Repeated Measures vs Randomized Blocks ANOVA

George H Olson, PhD
Appalachian State University
Original Version: Fall 2011
(Revised: Spring 2013, Spring 2014)

In the VassarStats textbook, Chapter 15 (One-way Analysis of Variance for Correlated Samples), Part 1, you were introduced to two versions of ANOVA for correlated samples: Repeated Measures ANOVA (RMAnova) and Randomized Blocks ANOVA (RBAnova).

One-way Designs

In a one-way RMAnova each subject is observed and measured under two or more conditions. A general research design for a RMAnova can be depicted as shown below. In the schematic, X_{ik} represents the measurement for individual *i* under treatment *k*. In the design, each of N subjects are exposed to all K treatments.

Table 1: Research Design for an N × K Repeated Measures ANOVA

	Treatment 1	Treatment 2	Treatment 3	· · · Treatment *k* · · ·	Treatment K
Subject 1	X_{11}	X_{12}	X_{13}	· · · X_{1k} · · ·	X_{1K}
Subject 2	X_{21}	X_{22}	X_{23}	· · · X_{2k} · · ·	X_{2K}
.	.	.	.	. . . .	.
.	.	.	.	. . . .	.
.	.	.	.	. . . .	.
Subject *i*	X_{i1}	X_{i2}	X_{i3}	· · · X_{ik} · · ·	X_{iK}
.	.	.	.	. . . .	.
.	.	.	.	. . . .	.
.	.	.	.	. . . .	.
Subject N	X_{N1}	X_{N2}	X_{N3}	· · · X_{Nk} · · ·	X_{NK}

The design for a one-way RBAnova is a little different, as shown in Table 2. In a true randomized blocks design, the number of Blocks is equal to the number of measurements, or times, that measurements are taken. Each Block contains K different subjects who *are matched* on some characteristic. Hence, all the observations within a Block are assumed correlated. Also, there is one subject in each Block × Treatment combination. Hence subjects are *confounded* with Block × Time of Measurement combinations (the notation, X_{iik}, indicates a measurement for **individual** *i* in **Block** *k* who is measured at **Time** *k*.) In a one-way RBAnova the notation could be simplified by eliminating the first subscript for each observation. In other words, the notation, X_{iik}, could have just as easily been written as X_{ik}, providing it is understood that the second subscript represents *both* the **individual**, *i*, and the Block, *i*. In this design, there are K×K subjects[1].

[1] When considering more complex research designs, like those shown here, it is important to diagram the design, including subscripts. Doing so helps the analyst keep track of what is being analyzed.

The advantage of the randomize blocks design is the same as that for a repeated measures design and is adequately explained in Part 1 of VassarStats Chapter 15.

Table 2: Research Design for an K × K Randomized Blocks ANOVA

	Measurement at Time k						
	1	2	3	· · ·	k	· · ·	K
Block 1	X_{111}	X_{212}	X_{313}	· · ·	X_{k1k}	· · ·	X_{K1K}
Block 2	X_{121}	X_{222}	X_{223}	· · ·	X_{22k}	· · ·	X_{22K}
Block 3	X_{331}	X_{332}	X_{333}	· · ·	X_{33k}	· · ·	X_{33K}
.	.	.	.	· · ·	.	· · ·	· · ·
.	.	.	.	· · ·	.	· · ·	· · ·
.	.	.	.	· · ·	.	· · ·	· · ·
Block j	X_{jj1}	X_{jj2}	X_{jj3}	· · ·	X_{jjk}	· · ·	X_{jjK}
.	.	.	.	· · ·	.	· · ·	· · ·
.	.	.	.	· · ·	.	· · ·	· · ·
.	.	.	.	· · ·	.	· · ·	· · ·
							· · ·
Block K	X_{KK1}	X_{KK2}	X_{KK3}	· · ·	X_{KKk}	· · ·	X_{KKK}

J-Between, K-Within Designs

Often, in practice, in an RMAnova design, we might have two or more (or J) groups of individuals –representing two or more levels of an independent variable—on whom a common set of K *repeated* measurements are taken. Similarly, in an RBAnova design, we might have two or J blocks if individuals–representing two or K levels of an independent variable—on whom a common set of measurements are taken. The difference is that in the RMAnova the *same* individuals are measured repeatedly, whereas in the RBAnova design *different* individuals matched on some characteristic are measured on the K different occasions.

