# Example of a Mixed Factorial ANOVA

A 3 × 3 (Lecture Type × Time) mixed between and within factorial analysis of vocabulary scores[1]

## Syntax

```
*these first lines of syntax compute the mean for the three measurements just to get the main effect mean for
descriptive purposes. It is not needed for the analyses.
COMPUTE vocab=mean(baseline,twowks,fourwks).
MEANS VARS=vocab BY lecture.
MEANS VARS=baseline twowks fourwks by lecture.

GLM  baseline twowks fourwks BY lecture
  /WSFACTOR=time 3
  /WSDESIGN=time
  /DESIGN=lecture
  /PRINT=DESCRIPTIVE ETASQ
  /EMMEANS=tables(lecture)
 /EMMEANS=TABLES(time*lecture)
  /PLOT=PROFILE(time*lecture).
```

## Menus

**Analyze → General Linear Model → Repeated Measures**

Enter number of levels of within-subjects factor and name it. Click **Add**. Click the **Define** button. Drag over the variables for the within-subjects factor (here, `baseline`, `twowks`, and `fourwks`). Move over **Between-Subjects Factor**.

**Report**

vocab

| lecture type of lecture | Mean | N | Std. Deviation |
|---|---|---|---|
| 1.000 physical science | 40.0000 | 4 | 5.23521 |
| 2.000 social science | 26.0000 | 4 | 7.10764 |
| 3.000 history | 34.5000 | 4 | 3.15642 |
| Total | 33.5000 | 12 | 7.75574 |

**Report**

| lecture type of lecture | | baseline | twowks | fourwks |
|---|---|---|---|---|
| 1.000 physical science | Mean | 47.75000 | 44.25000 | 28.00000 |
| | N | 4 | 4 | 4 |
| | Std. Deviation | 4.573474 | 7.410578 | 7.438638 |
| 2.000 social science | Mean | 41.25000 | 26.00000 | 10.75000 |
| | N | 4 | 4 | 4 |
| | Std. Deviation | 4.349329 | 13.832329 | 4.112988 |
| 3.000 history | Mean | 40.00000 | 38.50000 | 25.00000 |
| | N | 4 | 4 | 4 |
| | Std. Deviation | 3.915780 | 5.802298 | 5.228129 |
| Total | Mean | 43.00000 | 36.25000 | 21.25000 |
| | N | 12 | 12 | 12 |
| | Std. Deviation | 5.257030 | 11.817745 | 9.430367 |

[1] This numeric example is adapted from Keppel, G., & Zedeck S. (1989). *Data analysis for research designs*. New York: Freeman. Explanation of the mixed factorial using this example is in the handout "Factorial ANOVA for Mixed Designs Example: SPSS and R" for this class.

**Mauchly's Test of Sphericity** [a]

Measure: MEASURE_1

| Within Subjects Effect | Mauchly's W | Approx. Chi-Square | df | Sig. | Epsilon [b] | | |
|---|---|---|---|---|---|---|---|
| | | | | | Greenhouse-Geisser | Huynh-Feldt | Lower-bound |
| time | .691 | 2.952 | 2 | .229 | .764 | 1.000 | .500 |

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. Design: Intercept + lecture
Within Subjects Design: time

b. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

**Multivariate Tests** [a]

| Effect | | Value | F | Hypothesis df | Error df | Sig. | Partial Eta Squared |
|---|---|---|---|---|---|---|---|
| time | Pillai's Trace | .963 | 104.309 [b] | 2.000 | 8.000 | .000 | .963 |
| | Wilks' Lambda | .037 | 104.309 [b] | 2.000 | 8.000 | .000 | .963 |
| | Hotelling's Trace | 26.077 | 104.309 [b] | 2.000 | 8.000 | .000 | .963 |
| | Roy's Largest Root | 26.077 | 104.309 [b] | 2.000 | 8.000 | .000 | .963 |
| time * lecture | Pillai's Trace | .683 | 2.335 | 4.000 | 18.000 | .095 | .342 |
| | Wilks' Lambda | .332 | 2.942 [b] | 4.000 | 16.000 | .053 | .424 |
| | Hotelling's Trace | 1.966 | 3.441 | 4.000 | 14.000 | .037 | .496 |
| | Roy's Largest Root | 1.942 | 8.741 [c] | 2.000 | 9.000 | .008 | .660 |

a. Design: Intercept + lecture
Within Subjects Design: time

b. Exact statistic

c. The statistic is an upper bound on F that yields a lower bound on the significance level.

