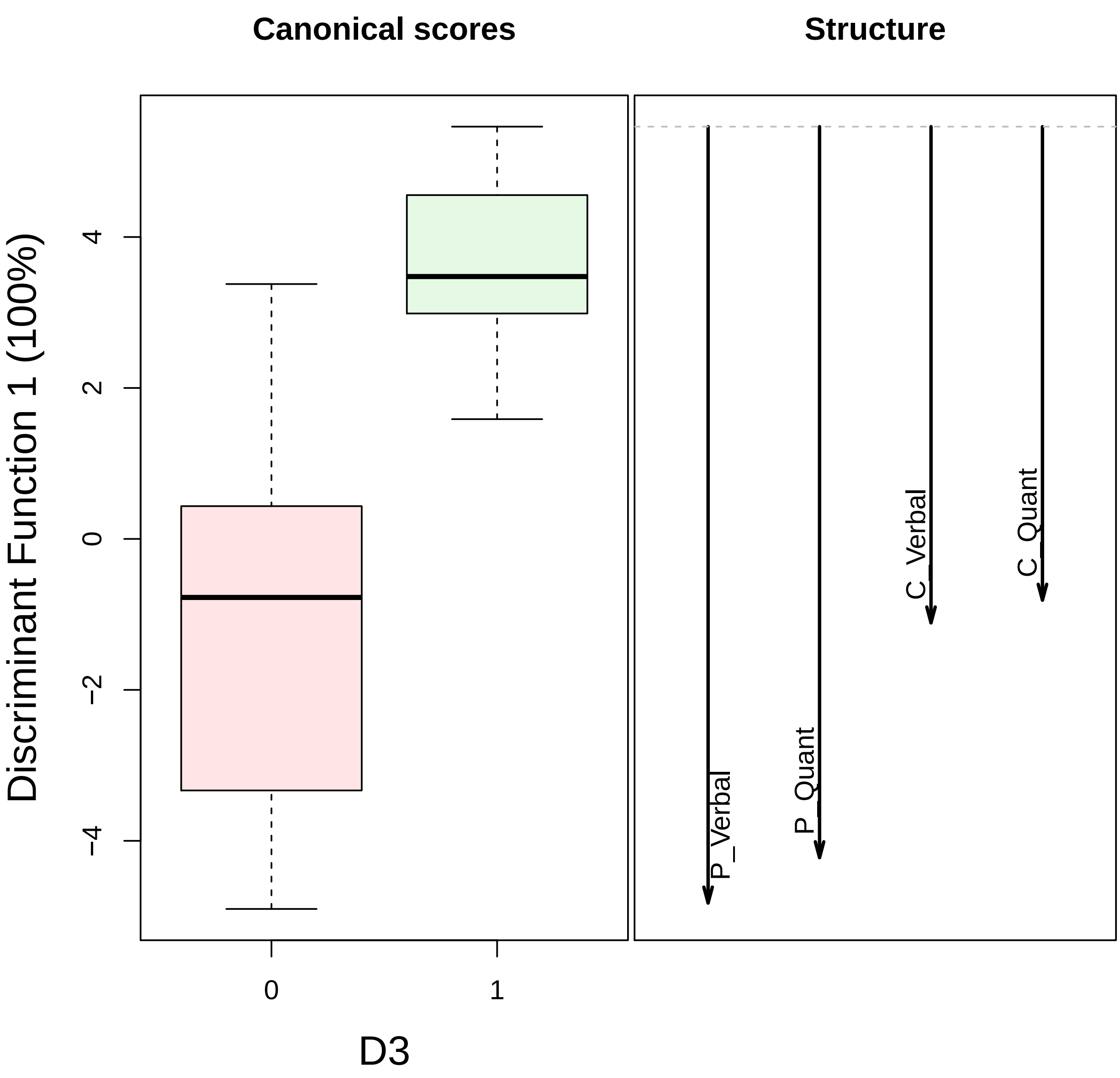
```r
LDA_10$coeffs.std
```

```
##             Can1
## P_Verbal -1.0273
## P_Quant  -0.1400
## C_Verbal  0.6316
## C_Quant  -0.3666
```

```r
LDA_10$structure
```

```
##              Can1
## P_Verbal -0.9923
## P_Quant  -0.9343
## C_Verbal -0.6342
## C_Quant  -0.6052
```

```
plot(LDA_10, cex = 1.25, font.axis = 2, font.lab = 2, cex.lab = 1.5,
    prefix = "Discriminant Function ", var.col = "black", var.lwd = 2)
```




```
LM_11 <- lm(cbind(P_Verbal, P_Quant, C_Verbal, C_Quant) ~ D1 + D2 +
    D3, data = Skills_Trimmed)
LDA_11 <- candisc(LM_11, term = "D1", data = Skills_Trimmed, type = "2")
LDA_11
```

```
##
## Canonical Discriminant Analysis for D1:
##
##   CanRsq Eigenvalue Difference Percent Cumulative
## 1  0.816       4.44                100        100
##
## Test of H0: The canonical