Multivariate Analysis of Variance

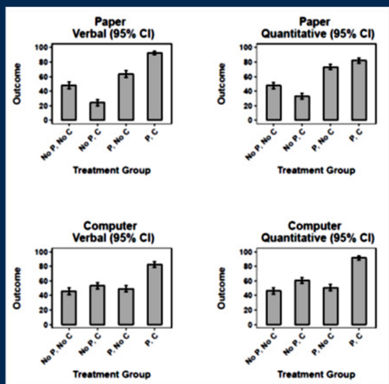
Today . . .

- MANOVA as a way to reveal new information about group differences
- MANOVA as a way to overcome problems with a univariate approach

Within-Subjects

Between-Groups

	Standard Measure of Verbal Ability	Standard Measure of Quantitative Ability	Computer Measure of Verbal Ability	Computer Measure of Quantitative Ability
No Training				
Standard Training Only				
Computer Training Only				
Both				



The four groups exhibited significant differences on each of the four measures.

Question

Measure	Computer Measure?	Quantitative Ability?	Both or Neither?
Standard Measure of Verbal	-1	-1	1
Standard Measure of Quantitative	-1	1	-1
Computer Measure of Verbal	1	-1	-1
Computer Measure of Quantitative	1	1	1

The factorial structure underlying the repeated measures represents one way to partition the variability in the repeated measures. It also represents a set of weighted linear combinations that separate the groups. But are they the best linear combinations? Could a different set of weights do a better job?

Questions

Groups	Computer Training?	Standard Training?	Both or None?
No Training	-1	-1	1
Standard Training	-1	1	-1
Computer Training	1	-1	-1
Standard and Computer Training	1	1	1

The factorial structure underlying the groups likewise partitions the variability in the group means. It also creates the simplest possible group distinctions—a set of independent two-group comparisons. But, are these the most interesting comparisons? Could the groups differ in other important ways?

We could relax the factorial structure and ask, "Is there a linear combination of the outcomes that produces the best or maximum separation of the groups?"

$$LC_i = W_1X_{1i} + W_2X_{2i} + \ . \ . \ . \ + W_kX_{ki}$$

$$\underbrace{\overline{LC}_{G_1} \quad \overline{LC}_{G_2} \quad \overline{LC}_{G_3} \quad . \ . \ . \quad \overline{LC}_{G_K}}$$

$$Maximize \ \sigma^2_{\overline{LC}}$$

Once that linear combination is found, is there another one—independent of the first one—that produces maximum separation of the groups?

Are there subsequent linear combinations, independent of previous ones, that provide maximum group separation?

A MANOVA (just a discriminant analysis) can answer these questions.

The simplest and least restrictive discriminant analysis imposes no structure on the groups and no structure on the measures. It represents a standard discriminant analysis applied to groups that happen to have been created by the researcher.

It is the most exploratory multivariate approach that we can take with these data, aimed at discovering (rather than imposing) how the groups can best be separated with weighted linear combinations of the measures. The candisc( ) function from the candisc package is particularly useful.

```
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)
```

```
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
```

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

Three discriminant functions are possible with four groups and four measures; all three produce significant discrimination among the groups. The first two account for 95% of the group separation.

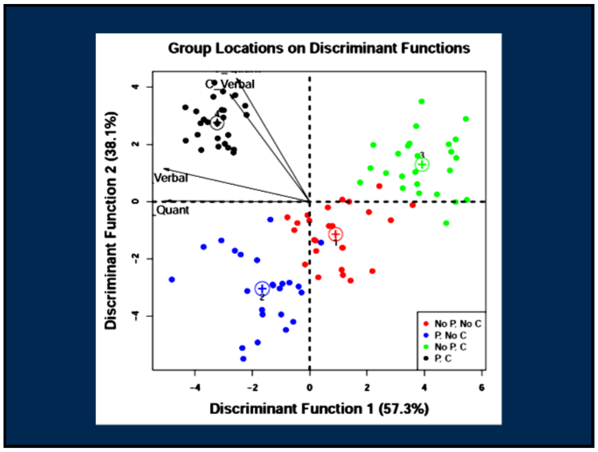
```
LDA_1$coeffs.raw

##              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

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
```