**Tests of Within-Subjects Effects**

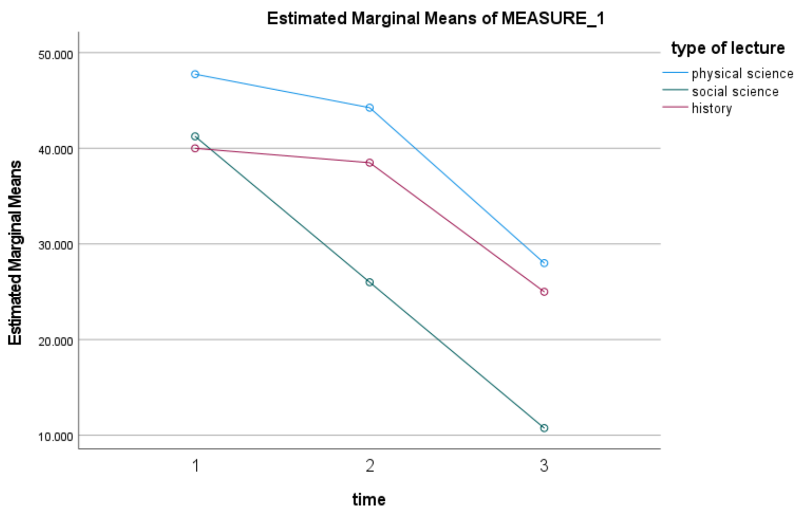
Measure: MEASURE_1

| Source | | Type III Sum of Squares | df | Mean Square | F | Sig. | Partial Eta Squared |
|---|---|---|---|---|---|---|---|
| time | Sphericity Assumed | 2974.500 | 2 | 1487.250 | 52.184 | .000 | .853 |
| | Greenhouse-Geisser | 2974.500 | 1.528 | 1946.128 | 52.184 | .000 | .853 |
| | Huynh-Feldt | 2974.500 | 2.000 | 1487.250 | 52.184 | .000 | .853 |
| | Lower-bound | 2974.500 | 1.000 | 2974.500 | 52.184 | .000 | .853 |
| time * lecture | Sphericity Assumed | 320.500 | 4 | 80.125 | 2.811 | .057 | .385 |
| | Greenhouse-Geisser | 320.500 | 3.057 | 104.847 | 2.811 | .078 | .385 |
| | Huynh-Feldt | 320.500 | 4.000 | 80.125 | 2.811 | .057 | .385 |
| | Lower-bound | 320.500 | 2.000 | 160.250 | 2.811 | .113 | .385 |
| Error(time) | Sphericity Assumed | 513.000 | 18 | 28.500 | | | |
| | Greenhouse-Geisser | 513.000 | 13.756 | 37.293 | | | |
| | Huynh-Feldt | 513.000 | 18.000 | 28.500 | | | |
| | Lower-bound | 513.000 | 9.000 | 57.000 | | | |