correlations in the
## current row and all that follow are zero
##
##   LR test stat approx F numDF denDF Pr(> F)
## 1        0.184      104     4    94  <2e-16
```

```
summary(LDA_11)
```

```
##
## Canonical Discriminant Analysis for D1:
##
##   CanRsq Eigenvalue Difference Percent Cumulative
## 1 0.8161      4.438                100        100
##
## Class means:
##
## [1] -0.5626  1.6654
##
##  std coefficients:
## P_Verbal  P_Quant C_Verbal  C_Quant
##  -0.6897   0.1663   0.5693  -1.0762
```

```
LDA_11$dfh
```

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

```
LDA_11$dfe
```

```
## [1] 95
```

```
LDA_11$pct
```

```
## [1]  1.000e+02  6.004e-14  4.003e-15 -1.151e-14
```

```
LDA_11$ndim
```

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

```
LDA_11$coeffs.raw
```

```
##              Can1
## P_Verbal -0.07027
## P_Quant   0.01796
## C_Verbal  0.05650
## C_Quant  -0.11867
```

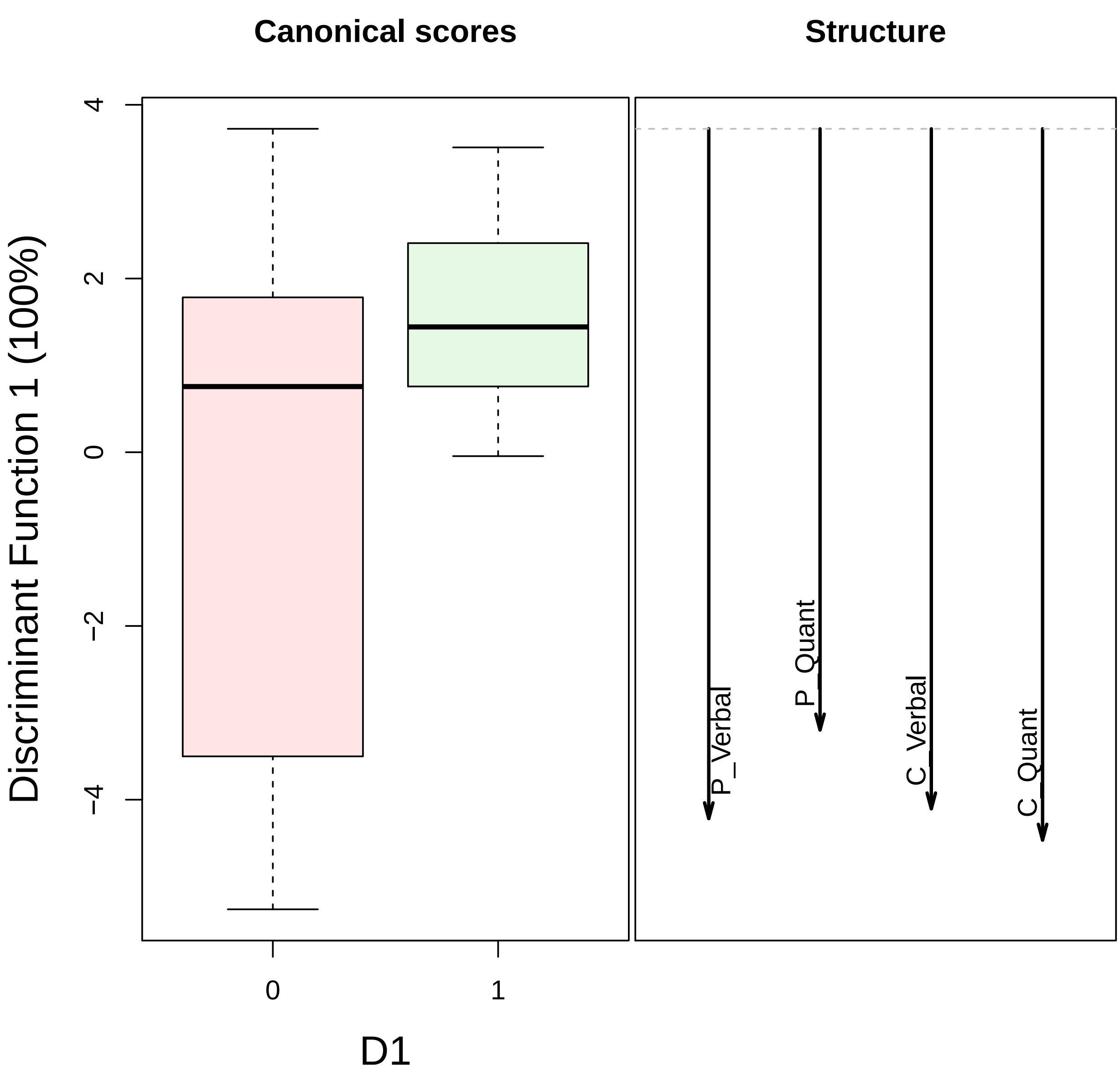
```
LDA_11$coeffs.std
```

```
##             Can1
## P_Verbal -0.6897
## P_Quant   0.1663
## C_Verbal  0.5693
## C_Quant  -1.0762
```

```r
LDA_11$structure
```

```
##              Can1
## P_Verbal -0.8838
## P_Quant  -0.7702
## C_Verbal -0.8712
## C_Quant  -0.9113
```

```r
plot(LDA_11, cex = 1.25, font.axis = 2, font.lab = 2, cex.lab = 1.5,
    prefix = "Discriminant Function ", var.col = "black", var.lwd = 2)
```



9 Univariate Repeated Measures with Sphericity Test

This method, however, does produce the univariate repeated measures F tests. They are based on separate error terms for each within-subjects effect. Note that a test of sphericity is not given because all within-subjects effects are 1 degree of freedom.