A research design for a J-Between, K-Within RMAnova is depicted in Table 3, where it is assumed that Factor A has K levels. Since different individuals are measured in each level of Factor A, Factor A is referred to as a *between* groups factor. The repeated measures, on the other hand, are measured within individuals; hence the repeated measures factor is referred to as a *within* groups factor.

In Table 3 the notation, X_{ijk}, denotes k^{th} measure taken on the i^{th} individual in group at level Aj. We often refer to this design as a one-between, one-within design. It should be obvious that more complex designs are possible. For instance, a two-between, one-within design would have two between-groups factors, A and B, for instance, and one within-groups factor (measures). Similarly, a two-between, two-within design would have two between-groups factors and two sets of repeated measures, taken under two conditions (before lunch and after lunch, for instance.)

Table 3: Research Design for an J × K Repeated Measures ANOVA (One Between, One Within)

	Measurement at Time k						
Factor A	1	2	3	· · ·	k	· · ·	K
Level A_1	X_{111}	X_{112}	X_{113}	· · ·	X_{11k}	· · ·	X_{11K}
	X_{211}	X_{212}	X_{213}	· · ·	X_{21k}	· · ·	X_{21K}
	X_{311}	X_{312}	X_{313}	· · ·	X_{31k}	· · ·	X_{31K}
	.	.	.	· · ·	.	· · ·	.
	.	.	.	· · ·	.	· · ·	.
	.	.	.	· · ·	.	· · ·	.
	$X_{n_1 11}$	$X_{n_1 12}$	$X_{n_1 13}$	· · ·	$X_{n_1 1k}$	· · ·	$X_{n_1 1K}$
Level A_2	X_{121}	X_{122}	X_{123}	· · ·	X_{12k}	· · ·	X_{12K}
	X_{221}	X_{222}	X_{223}	· · ·	X_{22k}	· · ·	X_{22K}
	X_{321}	X_{322}	X_{323}	· · ·	X_{32k}	· · ·	X_{32K}
	.	.	.	· · ·	.	· · ·	.
	.	.	.	· · ·	.	· · ·	.
	.	.	.	· · ·	.	· · ·	.
	$X_{n_2 21}$	$X_{n_2 22}$	$X_{n_2 23}$	· · ·	$X_{n_2 2k}$	· · ·	$X_{n_2 2K}$
.	.	.	.	· · ·	.	· · ·	· · ·
.	.	.	.	· · ·	.	· · ·	· · ·
.	.	.	.	· · ·	.	· · ·	· · ·
Level A_j	X_{ij1}	X_{ij2}	X_{ij3}	· · ·	X_{ijk}	· · ·	X_{ijK}
.	.	.	.	· · ·	.	· · ·	· · ·
.	.	.	.	· · ·	.	· · ·	· · ·
.	.	.	.	· · ·	.	· · ·	· · ·
Level A_J	X_{1J1}	X_{1J2}	X_{1J3}	· · ·	X_{1Jk}	· · ·	X_{1JK}
	X_{2J1}	X_{2J2}	X_{2J3}	· · ·	X_{2Jk}	· · ·	X_{2JK}
	X_{3J1}	X_{3J2}	X_{3J3}	· · ·	X_{3Jk}	· · ·	X_{32K}
	.	.	.	· · ·	.	· · ·	.
	.	.	.	· · ·	.	· · ·	.
	.	.	.	· · ·	.	· · ·	.
	$X_{n_J J1}$	$X_{n_J J2}$	$X_{n_J J3}$	· · ·	$X_{n_J Jk}$	· · ·	$X_{n_J JK}$

In the design, there are n_j individuals at each level, j, of Factor A. Each individual has a measurement on all K repeated measures. The notation, X_{324}, is the fourth measurement taken on the third individual at Level 2. Note that there can be a different number of individuals (n_j) at each level of A.

A research design of an J x K RBAnova design is similar to the RMAnova design except that instead of levels of a between-group factor we have different blocks of similar (presumably matched) individuals. The blocks could represent different levels of some independent variable as in the RMAnova design. The main difference between the two designs is that in the RBAnova design all the X_{ijk} represent different measures on *different* individuals. Hence, X_{ijk}, represents the k^{th} measurement on the i'th subject in Block j.

