Lecture 12: Mixed-Design ANOVA

Combining between-group & within-subject

In the simplest 2 x 2 design you would have subjects randomly assigned to one of two groups (two study systems, Kaplan vs Princeton), but each group would experience 2 conditions (measurements).

	GRE - before	GRE - after
Kaplan	S_1 S_2 S_3 S_4 S_5	S_1 S_2 S_3 S_4 S_5
Princeton	S_6 S_7 S_8 S_9 S_{10}	S_6 S_7 S_8 S_9 S_{10}

Analysis

Repeated measures

		b_1	b_2	b_3
Randomized Groups	a_1	S_1	S_1	S_1
		S_2	S_2	S_2
		S_3	S_3	S_3
		S_4	S_4	S_4
		S_5	S_5	S_5
	a_2	S_6	S_6	S_6
		S_7	S_7	S_7
		S_8	S_8	S_8
		S_9	S_9	S_9
		S_{10}	S_{10}	S_{10}
	a_3	S_{11}	S_{11}	S_{11}
		S_{12}	S_{12}	S_{12}
		S_{13}	S_{13}	S_{13}
		S_{14}	S_{14}	S_{14}
		S_{15}	S_{15}	S_{15}

Sources of Variance

- $SS_T = SS_{BG} + SS_{WS}$
- What are the sources of variance?
  - A (between-group factor)
  - S/A (deviation of each subject from its A group mean)
  - B (repeated-measure factor)
  - AB (interaction of two factors)
  - BxS/A (deviation of each subject from its cell mean, minus S/A)
  - T (total variance).
- Degrees of freedom?

Degree of freedom

- A = a - 1
- S/A = a(s – 1) = as – a

  (s is the number of subjects on each level of A)
- B = b - 1
- AB = (a – 1)(b – 1)
- BxS/A = a(b – 1)(s – 1)
- T = abs – 1 or N - 1