```
Mode <- factor(rep(c("Paper", "Computer"), c(2, 2)), levels = c("Paper",
    "Computer"))
Domain <- factor(rep(c("Verbal", "Quant"), 2), levels = c("Verbal",
    "Quant"))
idata <- data.frame(Mode, Domain)

LM_6 <- lm(cbind(P_Verbal, P_Quant, C_Verbal, C_Quant) ~ Tx_P * Tx_C,
    data = Skills_Trimmed)
LM_6

##
## Call:
## lm(formula = cbind(P_Verbal, P_Quant, C_Verbal, C_Quant) ~ Tx_P *
##     Tx_C, data = Skills_Trimmed)
##
## Coefficients:
##              P_Verbal  P_Quant  C_Verbal  C_Quant
## (Intercept)   56.94     58.84    57.68     62.29
## Tx_P1        -20.93    -18.69    -8.14     -8.84
## Tx_C1         -1.37      1.49   -10.22    -13.77
## Tx_P1:Tx_C1   13.21      5.88     6.40      6.61

ANOVA_1 <- Anova(LM_6, idata = idata, idesign = ~Mode * Domain, type = 2)
ANOVA_2 <- Anova(LM_6, idata = idata, idesign = ~Mode * Domain, type = 3)
summary(ANOVA_1, multivariate = FALSE)

##
## Univariate Type II Repeated-Measures ANOVA Assuming Sphericity
##
##                          Sum Sq num Df Error SS den Df F value
## (Intercept)             1375522      1    26463     95 4937.99
## Tx_P                      80207      1    26463     95  287.94
## Tx_C                      13712      1    26463     95   49.22
## Tx_P:Tx_C                 25488      1    26463     95   91.50
## Mode                        522      1     5005     95    9.90
## Tx_P:Mode                 12765      1     5005     95  242.29
## Tx_C:Mode                 14463      1     5005     95  274.53
## Tx_P:Tx_C:Mode              916      1     5005     95   17.38
## Domain                     1031      1     1541     95   63.55
## Tx_P:Domain                  64      1     1541     95    3.97
## Tx_C:Domain                  15      1     1541     95    0.92
## Tx_P:Tx_C:Domain           1252      1     1541     95   77.17
## Mode:Domain                 197      1     1743     95   10.71
## Tx_P:Mode:Domain            225      1     1743     95   12.25
## Tx_C:Mode:Domain            991      1     1743     95   54.02
## Tx_P:Tx_C:Mode:Domain      1407      1     1743     95   76.66
##                          Pr(>F)
## (Intercept)             < 2e-16
## Tx_P                    < 2e-16
```

```
## Tx_C                  3.3e-10
## Tx_P:Tx_C             1.4e-15
## Mode                  0.00221
## Tx_P:Mode             < 2e-16
## Tx_C:Mode             < 2e-16
## Tx_P:Tx_C:Mode        6.8e-05
## Domain                3.5e-12
## Tx_P:Domain           0.04927
## Tx_C:Domain           0.33939
## Tx_P:Tx_C:Domain      6.5e-14
## Mode:Domain           0.00149
## Tx_P:Mode:Domain      0.00071
## Tx_C:Mode:Domain      6.9e-11
## Tx_P:Tx_C:Mode:Domain 7.5e-14
```

```
summary(ANOVA_2, multivariate = FALSE)
```

```
##
## Univariate Type III Repeated-Measures ANOVA Assuming Sphericity
##
##                        Sum Sq num Df Error SS den Df F value
## (Intercept)           1375167      1    26463     95 4936.71
## Tx_P                    79269      1    26463     95  284.57
## Tx_C                    14098      1    26463     95   50.61
## Tx_P:Tx_C               25488      1    26463     95   91.50
## Mode                      432      1     5005     95    8.20
## Tx_P:Mode               12693      1     5005     95  240.93
## Tx_C:Mode               14386      1     5005     95  273.07
## Tx_P:Tx_C:Mode            916      1     5005     95   17.38
## Domain                   1047      1     1541     95   64.50
## Tx_P:Domain                59      1     1541     95    3.61
## Tx_C:Domain                12      1     1541     95    0.76
## Tx_P:Tx_C:Domain         1252      1     1541     95   77.17
## Mode:Domain               181      1     1743     95    9.88
## Tx_P:Mode:Domain          213      1     1743     95   11.63
## Tx_C:Mode:Domain         1016      1     1743     95   55.35
## Tx_P:Tx_C:Mode:Domain    1407      1     1743     95   76.66
##                        Pr(>F)
## (Intercept)           < 2e-16
## Tx_P                  < 2e-16
## Tx_C                  2.1e-10
## Tx_P:Tx_C             1.4e-15
## Mode                  0.00514
## Tx_P:Mode             < 2e-16
## Tx_C:Mode             < 2e-16
## Tx_P:Tx_C:Mode        6.8e-05
## Domain                2.6e-12
## Tx_P:Domain           0.06032
## Tx_C:Domain           0.38672
## Tx_P:Tx_C:Domain      6.5e-14
## Mode:Domain           0.00223
## Tx_P:Mode:Domain      0.00095
## Tx_C:Mode:Domain      4.5e-11
## Tx_P:Tx_C:Mode:Domain 7.5e-14
```

In this version, the factorial structure on the within-subjects side is ignored. Now the 3 degrees of freedom for the within-subjects effect require the sphericity assumption and that test is provided.