Table 4: Research Design for an J × K Randomized Blocks ANOVA

	Measurement at Time k						
Factor A	1	2	3	· · ·	k	· · ·	K
Block 1	X_{111}	X_{112}	X_{113}	· · ·	X_{11k}	· · ·	X_{11K}
	X_{211}	X_{212}	X_{213}	· · ·	X_{21k}	· · ·	X_{21K}
	X_{311}	X_{312}	X_{313}	· · ·	X_{31k}	· · ·	X_{31K}
	.	.	.	· · ·	.	· · ·	.
	.	.	.	· · ·	.	· · ·	.
	.	.	.	· · ·	.	· · ·	.
	$X_{n_1 11}$	$X_{n_1 12}$	$X_{n_1 13}$	· · ·	$X_{n_1 1k}$	· · ·	$X_{n_1 1K}$
Block 2	X_{121}	X_{122}	X_{123}	· · ·	X_{12k}	· · ·	X_{12K}
	X_{221}	X_{222}	X_{223}	· · ·	X_{22k}	· · ·	X_{22K}
	X_{321}	X_{322}	X_{323}	· · ·	X_{32k}	· · ·	X_{32K}
	.	.	.	· · ·	.	· · ·	.
	.	.	.	· · ·	.	· · ·	.
	.	.	.	· · ·	.	· · ·	.
	$X_{n_2 21}$	$X_{n_2 22}$	$X_{n_2 23}$	· · ·	$X_{n_2 2k}$	· · ·	$X_{n_2 2K}$
.	.	.	.	· · ·	.	· · ·	· · ·
.	.	.	.	· · ·	.	· · ·	· · ·
.	.	.	.	· · ·	.	· · ·	· · ·
Block j	X_{ij1}	X_{ij2}	X_{ij3}	· · ·	X_{ijk}	· · ·	X_{ijK}
.	.	.	.	· · ·	.	· · ·	· · ·
.	.	.	.	· · ·	.	· · ·	· · ·
.	.	.	.	· · ·	.	· · ·	· · ·
Block J	X_{1J1}	X_{1J2}	X_{1J3}	· · ·	X_{1Jk}	· · ·	X_{1JK}
	X_{2J1}	X_{2J2}	X_{2J3}	· · ·	X_{2Jk}	· · ·	X_{2JK}
	X_{3J1}	X_{3J2}	X_{3J3}	· · ·	X_{3Jk}	· · ·	X_{32K}
	.	.	.	· · ·	.	· · ·	.
	.	.	.	· · ·	.	· · ·	.
	.	.	.	· · ·	.	· · ·	.
	$X_{n_J J1}$	$X_{n_J J2}$	$X_{n_J J3}$	· · ·	$X_{n_J Jk}$	· · ·	$X_{n_J JK}$

As in the case of RMAnova the number of subjects within a block can vary across blocks. Hence, while block 2 may have n_2 subjects, block K might have n_K subjects.

An Example: Randomized Blocks Repeated Measures Design

In this factitious example we have a researcher who wants to investigate the effects of three types of instruction: Face to Face (F2F), Virtual Face to Face (V-F2F) via an immersive virtual environment (e.g., Appstate's Open Qwaq), and asynchronous online (AsyOL). She has 23 students in the particular class in which she wants to conduct her study. One approach she could use is to randomly assign the 23 students to the three conditions as follows:

Table 5: One Possible Scenario for Using 23 Subjects

Instructional Condition	F2F	V-F2F	AsyOL
Number of students	8	8	7

However, she realizes that there could be considerable variance due to individual differences among students assigned to each condition (random assignment does not mitigate this potential problem). This could result in inflated within-group variance, leading to weakened power. Instead, she opts to employ a repeated-measures design where all 23 students are exposed to all three instructional conditions.

She realizes, also, that order of exposure to the three conditions may have a systematic effect on the students' achievement outcomes. For instance, exposure to F3F first might influence how students later react to V-F2F. Additionally, it is not unreasonable to assume that there may be a sequential, cumulative effect to the instructional conditions. For instance, regardless of which condition students are exposed to first, that exposure might affect their reaction to the second condition exposure, and so on. With this realization, she decides to employ a variation of a randomized blocks with repeated measures design. The design looks like that depicted in the table on the next page.

