Within-Subjects ANOVA

General Comments

As with the rationale for the paired t-test, the within-subjects ANOVA is used under similar circumstances. The within-subjects ANOVA, however, is more general than the paired ("correlated-scores") t-test in that it also can be used with more than two repeated measures. The within-subjects ANOVA is appropriate for repeated measures designs (e.g., pretest-posttest designs), within-subjects experimental designs, matched designs, or multiple measures. It is sometimes difficult to understand that two dependent measures (e.g., the DV at pretest and posttest) function as two “levels” of the independent variable. So, for instance, in the case of the pretest-posttest design the independent variable is “time.”

Within-subjects designs have advantages over between-subjects designs, because, in general, they have greater power to detect significance. The fact that each participant serves as his or her own control (or there is a related other used as a control) leads to the advantage of eliminating variance due specifically to individual differences. Thus, the error term used in within-subjects ANOVA is a more precise one, because individual differences have been removed from it. The separate estimation of variance due to individual differences is explicit in the sum of squares for subject (SS_S). Most of the other general procedures are similar to what we have done before.

Definitional Formulas

The only new quantity in these formulas is SS_S, which is the sum of squares for subject. This is computed by finding the mean score for each case (averaging across rows of the data). SS_s represents individual variation. Individual variation is thus a function of variation of average scores for an individual around the average of all scores for the sample, SS_T.

Source	SS	df	MS	F
A	$SS_A = n\sum\left(\bar{Y}_{.j} - \bar{Y}_{..}\right)^2$	$a-1$	$MS_A = \frac{SS_A}{df_A}$	$F_A = \frac{MS_A}{MS_{AxS}}$
S	$SS_S = a\sum\left(\bar{Y}_{i.} - \bar{Y}_{..}\right)^2$	$n-1$	$MS_S = \frac{SS_S}{df_s}$	
AxS	$SS_{AxS} = \sum\sum\left(Y - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..}\right)^2$ or $SS_{AxS} = SS_T - SS_A - SS_S$	$(a-1)(n-1)$	$MS_{AxS} = \frac{SS_{AxS}}{df_{AxS}}$	
T	$SS_T = \sum\left(Y_{ij} - \bar{Y}_{..}\right)^2$	$(a)(n)$		

Example
Consider a hypothetical example in which eight non-native English-speaking students are tested before and after the implementation of an "immersion" approach to teaching proficiency in English. Thus, we test students on a language usage scale (say with values from 1-10) during the traditional bilingual education program and then after the conversion to immersion.[1]

Student	**Bilingual**	$Y_{ij}-\bar{Y}_{..}$	$\left(Y_{ij}-\bar{Y}_{..}\right)^2$	**Immersion**	$Y_{ij}-\bar{Y}_{..}$	$\left(Y_{ij}-\bar{Y}_{..}\right)^2$	$\bar{Y}_{i.}$	$\bar{Y}_{i.}-\bar{Y}_{..}$	$\left(\bar{Y}_{i.}-\bar{Y}_{..}\right)^2$
1	4	.5	.25	2	-1.5	2.25	3	-.5	.25
2	3	-.5	.25	3	-.5	.25	3	-.5	.25
3	6	2.5	6.25	4	.5	.25	5	1.5	2.25
4	5	1.5	2.25	5	1.5	2.25	5	1.5	2.25
5	6	2.5	6.25	4	.5	.25	5	1.5	2.25
6	3	-.5	.25	3	-.5	.25	3	-.5	.25
7	2	-1.5	2.25	2	-1.5	2.25	2	-1.5	2.25
8	3	-.5	.25	1	-2.5	6.25	2	-1.5	2.25
	$\bar{Y}_{.1}=4$			$\bar{Y}_{.2}=3$					$\sum\left(\bar{Y}_{i.}-\bar{Y}_{..}\right)^2=12$

$\bar{Y}_{..}=3.5$

$$SS_A=n\sum\left(\bar{Y}_{.j}-\bar{Y}_{..}\right)^2=8\left[(4-3.5)^2+(3-3.5)^2\right]=4$$

$$SS_S=a\sum\sum\left(\bar{Y}_{i.}-\bar{Y}_{..}\right)^2=2[.25+.25+2.25+\ldots+2.25+2.25]=24$$

$$SS_{AxS}=SS_T-SS_A-SS_S=32-24-4=4$$

$$SS_T=\sum\left(Y_{ij}-\bar{Y}_{..}\right)^2=(4-3.5)^2+(3-3.5)^2+\ldots+(2-3.5)^2+(1-3.5)^2=32$$

Source	SS	df	MS	F
A (Language Program)	4	$a-1=2-1=1$	$MS_A=\frac{SS_A}{df_A}=\frac{4}{1}=4$	$F_A=\frac{MS_A}{MS_{AxS}}=\frac{4}{.57}=7.02$
S	24	$n-1=8-1=7$	$MS_S=\frac{SS_S}{df_s}=\frac{24}{7}=3.4$	
AxS	4	$(a-1)(n-1)=(1)(7)=7$	$MS_{AxS}=\frac{SS_{AxS}}{df_{AxS}}=\frac{4}{7}=.57$	
T	32	$(a)(n)=(2)(8)=16$		

$F_{crit}\left(df_A=1, df_{AxS}=7\right)=5.59$

As in the between-subjects, case, the within-subjects ANOVA is equivalent to the paired (correlated scores) t-test, and $t^2=F$.