```
Measure <- factor(c("P_V", "P_Q", "C_V", "C_Q"), levels = c("P_V",
    "P_Q", "C_V", "C_Q"))
idata <- data.frame(Measure)

LM_7 <- lm(cbind(P_Verbal, P_Quant, C_Verbal, C_Quant) ~ Tx_P * Tx_C,
    data = Skills_Trimmed)
LM_7

##
## Call:
## lm(formula = cbind(P_Verbal, P_Quant, C_Verbal, C_Quant) ~ Tx_P *
##     Tx_C, data = Skills_Trimmed)
##
## Coefficients:
##              P_Verbal  P_Quant  C_Verbal  C_Quant
## (Intercept)   56.94     58.84    57.68     62.29
## Tx_P1        -20.93    -18.69    -8.14     -8.84
## Tx_C1         -1.37      1.49   -10.22    -13.77
## Tx_P1:Tx_C1   13.21      5.88     6.40      6.61

ANOVA_3 <- Anova(LM_7, idata = idata, idesign = ~Measure, type = 2)
ANOVA_4 <- Anova(LM_7, idata = idata, idesign = ~Measure, type = 3)
summary(ANOVA_3, multivariate = FALSE)

##
## Univariate Type II Repeated-Measures ANOVA Assuming Sphericity
##
##                     Sum Sq num Df Error SS den Df F value  Pr(>F)
## (Intercept)        1375522      1    26463     95  4938.0 < 2e-16
## Tx_P                 80207      1    26463     95   287.9 < 2e-16
## Tx_C                 13712      1    26463     95    49.2 3.3e-10
## Tx_P:Tx_C            25488      1    26463     95    91.5 1.4e-15
## Measure               1749      3     8290    285    20.1 8.1e-12
## Tx_P:Measure         13054      3     8290    285   149.6 < 2e-16
## Tx_C:Measure         15469      3     8290    285   177.3 < 2e-16
## Tx_P:Tx_C:Measure     3575      3     8290    285    41.0 < 2e-16
##
##
## Mauchly Tests for Sphericity
##
##                     Test statistic  p-value
## Measure                      0.577 6.59e-10
## Tx_P:Measure                 0.577 6.59e-10
## Tx_C:Measure                 0.577 6.59e-10
## Tx_P:Tx_C:Measure            0.577 6.59e-10
##
##
## Greenhouse-Geisser and Huynh-Feldt Corrections
##  for Departure from Sphericity
##
##                     GG eps Pr(>F[GG])
```

```
## Measure             0.717    4.2e-09
## Tx_P:Measure        0.717    < 2e-16
## Tx_C:Measure        0.717    < 2e-16
## Tx_P:Tx_C:Measure  0.717    < 2e-16
##
##                    HF eps Pr(>F[HF])
## Measure            0.7339  2.853e-09
## Tx_P:Measure       0.7339  1.779e-43
## Tx_C:Measure       0.7339  2.420e-48
## Tx_P:Tx_C:Measure 0.7339  7.766e-17
```

```
summary(ANOVA_4, multivariate = FALSE)
```

```
##
## Univariate Type III Repeated-Measures ANOVA Assuming Sphericity
##
##                     Sum Sq num Df Error SS den Df F value  Pr(>F)
## (Intercept)       1375167      1    26463     95  4936.7 < 2e-16
## Tx_P                79269      1    26463     95   284.6 < 2e-16
## Tx_C                14098      1    26463     95    50.6 2.1e-10
## Tx_P:Tx_C           25488      1    26463     95    91.5 1.4e-15
## Measure              1660      3     8290    285    19.0 2.8e-11
## Tx_P:Measure        12965      3     8290    285   148.6 < 2e-16
## Tx_C:Measure        15414      3     8290    