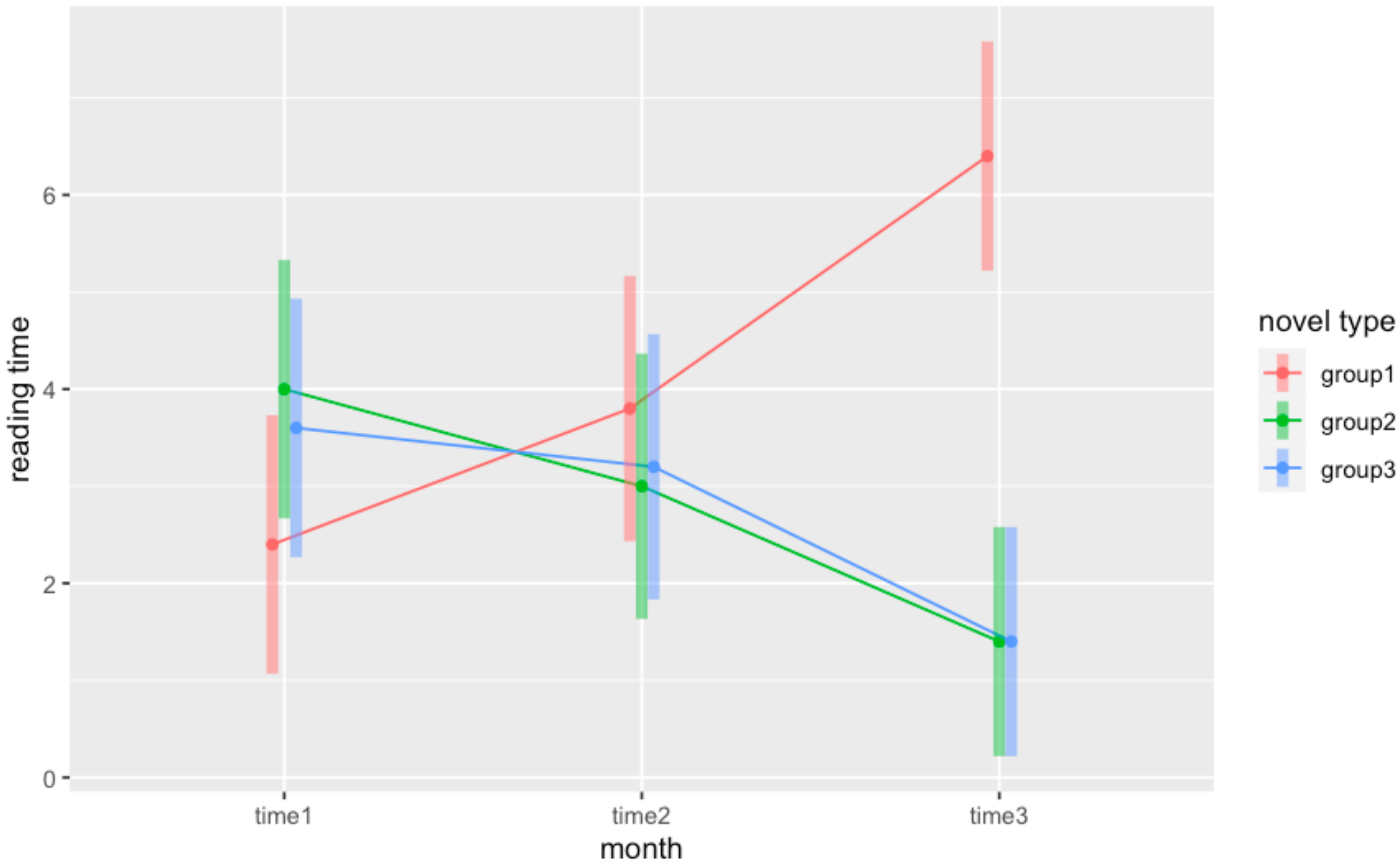
Example – Books by Month

Example:

- Imagine if we designed a study concerning reading different novels over time
- we randomly assign subjects to three types of books and have them read for three months
- Book type is the group factor, month is the repeated-measure factor.
- How many they read are measured.

			B: *Month*			Case Means
			b_1 : *Month 1*	b_2 : *Month 2*	b_3 : *Month 3*	
A: *Type of Novel*	: *Science Fiction*	S_1	1	3	6	**S_1 = 3.333**
		S_2	1	4	8	**S_2 = 4.333**
		S_3	3	3	6	**S_3 = 4**
		S_4	5	5	7	**S_4 = 5.667**
		S_5	2	4	5	**S_5 = 3.667**
			$a_1 b_1$ = 2.4	**$a_1 b_2$ = 3.8**	**$a_1 b_3$ = 6.4**	**a_1 = 4.2**
	a_2 : *Mystery*	S_6	3	1	0	**S_6 = 1.333**
		S_7	4	4	2	**S_7 = 3.333**
		S_8	5	3	2	**S_8 = 3.333**
		S_9	4	2	0	**S_9 = 2**
		S_{10}	4	5	3	**S_{10} = 4**
		a2	**$a_2 b_1$ = 4**	**$a_2 b_2$ = 3**	**$a_2 b_3$ = 1.4**	**a_2 = 2.8**
	a_3 : *Romance*	S_{11}	4	2	0	**S_{11} = 2**
		S_{12}	2	6	1	**S_{12} = 3**
		S_{13}	3	3	3	**S_{13} = 3**
		S_{14}	6	2	1	**S_{14} = 3**
		S_{15}	3	3	2	**S_{15} = 2.667**
			$a_3 b_1$ = 3.6	**$a_3 b_2$ = 3.2**	**$a_3 b_3$ = 1.4**	**a_3 = 2.733**
			b_1 = 3.333	**b_2 = 3.333**	**b_3 = 3.067**	**GM = 3.244**

Eyeball it first



Sums of Squares

i: subject, 1~5
j: group factor A (novel), 1~3
k: within-subject factor B (month), 1~3