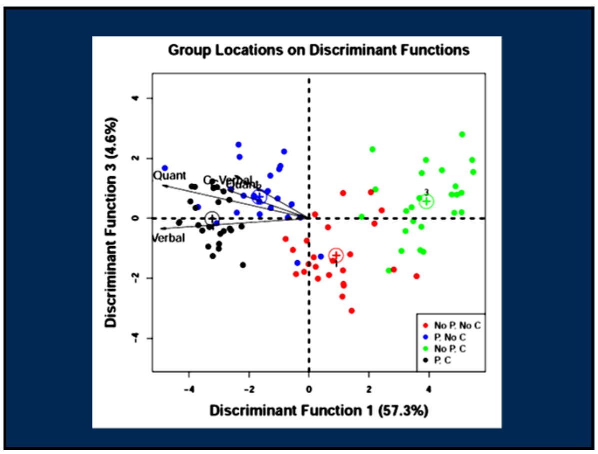
The raw coefficients are used to generate function scores. The standardized coefficients and structure coefficients (correlations) are used to determine the nature of the linear combinations.

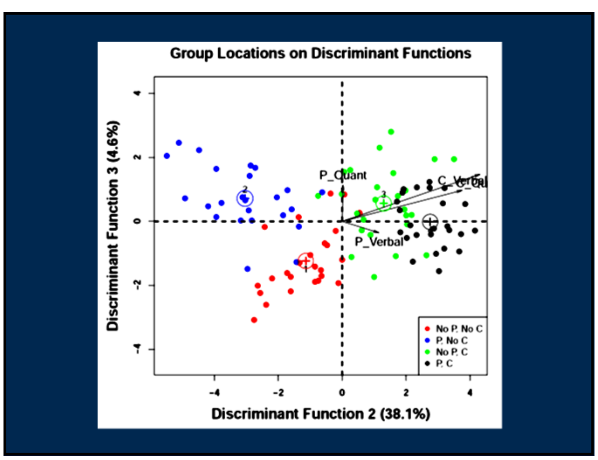
```
LDA_1$coeffs.std

##             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
```

```
Class means:

    Can1   Can2     Can3
1  0.903 -1.137 -1.23762   No P, No C
2 -1.655 -3.040  0.71859   P, No C
3  3.916  1.301  0.56375   No P, C
4 -3.230  2.755 -0.01598   P, C
```

The first function is dominated by paper measures; the second is dominated by computer measures. The first function has the greatest separation for the computer trained groups that differed in whether they received training with standard paper measures. The second function separates standard trained groups that differed in whether they received computer training.


Group Locations on Discriminant Functions
C_Verbal
Verbal
Quant
Discriminant Function 2 (38.1%)
Discriminant Function 1 (57.3%)
No P, No C
P, No C
No P, C
P, C


Group Locations on Discriminant Functions
Quant
C_Verbal
Verbal
Discriminant Function 3 (4.6%)
Discriminant Function 1 (57.3%)
No P, No C
P, No C
No P, C
P, C


Group Locations on Discriminant Functions
P_Quant
P_Verbal
Discriminant Function 3 (4.6%)
Discriminant Function 2 (38.1%)
No P, No C
P, No C
No P, C
P, C

The manova( ) function from the stats package can provide some additional information. 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)$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
```

H (or B)

E (or W)

These are the two major matrices that are used in the calculation of the significance tests.

Different combinations of **H** and **E** define the four most common tests of significance:

- Wilk's likelihood test is based on Λ = |E|/|H+E|
- Pillai's Trace is a function of tr(H[H+E]$^{-1}$)
- Lawley and Hotelling's Trace = tr(HE^{-1})
- Roy's test is a function of the largest eigenvalue of HE^{-1}

The first three usually agree because they use much the same information.

The four can also be defined in terms of the eigenvalues:

$$Wilks'\ \Lambda = \prod_{d=1}^{D} \frac{1}{1+\lambda_d}$$

$$Pillai's\ Trace = \sum_{d=1}^{D} \frac{\lambda_d}{1+\lambda_d}$$

$$Lawley - Hotelling\ Trace = \sum_{d=1}^{D} \lambda_d$$

$$Roy's\ Largest\ Root = \lambda_1$$

The Manova( ) function (note capitalization) from the car package 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 and all four significance tests.

```
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")
```

```
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
```

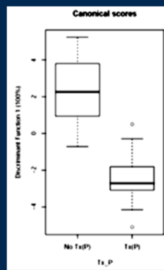
The discriminant analysis can also be performed on "groups" defined by user-chosen weights. These might represent a factorial structure or other comparisons of interest.

```
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")
```

Here we use the original factorial structure for the groups, but determine the weights that best separate the groups defined by that structure. In this case, we isolate the main effect for standard paper-and-pencil training.