There are three blocks, each having a different sequence of instructional conditions:

BLOCK 1: F2F, first, followed by V-F2F, followed by AsyOL,
BLOCK 2: V-F2F, first, followed by AsyOL followed by F2F,
BLOCK 3: AsyOL first, followed by F2F, followed by V-F2F.

Other sequences are, of course, possible. However, one of the things she is particularly interested in is the carry-over effect of F2F on V-F2F and AsyOL. Based on her previous experience with online instruction, she hypothesizes that F2F instruction has a positive influence on students' reaction to exposure to the two types of online instruction.

She randomly assigns students to the three blocks (eight students to each of the first two blocks, and seven to the BLOCK 3. Within each block, students are exposed to each of the three instructional conditions, in order, for five weeks. At the end of each five weeks, she administers a 25-item achievement test over the content covered during the previous five weeks. Hence, she collects achievement data three times over the 15-week semester. It should be noted that as the semester progresses, all students cover the same content regardless of their sequence of instructional exposure.

Table 7: Randomized Blocks Repeated Measures Design

Block (Instructional Sequence)	Achievement Assessment 1	2	3
Block 1 (F2F, V-F2F, AsyOL)	X_{111}	X_{112}	X_{113}
	X_{211}	X_{212}	X_{213}
	X_{311}	X_{312}	X_{313}
	.	.	.
	.	.	.
	.	.	.
	X_{811}	X_{812}	X_{813}
Block 2 (V-F2F, AsyOL, F2F)	X_{121}	X_{122}	X_{123}
	X_{221}	X_{222}	X_{223}
	X_{321}	X_{322}	X_{323}
	.	.	.
	.	.	.
	.	.	.
	X_{821}	X_{822}	X_{823}
Block 3 (AsyOL, F2F, V-F2F)	X_{131}	X_{132}	X_{133}
	X_{231}	X_{232}	X_{233}
	X_{331}	X_{332}	X_{333}
	.	.	.
	.	.	.
	.	.	.
	X_{731}	X_{732}	X_{733}

In the table above, an entry, X_{ijk}, denotes an achievement score on the *k*th achievement assessment for the *i*th student in block *j*. For example, the entry, X_{322}, denotes the third achievement assessment score for the second student in block 2.

The fictitious data used in this example are given in Table 8. To analyze those data, I used the SPSS General Linear Model, Repeated Measures procedure.[2]

From the SPSS output, I was able to construct Table 9, the Table of descriptive statistics. It appears, from an inspection of the table, that any sequence of instruction beginning with F2F resulted in higher test scores on all three tests.

[2] A copy of the Syntax commands and SPSS Output are given in the APPENDIX.

Table 8: Data for The Example

Randomized Blocks Repeated Measures Design			
	Achievement Assessment		
	1	2	3
Block 1 (F2F, V-F2F, AsyOL)	20	24	22
	19	23	24
	20	21	23
	17	24	21
	21	25	21
	18	24	22
	22	22	20
	17	23	23
Block 2 (V-F2F, AsyOL, F2F)	17	15	18
	18	18	19
	18	17	22
	16	15	20
	15	15	17
	18	18	20
	15	17	21
	20	14	18
Block 3 (AsyOL, F2F, V-F2F)	17	23	22
	15	22	21
	16	20	20
	13	21	22
	15	19	18
	12	20	18
	17	24	21

Table 9: Descriptive Statistics

	Assessment					
Instructional Sequence	Measurement 1		Measurement 2		Measurement 3	
	Mean	SD	Mean	SD	Mean	SD
F2F,V-F2F,AsyOL	19.25	1.832	23.25	1.282	22.00	1.309
V-F2F,AsyOL ,F2F	17.13	1.727	16.13	1.553	19.38	1.685
AsyOL ,F2F,V-F2F	15.00	1.915	21.29	1.799	20.29	1.704
TOTAL	17.22	2.467	20.17	3.460	20.57	1.879

This observation was confirmed by the analysis (Table 10). Statistically significant differences found among blocks. $F(2.20)=26.091$ and among measures, $F(2,40)=36.823$. Both F tests were significant ($p<.0005$). There was, however, a significant Block by Measure interaction, $F(4,40)=13.145$; $p<.0005$), which complicated the interpretation of the results. A plot of the means for each instructional sequence block (Figure 1) shows the interaction clearly. While the sequences beginning with F2F yielded higher levels of performance across the board, the effects of the other two instructional sequences were mixed.