The total variability can be partitioned into A, B, AB, S/A, and B*S/A

$$SS_{Total} = SS_A + SS_B + SS_{AB} + SS_{S/A} + SS_{B*S/A}$$

$$\sum\left(Y_{ijk} - \bar{Y}_{...}\right)^2 = \sum n_j\left(\bar{Y}_{.j.} - \bar{Y}_{...}\right)^2 + \sum n_k\left(\bar{Y}_{..k} - \bar{Y}_{...}\right)^2 +$$

$$+\left[\sum n_{jk}\left(\bar{Y}_{.jk} - \bar{Y}_{...}\right)^2 - \sum n_j\left(\bar{Y}_{.j.} - \bar{Y}_{...}\right)^2 - \sum n_k\left(\bar{Y}_{..k} - \bar{Y}_{...}\right)^2\right]$$

$$+k\sum(\bar{Y}_{i..} - \bar{Y}_{.j.})^2 + \left[\sum(\mathrm{Y}_{ijk} - \bar{Y}_{.jk})^2 - k\sum(\bar{Y}_{i..} - \bar{Y}_{.j.})^2\right]$$

$$SS_A = \sum n_j \left(\bar{Y}_{.j.} - \bar{Y}_{...}\right)^2 = 15 * [\left(4.2 - 3.244\right)^2 + \left(2.8 - 3.244\right)^2 + \left(2.733 - 3.244\right)^2] = 20.583$$

$$SS_B = \sum n_k \left(\bar{Y}_{..k} - \bar{Y}_{...}\right)^2 = 15 * [\left(3.333 - 3.244\right)^2 + \left(3.333 - 3.244\right)^2 + \left(3.067 - 3.244\right)^2] = .708$$

$$SS_{AB} = \left[\sum n_{jk} \left(\bar{Y}_{.jk} - \bar{Y}_{...}\right)^2 - \sum n_j \left(\bar{Y}_{.j.} - \bar{Y}_{...}\right)^2 - \sum n_k \left(\bar{Y}_{..k} - \bar{Y}_{...}\right)^2\right] =$$

$$\sum n_{jk} \left(\bar{Y}_{.jk} - \bar{Y}_{...}\right)^2 = 5 * [\left(2.4 - 3.244\right)^2 + \left(3.8 - 3.244\right)^2 + \left(6.4 - 3.244\right)^2 +$$

$$+ \left(4 - 3.244\right)^2 + \left(3 - 3.244\right)^2 + \left(1.4 - 3.244\right)^2 +$$

$$+ \left(3.6 - 3.244\right)^2 + \left(3.2 - 3.244\right)^2 + \left(1.4 - 3.244\right)^2] = 92.711$$

$$SS_{AB} = 92.711 - 20.583 - .708 = 71.420$$

This k is necessary; 3 months.
The deviation of each subject from his group mean

$$SS_{S/A} = \boxed{k}\sum\left(\bar{Y}_{i..} - \bar{Y}_{.j.}\right)^2 = 3*[\left(3.333-4.2\right)^2 + \left(4.333-4.2\right)^2 + \left(4-4.2\right)^2 +$$

$$+\left(5.667-4.2\right)^2 + \left(3.667-4.2\right)^2 + \left(3.333-2.8\right)^2 + \left(1.333-2.8\right)^2 + \left(3.333-2.8\right)^2 +$$

$$+\left(3.333-2.8\right)^2 + \left(2-2.8\right)^2 + \left(2-2.733\right)^2 + \left(3-2.733\right)^2 + \left(3-2.733\right)^2 +$$

$$+\left(3-2.733\right)^2 + \left(2.667-2.733\right)^2] = 26.400$$

The deviation of each subject (within a cell) from its cell mean

$$SS_{B*S/A} = \left[\sum \left(Y_{ijk} - \bar{Y}_{.jk} \right)^2 - k \sum \left(\bar{Y}_{i..} - \bar{Y}_{.j.} \right)^2 \right] =$$

$$\sum \left(Y_{ijk} - \bar{Y}_{.jk} \right)^2 = (1-2.4)^2 + (1-2.4)^2 + (3-2.4)^2 + (5-2.4)^2 + (2-2.4)^2 +$$

