

## 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 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	$SS_A = n \sum (\bar{Y}_j - \bar{Y}_{..})^2$	$a - 1$	$MS_A = \frac{SS_A}{df_A}$	$F_A = \frac{MS_A}{MS_{AxS}}$
S	$SS_S = a \sum (\bar{Y}_i - \bar{Y}_{..})^2$	$n - 1$	$MS_S = \frac{SS_S}{df_s}$	
AxS	$SS_{AxS} = \sum \sum (Y - \bar{Y}_i - \bar{Y}_j + \bar{Y}_{..})^2$ or $SS_{AxS} = SS_T - SS_A - SS_S$	$(a - 1)(n - 1)$	$MS_{AxS} = \frac{SS_{AxS}}{df_{AxS}}$	
T	$SS_T = \sum (Y_{ij} - \bar{Y}_{..})^2$	$(a)(n)$		

**Example**

Consider a hypothetical example in which non-native English speaking students are tested before and after the implementation of an “immersion” approach to teaching proficiency in English to eight students. 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.<sup>1</sup>

Student	Bilingual	$Y_{ij} - \bar{Y}_{..}$	$(Y_{ij} - \bar{Y}_{..})^2$	Immersion	$Y_{ij} - \bar{Y}_{..}$	$(Y_{ij} - \bar{Y}_{..})^2$	$\bar{Y}_{i.}$	$\bar{Y}_{i.} - \bar{Y}_{..}$	$(\bar{Y}_{i.} - \bar{Y}_{..})^2$
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
	$\bar{Y}_{.1} = 4$			$\bar{Y}_{.2} = 3$					$\sum (\bar{Y}_{i.} - \bar{Y}_{..})^2 = 12$

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

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

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

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

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

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	$(a)(n) = (2)(8) = 16$		

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

Language usage scores were significantly higher when students received bilingual education,  $F(1,7)=7.01, p<.05$ . The average difference between language scores when students were in the bilingual versus the immersion program was 1.00.

<sup>1</sup> As discussed earlier in the course, 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.

**SPSS Output**

**General Linear Model**

**Within-Subjects Factors**

Measure: MEASURE\_1

langprog	Dependent Variable
1	biling
2	immerse

**Tests of Within-Subjects Effects**

Measure: MEASURE\_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
langprog	Sphericity Assumed	4.000	1	4.000	7.000	.033
	Greenhouse-Geisser	4.000	1.000	4.000	7.000	.033
	Huynh-Feldt	4.000	1.000	4.000	7.000	.033
	Lower-bound	4.000	1.000	4.000	7.000	.033
Error(langprog)	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		

**Tests of Within-Subjects Contrasts**

Measure: MEASURE\_1

Source	langprog	Type III Sum of Squares	df	Mean Square	F	Sig.
langprog	Linear	4.000	1	4.000	7.000	.033
Error(langprog)	Linear	4.000	7	.571		

**Tests of Between-Subjects Effects**

Measure: MEASURE\_1

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	196.000	1	196.000	57.167	.000
Error	24.000	7	3.429		