We could compute η^2 by simply dividing the sum of squares for the effect by the sum of squares total, $\eta^2 = SS_A/SS_T = 4/32 = .125$. Partial η^2 removes the subject effect from the denominator, $\eta^2 = SS_A/(SS_A + SS_{AxS}) = 4/(4 + 4) = .50$. And Cohen's f can be computed from eta-squared,

$$f=\sqrt{\frac{\eta^2}{1-\eta^2}}$$

[1] This kind of pretest-posttest design with only one group is subject to many potential threats to internal validity, so it is not a very good research design. I use it here just to illustrate a simple analysis. More on this and the issue of counterbalancing later.

SPSS
Syntax

```
glm biling immerse
  /wsfactor=edtype 2
  /wsdesign=edtype
  /print=parameter.
```

Menus
Analyze→ General Linear Model→ Repeated Measures
Name the factor, enter number of levels, click **Add**. Click **Define**. Drag over the variables for each level.

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
edtype	Sphericity Assumed	4.000	1	4.000	7.000	.033	.500
	Greenhouse-Geisser	4.000	1.000	4.000	7.000	.033	.500
	Huynh-Feldt	4.000	1.000	4.000	7.000	.033	.500
	Lower-bound	4.000	1.000	4.000	7.000	.033	.500
Error(edtype)	Sphericity Assumed	4.000	7	.571			
	Greenhouse-Geisser	4.000	7.000	.571			
	Huynh-Feldt	4.000	7.000	.571			
	Lower-bound	4.000	7.000	.571			

Here is the result from a paired t test that shows $t^2 = (3.646)^2 = F = 7.00$ and the same p-value

t	df	Sig. (2-tailed)
2.646	7	.033

There are several tables in the output we can ignore in this case and that I omitted: **Mauchly's Test of Sphericity**, **Tests of Within-Subjects Contrasts** (which contains partial eta-squared, here was equal to .50), **Tests of Between-Subjects Effects**, and **Parameter Estimates**. Some of these we will use later.

R
I have not found any great within-subjects functions in R. All of the ANOVA methods for repeated measures that I know of require a reconfiguration of the data set, which is inconvenient. The usual data set construction is to have multiple variables for each case representing each level of the within-subjects factor, sometimes called "wide" format. In the SPSS data set, the variables `biling` and `immerse` represent the two measurements of the same English proficiency (taken under different conditions). The data initially look like this:

```
> rm(d)
> rm(longdata)
> rm(mymodel)
>
> library(haven)
> d = read_sav("c:/jason/spsswin/uvclass/bilingual.sav")
> d

BILING IMMERSE
1       4       2
2       3       3
3       6       4
4       5       5
5       6       4
6       3       3
7       2       2
8       3       1
```

To obtain a within-subjects ANOVA, however, we need to reshape the data set so that `biling` and `immerse` are stacked vertically into one variable, resulting in 16 records (rows) instead of 8, sometimes called "long" format. In the long format, there will be two records for each case. The `reshape2` package makes this pretty easy to do. I will need ID numbers later, so I create them first.

```
> #create id numbers
> d$id <- 1:nrow(d)
>
> #reshape data going from wide to long format
> library(reshape2)
> longdata <- melt(d,
+                  measure.vars = c("BILING", "IMMERSE"),  #identify the two old variables
+                  variable.name = "level",        #name new variable for the value labels
+                  value.name = "edtype")          #name a new variable for the values
>
> longdata
```

The data now look like this.

```
   id   level edtype
1   1  BILING      4
2   2  BILING      3
3   3  BILING      6
4   4  BILING      5
5   5  BILING      6
6   6  BILING      3
7   7  BILING      2
8   8  BILING      3
9   1 IMMERSE      2
10  2 IMMERSE      3
11  3 IMMERSE      4
12  4 IMMERSE      5
13  5 IMMERSE      4
14  6 IMMERSE      3
15  7 IMMERSE      2
16  8 IMMERSE      1
```

The within-subjects ANOVA is then relatively easy to obtain with a `rstatix` R function. `ges` is eta-squared.

```
> #repeated measures anova
> library(rstatix)
> res.aov <- anova_test(data = longdata, dv = edtype, wid = id, within = level,detailed=TRUE)
> get_anova_table(res.aov)
ANOVA Table (type III tests)

       Effect DFn DFd SSn SSd      F       p p<.05   ges
1 (Intercept)   1   7 196  24 57.167 0.00013     * 0.875
2       level   1   7   4   4  7.000 0.03300     * 0.125
```

Example Write-Up

A repeated-measures ANOVA was used to compare bilingual and immersion conditions. Language usage scores were significantly higher when students received bilingual education than when they received immersion education, F(1,7)=7.01, p<.05, η^2 = .125. The average difference between language scores when students were in the bilingual versus the immersion program was 1.00.

It would also be fine to present the two means. And I report regular eta-squared (based on hand computation), but most people would report the partial eta-squared produced in the output in SPSS. That partial eta-squared was equal to .50, suggesting that 50% of the variances in accounted for by the independent variable, whereas the percent of total variation accounted for as indicated by the regular eta-squared was 12.5%.