$$+(3-3.8)^2 + (4-3.8)^2 + (3-3.8)^2 + (5-3.8)^2 + (4-3.8)^2 +$$

$$+(6-6.4)^2 + (8-6.4)^2 + (6-6.4)^2 + (7-6.4)^2 + (5-6.4)^2 +$$

$$+(3-4)^2 + (4-4)^2 + (5-4)^2 + (4-4)^2 + (4-4)^2 +$$

$$+(1-3)^2 + (4-3)^2 + (3-3)^2 + (2-3)^2 + (5-3)^2 +$$

$$+(0-1.4)^2 + (2-1.4)^2 + (2-1.4)^2 + (0-1.4)^2 + (3-1.4)^2 +$$

$$+(4-3.6)^2 + (2-3.6)^2 + (3-3.6)^2 + (6-3.6)^2 + (3-3.6)^2 +$$

$$+(2-3.2)^2 + (6-3.2)^2 + (3-3.2)^2 + (2-3.2)^2 + (3-3.2)^2 +$$

$$+(0-1.4)^2 + (1-1.4)^2 + (3-1.4)^2 + (1-1.4)^2 + (2-1.4)^2 = 63.6$$

$$SS_{B*S/A} = 63.6 - 26.4 = 37.2$$

The deviation of each data point (within a cell) from the global mean

$$SS_{Total} = \sum \left(Y_{ijk} - \bar{Y}_{...}\right)^2 =$$

$$SS_{Total} = (1-3.244)^2 + (1-3.244)^2 + (3-3.244)^2 + (5-3.244)^2 + (2-3.244)^2 +$$
$$+(3-3.244)^2 + (4-3.244)^2 + (3-3.244)^2 + (5-3.244)^2 + (4-3.244)^2 +$$
$$+(6-3.244)^2 + (8-3.244)^2 + (6-3.244)^2 + (7-3.244)^2 + (5-3.244)^2 +$$
$$+(3-3.244)^2 + (4-3.244)^2 + (5-3.244)^2 + (4-3.244)^2 + (4-3.244)^2 +$$
$$+(1-3.244)^2 + (4-3.244)^2 + (3-3.244)^2 + (2-3.244)^2 + (5-3.244)^2 +$$
$$+(0-3.244)^2 + (2-3.244)^2 + (2-3.244)^2 + (0-3.244)^2 + (3-3.244)^2 +$$
$$+(4-3.244)^2 + (2-3.244)^2 + (3-3.244)^2 + (6-3.244)^2 + (3-3.244)^2 +$$
$$+(2-3.244)^2 + (6-3.244)^2 + (3-3.244)^2 + (2-3.244)^2 + (3-3.244)^2 +$$
$$+(0-3.244)^2 + (1-3.244)^2 + (3-3.244)^2 + (1-3.244)^2 + (2-3.244)^2 = 156.311$$

			B: *Month*			
			b_1: *Month 1*	b_2: *Month 2*	b_3: *Month 3*	Case Total
A: *Type of Novel*	a_1: *Science Fiction*	S_1	1	3	6	**S_1 = 10**
		S_2	1	4	8	**S_2 = 13**
		S_3	3	3	6	**S_3 = 12**
		S_4	5	5	7	**S_4 = 17**
		S_5	2	4	5	**S_5 = 11**
			a_1b_1 = 12	**a_1b_2 = 19**	**a_1b_3 = 32**	**a_1 = 63**
	a_2: *Mystery*	S_6	3	1	0	**S_6 = 4**
		S_7	4	4	2	**S_7 = 10**
		S_8	5	3	2	**S_8 = 10**
		S_9	4	2	0	**S_9 = 6**
		S_{10}	4	5	3	**S_{10} = 12**
			a_2b_1 = 20	**a_2b_2 = 15**	**a_2b_3 = 7**	**a_2 = 42**
	a_3: *Romance*	S_{11}	4	2	0	**S_{11} = 6**
		S_{12}	2	6	1	**S_{12} = 9**
		S_{13}	3	3	3	**S_{13} = 9**
		S_{14}	6	2	1	**S_{14} = 9**
		S_{15}	3	3	2	**S_{15} = 8**
			a_3b_1 = 18	**a_3b_2 = 16**	**a_3b_3 = 7**	**a_3 = 41**
			b_1 = 50	**b_2 = 50**	**b_3 = 46**	**Total = 146**