285   176.7 < 2e-16
## Tx_P:Tx_C:Measure    3575      3     8290    285    41.0 < 2e-16
##
##
## Mauchly Tests for Sphericity
##
##                    Test statistic  p-value
## Measure                     0.577 6.59e-10
## Tx_P:Measure                0.577 6.59e-10
## Tx_C:Measure                0.577 6.59e-10
## Tx_P:Tx_C:Measure           0.577 6.59e-10
##
##
## Greenhouse-Geisser and Huynh-Feldt Corrections
##  for Departure from Sphericity
##
##                    GG eps Pr(>F[GG])
## Measure             0.717      1e-08
## Tx_P:Measure        0.717     <2e-16
## Tx_C:Measure        0.717     <2e-16
## Tx_P:Tx_C:Measure  0.717     <2e-16
##
##                    HF eps Pr(>F[HF])
## Measure            0.7339  7.222e-09
## Tx_P:Measure       0.7339  2.753e-43
## Tx_C:Measure       0.7339  3.081e-48
## Tx_P:Tx_C:Measure 0.7339  7.766e-17
```

10 Roy-Bargman Step-Down Tests

The redundancy question can be addressed using the Roy-Bargman stepdown procedure. The dependent variables are tested in a univariate fashion, in a specific order, with earlier-considered dependent variables used as covariates for later dependent variables. All dependent variables add significantly to group differentiation.

```
summary(aov(P_Verbal ~ P_Quant + C_Verbal + C_Quant + as.factor(Group),
    data = Skills_Trimmed))

##                  Df Sum Sq Mean Sq F value  Pr(>F)
## P_Quant           1  61297   61297 1649.98 < 2e-16
## C_Verbal          1   1413    1413   38.02 1.8e-08
## C_Quant           1    237     237    6.38   0.013
## as.factor(Group)  3   4196    1399   37.65 5.8e-16
## Residuals        92   3418      37

summary(aov(P_Quant ~ P_Verbal + C_Verbal + C_Quant + as.factor(Group),
    data = Skills_Trimmed))

##                  Df Sum Sq Mean Sq F value  Pr(>F)
## P_Verbal          1  40450   40450 1620.91 < 2e-16
## C_Verbal          1     33      33    1.31 0.25608
## C_Quant           1    344     344   13.77 0.00035
## as.factor(Group)  3   3440    1147   45.95 < 2e-16
## Residuals        92   2296      25

summary(aov(C_Verbal ~ P_Quant + P_Verbal + C_Quant + as.factor(Group),
    data = Skills_Trimmed))

##                  Df Sum Sq Mean Sq F value  Pr(>F)
## P_Quant           1  12328   12328  501.99 < 2e-16
## P_Verbal          1   2806    2806  114.26 < 2e-16
## C_Quant           1  12891   12891  524.90 < 2e-16
## as.factor(Group)  3    444     148    6.03 0.00085
## Residuals        92   2259      25

summary(aov(C_Quant ~ P_Quant + C_Verbal + P_Verbal + as.factor(Group),
    data = Skills_Trimmed))

##                  Df Sum Sq Mean Sq F value  Pr(>F)
## P_Quant           1  11863   11863  469.76 < 2e-16
## C_Verbal          1  23005   23005  910.98 < 2e-16
## P_Verbal          1    120     120    4.73   0.032
## as.factor(Group)  3   1519     506   20.06 4.4e-10
## Residuals        92   2323      25
```

```
Sys.time() - how_long

## Time difference of 22.25 secs
```