Table 10: Analysis of Variance Summary Table

Source	SS	df	MS	F	Sig
Between Subjects					
Blocks (B)	194.456	2	97.228	26.091	<.0005
Error	74.530	20	3.726		
Within Subjects					
Measures. (M)	163.775	2	81.887	36.823	<.0005
B x M	116.932	4	29.233	13.145	<.0005
Error (w/ groups)	188.952	40	.4082.224		

On the first test, the group receiving virtual F2F instruction first outperformed the group receiving asynchronous online instruction. By the second assessment, the effects of V-F2F and AsyOL were reversed, with the group receiving asynchronous instruction out performing those receiving virtual F2F instruction. While this difference persisted at the third assessment, the difference between the two groups was not as great.

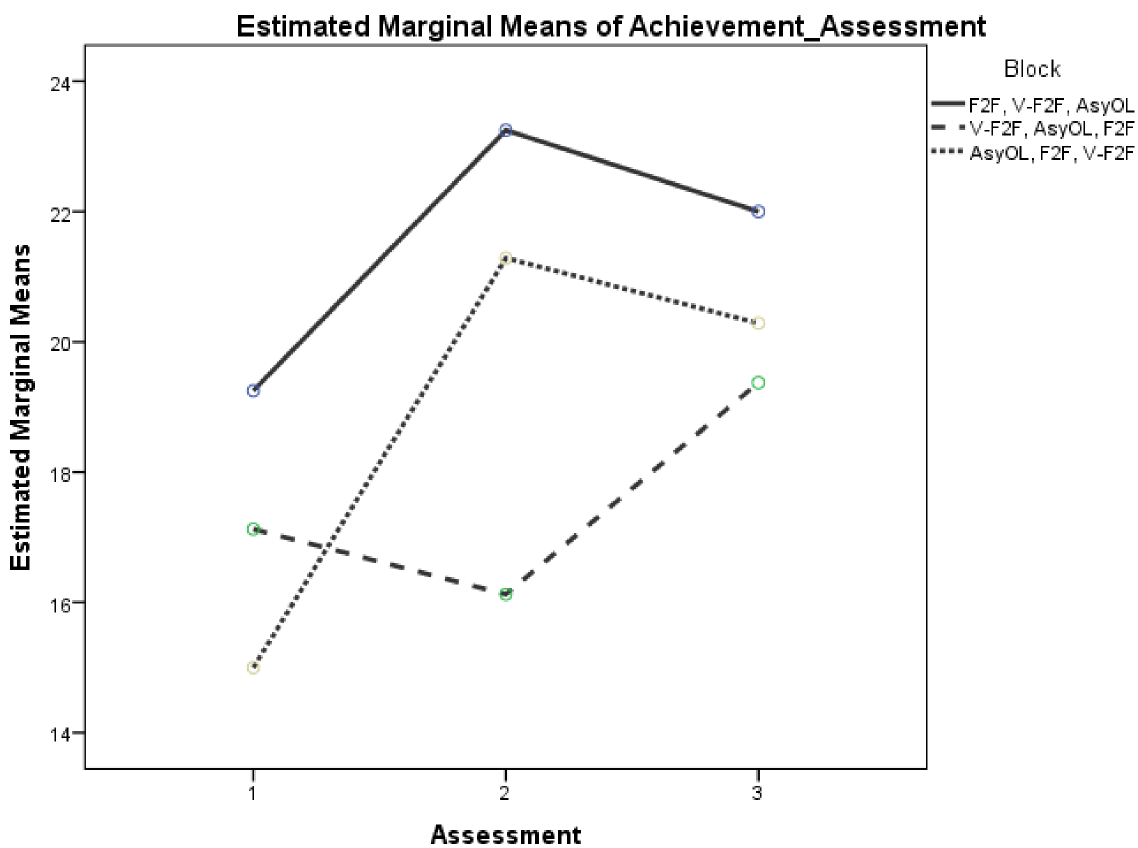

Figure 1: Assessment scores by Instructional Sequence Block

Appendix

SPSS syntax for analyzing the data for the example.

```
GLM Test1 Test2 Test3 BY Block
  /WSFACTOR=Assessment 3 Polynomial
  /MEASURE=Achievement_Assessment
  /METHOD=SSTYPE(3)
  /PLOT=PROFILE(Assessment*Block)
  /EMMEANS=TABLES(OVERALL)
  /EMMEANS=TABLES(Block)
  /EMMEANS=TABLES(Assessment)
  /EMMEANS=TABLES(Block*Assessment)
  /PRINT=DESCRIPTIVE
  /CRITERIA=ALPHA(.05)
  /WSDESIGN=Assessment
  /DESIGN=Block.
```