The *simplified* calculation formula

$$SS_A = \frac{\sum A^2}{bs} - \frac{T^2}{abs} = \frac{63^2 + 42^2 + 41^2}{3(5)} - \frac{146^2}{3(3)(5)}$$

$$SS_{S/A} = \frac{\sum (AS)^2}{b} - \frac{\sum A^2}{bs} = \frac{10^2 + 13^2 + 12^2 + \cdots + 8^2}{3} - \frac{63^2 + 42^2 + 41^2}{3(5)}$$

$$SS_B = \frac{\sum B^2}{as} - \frac{T^2}{abs} = \frac{50^2 + 50^2 + 46^2}{3(5)} - \frac{146^2}{3(3)(5)}$$

$$SS_{AB} = \frac{\sum (AB)^2}{s} - \frac{\sum A^2}{bs} - \frac{\sum B^2}{as} + \frac{T^2}{abs} = \frac{12^2 + 19^2 + 32^2 + 20^2 + 15^2 + 7^2 + 18^2 + 16^2 + 7^2}{5} - \frac{63^2 + 42^2 + 41^2}{3(5)} - \frac{50^2 + 50^2 + 46^2}{3(5)} + \frac{146^2}{3(3)(5)}$$

$$SS_{B \times S/A} = \sum Y^2 - \frac{\sum (AB)^2}{s} - \frac{\sum (AS)^2}{b} + \frac{\sum A^2}{bs} = 1^2 + 1^2 + 3^2 + 5^2 + 2^2 + 3^2 + 4^2 + 5^2 + \cdots + 1^2 + 2^2 - \frac{12^2 + 19^2 + 32^2 + 20^2 + 15^2 + 7^2 + 18^2 + 16^2 + 7^2}{5} - \frac{10^2 + 13^2 + \cdots + 8^2}{3} + \frac{63^2 + 42^2 + 41^2}{3(5)}$$

$$SS_T = \sum Y^2 - \frac{T^2}{abs} = 1^2 + 1^2 + 3^2 + 5^2 + 2^2 + 3^2 + \cdots + 1^2 + 2^2 - \frac{146^2}{3(3)(5)}$$

$$
\begin{aligned}
SS_A &= 494.27 - 473.69 &&= 20.58 \\
SS_{S/A} &= 520.67 - 494.27 &&= 26.40 \\
SS_B &= 474.40 - 473.69 &&= 0.71 \\
SS_{AB} &= 566.40 - 494.27 - 474.40 + 473.69 &&= 71.42 \\
SS_{B \times S/A} &= 630 - 566.40 - 520.67 + 494.27 &&= 37.20 \\
SS_T &= 630 - 473.69 &&= 156.31
\end{aligned}
$$

$$df_A = a - 1 = 3 - 1 = 2$$

$$df_{S/A} = a(s - 1) = 3(5 - 1) = 12$$

$$df_B = b - 1 = 3 - 1 = 2$$

$$df_{AB} = (a - 1)(b - 1) = (3 - 1)(3 - 1) = 4$$

$$df_{BxS/A} = a(b - 1)(s - 1) = 3(3 - 1)(5 - 1) = 24$$

$$df_T = abs - 1 = N - 1 = 3(3)(5) - 1 = 44$$

Results – ANOVA summary table

Source	SS	df	MS	F
Randomized Groups				
A	20.58	2	10.29	$\frac{10.29}{2.20} = 4.68$
S/A	26.40	12	2.20	
Repeated Measures				
B	0.71	2	0.36	$\frac{0.36}{1.55} = 0.23$
$A \times B$	71.42	4	17.86	$\frac{17.86}{1.55} = 11.52$
$B \times S/A$	37.20	24	1.55	
T	156.31	44		

Post-hoc tests after obtaining significant effects

Between Groups

- If you have a significant BG main effect(s) they need to be broken down to find which levels are different
- The comparisons are done the same way as complete BG comparisons (remember one-way or two-way independent-measures ANOVA?)