**Tests of Between-Subjects Effects**

Measure: MEASURE_1
Transformed Variable: Average

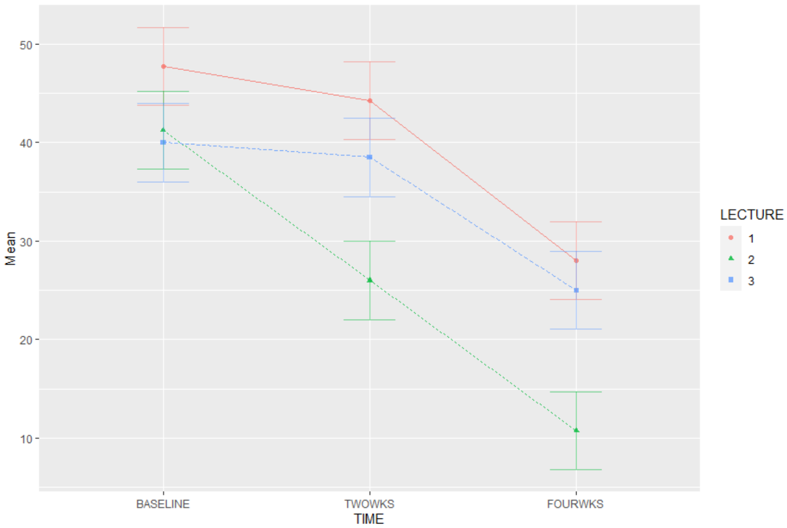
| Source | Type III Sum of Squares | df | Mean Square | F | Sig. | Partial Eta Squared |
|---|---|---|---|---|---|---|
| Intercept | 40401.000 | 1 | 40401.000 | 459.683 | .000 | .981 |
| lecture | 1194.000 | 2 | 597.000 | 6.793 | .016 | .602 |
| Error | 791.000 | 9 | 87.889 | | | |




**R**

```
> cat("\014")#clear console
> if(!is.null(dev.list())) dev.off(dev.list()["RStudioGD"])  #clear plots
> rm(d) #clear data frames
> rm(longdata)
> library(haven)
> d = read_sav("c:/jason/spsswin/uvclass/tabl17-5.sav")

>library(reshape2)
>longdata <- melt(d,
                measure.vars = c("BASELINE", "TWOWKS", "FOURWKS"),  #old variables
                variable.name = "TIME",        #name new variable for the value labels
                value.name = "SCORE")          #name a new variable for the values

> #convert LECTURE to factor, was not need for TIME for these data
> longdata$LECTURE <- factor(longdata$LECTURE)

library('ez')

> mymodel = ezANOVA(data = longdata,
+     dv = SCORE,         #dependent variable
+     wid = id,           #id variable
+     within = TIME,      #within subjects factor
+     between = LECTURE,  #between-subjects factor
+     detailed = TRUE)    #print some extra details
Warning: Converting "id" to factor for ANOVA.  (convert id to factor beforehand or tolerate this warning)
> print(mymodel)

$ANOVA
        Effect DFn DFd     SSn SSd          F                p p<.05       ges
1  (Intercept)   1   9 40401.0 791 459.682680 0.000000004913104     * 0.9687328
2      LECTURE   2   9  1194.0 791   6.792668 0.015917390678695     * 0.4779824
3         TIME   2  18  2974.5 513  52.184211 0.000000032242473     * 0.6952203
4 LECTURE:TIME   4  18   320.5 513   2.811404 0.056533988965637       0.1972915

$`Mauchly's Test for Sphericity`
        Effect         W         p p<.05
3         TIME 0.6914584 0.2285937
4 LECTURE:TIME 0.6914584 0.2285937

$`Sphericity Corrections`
        Effect       GGe           p[GG] p[GG]<.05       HFe           p[HF] p[HF]<.05
3         TIME 0.7642096 0.000001009711         * 0.8889813 0.0000001627447         *
4 LECTURE:TIME 0.7642096 0.077784352860           0.8889813 0.0656228251801


#Get cell means
> descrip = ezStats(data = longdata,
+     dv = SCORE,              #dependent variable scores
+     wid = id,                #id variable
+     within = TIME,           #within-subjects factor
+     between = LECTURE)       #between-subjects factor
> print(descrip)

          LECTURE     TIME N  Mean        SD     FLSD
1 physical science BASELINE 4 47.75  4.573474 7.930807
2 physical science   TWOWKS 4 44.25  7.410578 7.930807
3 physical science  FOURWKS 4 28.00  7.438638 7.930807
4   social science BASELINE 4 41.25  4.349329 7.930807
5   social science   TWOWKS 4 26.00 13.832329 7.930807
6   social science  FOURWKS 4 10.75  4.112988 7.930807
```

```
7          history BASELINE 4 40.00  3.915780 7.930807
8          history   TWOWKS 4 38.50  5.802298 7.930807
9          history  FOURWKS 4 25.00  5.228129 7.930807

#plot the data
#ezplot requires factor rather than double precision (continuous) used by havenl
> mixplot = ezPlot(data = longdata,
+        dv = SCORE,          #dependent variable scores
+        wid = id,            #id variable
+        x = TIME,            #x-axis variable
+        split = LECTURE,     #separate lines
+        within = TIME,       #within factor
+        between = LECTURE) #between factor
> print(mixplot)
```