SPSS Output, using the syntax given above:

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
factor1	Sphericity Assumed	.032	2	.016	.040	.961
	Greenhouse-Geisser	.032	1.589	.020	.040	.932
	Huynh-Feldt	.032	1.877	.017	.040	.954
	Lower-bound	.032	1.000	.032	.040	.844
factor1 * Block	Sphericity Assumed	10.502	4	2.626	6.428	.000
	Greenhouse-Geisser	10.502	3.179	3.304	6.428	.001
	Huynh-Feldt	10.502	3.754	2.798	6.428	.001
	Lower-bound	10.502	2.000	5.251	6.428	.007
Error(factor1)	Sphericity Assumed	16.339	40	.408		
	Greenhouse-Geisser	16.339	31.789	.514		
	Huynh-Feldt	16.339	37.541	.435		
	Lower-bound	16.339	20.000	.817		

Within-Subjects Factors

Measure: Achievement_Assessment

Assessment	Dependent Variable
1	Test1
2	Test2
3	Test3

Between-Subjects Factors

		Value Label	N
Block	1	F2F, V-F2F, AsyOL	8
	2	V-F2F, AsyOL, F2F	8
	3	AsyOL, F2F, V-F2F	7

Descriptive Statistics

	Block	Mean	Std. Deviation	N
Test 1	F2F, V-F2F, AsyOL	19.25	1.832	8
	V-F2F, AsyOL, F2F	17.13	1.727	8
	AsyOL, F2F, V-F2F	15.00	1.915	7
	Total	17.22	2.467	23
Test 2	F2F, V-F2F, AsyOL	23.25	1.282	8
	V-F2F, AsyOL, F2F	16.13	1.553	8
	AsyOL, F2F, V-F2F	21.29	1.799	7
	Total	20.17	3.460	23
Test 3	F2F, V-F2F, AsyOL	22.00	1.309	8
	V-F2F, AsyOL, F2F	19.38	1.685	8
	AsyOL, F2F, V-F2F	20.29	1.704	7
	Total	20.57	1.879	23

Multivariate Tests[a]

Effect		Value	F	Hypothesis df	Error df	Sig.
Assessment	Pillai's Trace	.726	25.126[b]	2.000	19.000	.000
	Wilks' Lambda	.274	25.126[b]	2.000	19.000	.000
	Hotelling's Trace	2.645	25.126[b]	2.000	19.000	.000
	Roy's Largest Root	2.645	25.126[b]	2.000	19.000	.000
Assessment * Block	Pillai's Trace	.899	8.166	4.000	40.000	.000
	Wilks' Lambda	.210	11.241[b]	4.000	38.000	.000
	Hotelling's Trace	3.248	14.616	4.000	36.000	.000
	Roy's Largest Root	3.080	30.796[c]	2.000	20.000	.000

a. Design: Intercept + Block
Within Subjects Design: Assessment

b. Exact statistic

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

Mauchly's Test of Sphericity[a]

Measure: Achievement_Assessment

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[b] Greenhouse-Geisser	Huynh-Feldt	Lower-bound
Assessment	.842	3.264	2	.195	.864	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 + Block
Within Subjects Design: Assessment

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.

Tests of Within-Subjects Contrasts

Measure: Achievement_Assessment

Source	Assessment	Type III Sum of Squares	df	Mean Square	F	Sig.
Assessment	Linear	134.649	1	134.649	47.068	.000
	Quadratic	29.126	1	29.126	18.354	.000
Assessment * Block	Linear	19.394	2	9.697	3.390	.054
	Quadratic	97.537	2	48.769	30.732	.000
Error(Assessment)	Linear	57.214	20	2.861		
	Quadratic	31.738	20	1.587		

Tests of Between-Subjects Effects

Measure: Achievement_Assessment
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	25599.169	1	25599.169	6869.516	.000
Block	194.456	2	97.228	26.091	.000
Error	74.530	20	3.726