Within Groups comparisons

- If a WG main effect is significant, it also needs to be followed by comparisons
- In the same way as in one-way repeated-measures ANOVA

Assumptions

- **Normality** of the distribution of residuals:
    - Applies to each condition (cell)
    - Possible outliers might lead to violations
    - Practically, only requires approximately normal
    - Shapiro-Wilk test (or Q-Q plot)
- **Homogeneity** of Variance
    - Each group (combination or cell) has the same variance
    - Should hold on each level of the within-group factor
    - Levene's F test

Assumptions

- **Sphericity** for within-group differences
  - Variance for the difference score between within-group conditions (factor B in our example) should be equal.
  - This is also an assumption for any repeated-measure ANOVA, including the one-way repeated-measure ANOVA we covered before.
  - This time, this equal variance of difference score is only tested for the repeated measures of the design, including interaction terms with repeated measures.
  - Using Mauchly’s sphericity, as we covered before.

R outputs for the example question

Levene's Test for Homogeneity of Variance:

	Levene's F	df1	df2	p
DV: time1	1.848	2	12	.200
DV: time2	0.681	2	12	.525
DV: time3	0.321	2	12	.731

Mauchly's Test of Sphericity:

	Mauchly's W	p
time	0.7890	.272
group x time	0.7890	.272

	MS	MSE	df1	df2	F	p		η^2p [90% CI of η^2p]	η^2G
group	10.289	2.200	2.000	12.000	4.677	.031	*	.438 [.034, .659]	.244
time	0.431	1.877	1.651	19.818	0.229	.755		.019 [.000, .152]	.011
group x time	21.624	1.877	3.303	19.818	11.520	<.001	***	.658 [.388, .775]	.529

What if the interaction is significant?

- ***Only*** reporting significant main effect is misleading (remember we cannot take the face value of main effect?).
- We will need to look at the profile plot and interpret the difference between groups (between-subject factor) accordingly.
- Use **simple main effect** to check group difference on each level (i.e., each month here).

Post-hoc tests for main effect

Note here main effect should be interpreted with caution since the interaction is significant.

t

```
Pairwise Comparisons of "group":
────────────────────────────────────────────────────────────────────────────────────
        Contrast  Estimate    S.E.  df       t      p      Cohen’s d [95% CI of d]
────────────────────────────────────────────────────────────────────────────────────
  group2 - group1    -1.400  (0.542)  12  -2.585   .072  .     -0.859 [-1.782, 0.065]
  group3 - group1    -1.467  (0.542)  12  -2.708   .057  .     -0.900 [-1.823, 0.024]
  group3 - group2    -0.067  (0.542)  12  -0.123  1.000        -0.041 [-0.964, 0.883]
────────────────────────────────────────────────────────────────────────────────────
Pooled SD for computing Cohen’s d: 1.630
Results are averaged over the levels of: time
P-value adjustment: Bonferroni method for 3 tests.
```

Simple main effect on **time**

------ EMMEANS (effect = "time") ------

Joint Tests of "time":

Effect	group	df1	df2	F	p		η²p [90% CI of η²p]
time	group1	2	12	18.219	<.001	***	.752 [.465, .857]
time	group2	2	12	7.356	.008	**	.551 [.148, .734]
time	group3	2	12	6.941	.010	**	.536 [.131, .725]

Post-hoc tests for simple main effect on **time**

Pairwise Comparisons of "time":