The sample size was too small to use MANOVA for repeated measures here, but the R code for obtaining the correct results with the `car` package is:

```
> library(car)
#need to create a factor and frame for appropriate number of IV levels, called "time" here
> time <- c(1,2,3)
> time <- as.factor(time)
> condframe <- data.frame(time)

> d$LECTURE <- factor(d$LECTURE) #convert LECTURE to factor--car requires for contrasts command
> model2 <- lm(cbind(BASELINE, TWOWKS, FOURWKS) ~ LECTURE, data=d, contrasts=list(time=contr.sum,
+      LECTURE=contr.sum))
> analysis <- Manova(model2, idata=condframe, idesign=~time, type="III")
> summary(analysis)
```

Obtaining partial eta squared:

```
> library(effectsize)
> eta_squared(analysis)
# Effect Size for ANOVA (Type III)

Parameter     | Eta2 (partial) |       95% CI
-----------------------------------------------
LECTURE       |           0.60 | [0.13, 1.00]
time          |           0.85 | [0.72, 1.00]
LECTURE:time  |           0.38 | [0.00, 1.00]
```

**Example Write-Up**

A 3 (physical science, social science, history) × 3 (baseline, two weeks, four weeks) mixed factorial ANOVA, with one between-subjects factor and one within-subjects factor, was conducted to investigate whether changes in vocabulary over time differed by lecture type. The univariate repeated-measures tests and their sphericity corrections were examined to determine significance because of the small sample size (Algina & Kesselman, 1997). There was a significant main effect for lecture type, $F(2,9) = 6.79$, $p$ =, partial $\eta^2 = .60$, with the highest vocabulary scores in physical science M = 40, SD = 5.24, followed by history (M = 34.50, SD = 3.16), and social science (M = 26.00, SD = 7.11). A significant main effect for time, $F(2,18) = 52.18$, $p < .001$

(for sphericity assumed and corrections), partial $\eta^2$ = .85, indicated that there was a decline in vocabulary scores over time (baseline M = 43.00, SD = 5.26, two weeks 36.25, SD = 11.42, four weeks M = 21.25, SD = 9.42). The interaction did not reach conventional levels of significance for the the sphericity assumed results $F(4,18) = 2.81$, $p = .057$, the Greenhouse-Geisser ($p = .08$), or the Huynh-Feldt corrections ($p = .06$), partial $\eta^2$ = .39. The results indicated that there may be some tendency for vocabulary scores to decline at different rates across lecture types, although caution is warranted in this interpretation because the results may not be statistically reliable.

*Given the very small sample size in this example, the univariate repeated tests of repeated measures are more appropriate. But had we had even 20 or so cases or more, I would have recommended using the multivariate repeated measures test statistics, such as the Pillai's trace, given the Algina-Keselmen guidelines.*

*Several possible follow-up tests might explore the main effects for lecture or time by conducting main effect contrasts. For example, it might of interest to know whether there was a significant difference between baseline and two-week vocabulary scores overall or whether history and social science vocabulary scores differed significantly overall. In addition, because the interaction effect was marginally significant, some authors might choose to conduct simple effect tests, such as exploring whether baseline scores differed among the three lecture types (using one-way ANOVA) or social science scores changed significantly over time (within-subjects ANOVA). Those simple effect tests might be followed by simple contrasts, such as examining whether social science scores differed between baseline and two weeks (paired $t$ test) or whether social and physical science scores differed at four weeks (one-way ANOVA planned contrast).*