```
 LR test stat approx F num Df den Df Pr(> F)
1        0.137      601      1     95  <2e-16
```

```
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
```



A significant separation of the two groups is possible and the linear combination is dominated by the paper measures.

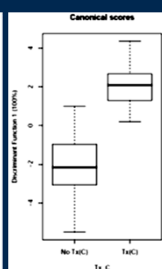
```
LM_3 <- lm(cbind(P_Verbal, P_Quant, C_Verbal, C_Quant) ~ Tx_P + Tx_C +
    Tx_P:Tx_C, data = Skills_Trimmed)
LDA_3 <- candisc(LM_3, term = "Tx_C", data = Skills_Trimmed, type = "2")
```

This analysis identifies the best way to separate the groups that either received computer training or not.

```
 LR test stat approx F num Df den Df Pr(> F)
1        0.179      437      1     95  <2e-16
```

These groups can be significantly separated by a linear combination dominated by computer measures.

```
LDA_3$coeffs.std

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

LDA_3$structure

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




```
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
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
```

Other “groups” can be defined with special contrasts or dummy codes.

```
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
```

```r
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")
```

This analysis identifies the linear combination that best separates treated participants from the untreated controls.

```
  LR test stat approx F num Df den Df Pr(> F)
1        0.441      120      1     95  <2e-16
```

Treatment of any kind improves performance, but especially on computer-based measures.

```
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
```

```r
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")
```

This analysis uses dummy codes with the reference group defined as participants who received both kinds of training. Here we identify the linear combination that best separates this group from participants who received computer training only.

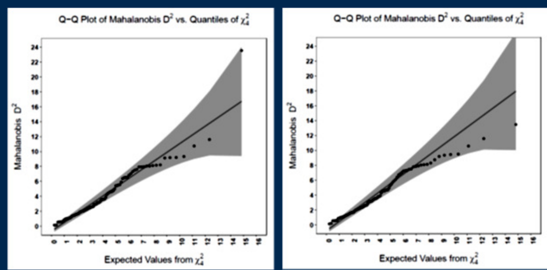
```
  LR test stat approx F num Df den Df Pr(> F)
1        0.124      669      1     95  <2e-16
```

The additional standard training for the reference group is apparent in the linear combination dominated by paper measures.

```
LDA_10$coeffs.std

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

LDA_10$structure
##              Can1
## P_Verbal  -0.9923
## P_Quant   -0.9343
## C_Verbal  -0.6342
## C_Quant   -0.6052
```

MANOVA (and discriminant analysis) has assumptions that need to be checked:

- Multivariate normality
- Homogeneity of variance-covariance matrices
- Absence of outliers



We eliminated a multivariate outlier.

Outlier influence on normality:

Full sample

```
$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
```

Outlier excluded

```
$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
```

```
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
```

```
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
```

The homogeneity assumption fails, but this test is notoriously sensitive and MANOVA is generally robust to this violation, especially as sample size increases.

The multivariate approach can add to the univariate analyses by addressing some additional questions.

- With multiple outcomes, there is a danger of redundant findings. Which outcomes are most important? Do all add to the analyses in a meaningful way?
- Repeated measures ANOVA assumes sphericity. If this assumption is not met, the Type I error rate could be inflated. One remedy is adjustment of the degrees of freedom. Another is MANOVA, which does not have a sphericity assumption.

```
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
```

The multiple outcomes are highly related, especially the different abilities measured by the same method. These correlations underscore the redundancy in multiple univariate tests. MANOVA can determine how many independent ways the groups can actually be distinguished.

```
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

  Chi_Sq df          p
1 418.68 12 4.180e-82
2 215.44  6 9.737e-44
3  45.19  2 1.541e-10
```

MANOVA is sometimes used to justify the individual ANOVAs. The rationale for using this omnibus test as a Type I error protection approach is that included among the possible linear combinations are those in which each outcome variable is the only variable receiving a weight.

The redundancy question can also be addressed using the Roy-Bargman step-down procedure. The dependent variables are tested in a univariate fashion, with other dependent variables used as covariates in an analysis of a target dependent variables.

Dependent variables that significantly distinguish groups after controlling for other dependent variables are adding uniquely 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
```

Each outcome measure has a unique contribution.

```
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)
```

```
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

                       SS num Df Error SS den Df      F  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
```

When the repeated measures have more than 1 df, sphericity becomes an important assumption.

```
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
```

If sphericity fails, the F tests can be adjusted to be more conservative.

MANOVA is another solution.

Next time . . .

Additional extensions and issues