Contrast	group	Estimate	S.E.	df	t	p		Cohen's d [95% CI of d]
time2 - time1	group1	1.400	(0.917)	12	1.528	.458		0.859 [-0.704, 2.422]
time3 - time1	group1	4.000	(0.816)	12	4.899	.001	**	2.454 [1.062, 3.846]
time3 - time2	group1	2.600	(0.594)	12	4.374	.003	**	1.595 [0.581, 2.609]
time2 - time1	group2	-1.000	(0.917)	12	-1.091	.890		-0.613 [-2.176, 0.949]
time3 - time1	group2	-2.600	(0.816)	12	-3.184	.024	*	-1.595 [-2.987, -0.203]
time3 - time2	group2	-1.600	(0.594)	12	-2.692	.059	.	-0.982 [-1.995, 0.032]
time2 - time1	group3	-0.400	(0.917)	12	-0.436	1.000		-0.245 [-1.808, 1.317]
time3 - time1	group3	-2.200	(0.816)	12	-2.694	.059	.	-1.350 [-2.742, 0.043]
time3 - time2	group3	-1.800	(0.594)	12	-3.028	.032	*	-1.104 [-2.118, -0.091]

Simple main effect on **group**

```
------ EMMEANS (effect = "group") ------
```

Joint Tests of "group":

Effect	time	df1	df2	F	p		η^2p [90% CI of η^2p]
group	time1	2	12	1.857	.198		.236 [.000, .500]
group	time2	2	12	0.441	.654		.068 [.000, .291]
group	time3	2	12	28.409	<.001	***	.826 [.612, .900]

Post-hoc tests for simple main effect on **group**

Pairwise Comparisons of "group":

Contrast	time	Estimate	S.E.	df	t	p		Cohen's d [95% CI of d]
group2 - group1	time1	1.600	(0.864)	12	1.852	.266		0.982 [-0.492, 2.455]
group3 - group1	time1	1.200	(0.864)	12	1.389	.570		0.736 [-0.737, 2.210]
group3 - group2	time1	-0.400	(0.864)	12	-0.463	1.000		-0.245 [-1.719, 1.228]
group2 - group1	time2	-0.800	(0.887)	12	-0.902	1.000		-0.491 [-2.003, 1.022]
group3 - group1	time2	-0.600	(0.887)	12	-0.676	1.000		-0.368 [-1.880, 1.144]
group3 - group2	time2	0.200	(0.887)	12	0.225	1.000		0.123 [-1.390, 1.635]
group2 - group1	time3	-5.000	(0.766)	12	-6.528	<.001	***	-3.067 [-4.373, -1.761]
group3 - group1	time3	-5.000	(0.766)	12	-6.528	<.001	***	-3.067 [-4.373, -1.761]
group3 - group2	time3	-0.000	(0.766)	12	-0.000	1.000		-0.000 [-1.306, 1.306]

Higher order mixed designs

between within

Design	Randomized-Groups IVs	Repeated-Measures IVs	Sources of Variability		
			Randomized Groups	*Repeated Measures*	*Error Terms*
Two-way mixed	A	B	A		S/A
				$B, A \times B$	$B \times S/A$
Three-way mixed	A, B	C	$A, B, A \times B$		S/AB
				$C, A \times C, B \times C, A \times B \times C$	$C \times S/AB$
	A	B, C	A		S/A
				$B, A \times B$	$B \times S/A$
				$C, A \times C$	$C \times S/A$
				$B \times C, A \times B \times C$	$B \times C \times S/A$
Four-way mixed	A, B	C, D	$A, B, A \times B$		S/AB
				$C, A \times C, B \times C, A \times B \times C$	$C \times S/AB$
				$D, A \times D, B \times D, A \times B \times D$	$D \times S/AB$
				$C \times D, A \times C \times D,$ $B \times C \times D, A \times B \times C \times D$	$C \times D \times S/AB$

Another example for R coding

no	group	day1	day2	day3	day4	day5
1	1	26	20	18	11	10
2	1	34	35	29	22	23
3	1	41	37	25	18	15
4	1	29	28	22	15	13
5	1	35	34	27	21	17
6	1	28	22	17	14	10
7	1	38	34	28	25	22
8	1	43	37	30	27	25
9	2	42	38	26	20	15
10	2	31	27	21	18	13
11	2	45	40	33	25	18
12	2	29	25	17	13	8
13	2	29	32	28	22	18
14	2	33	30	24	18	7
15	2	34	30	25	24	23
16	2	37	31	25	22	20

How learning a memory technique improves memory:

- Two memory techniques are compared
- 16 subjects, divided into two groups
- Measured 5 times in 5 days