kxk Within-Groups Factorial ANOVA

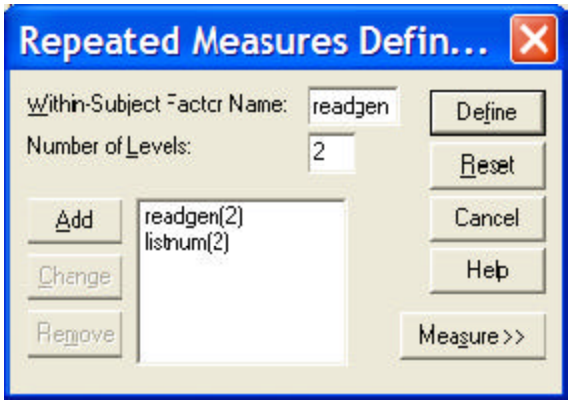
Application: Examination of main effects and interaction relating two IVs (with 2 or more within-groups conditions each) to a single quantitative DV.

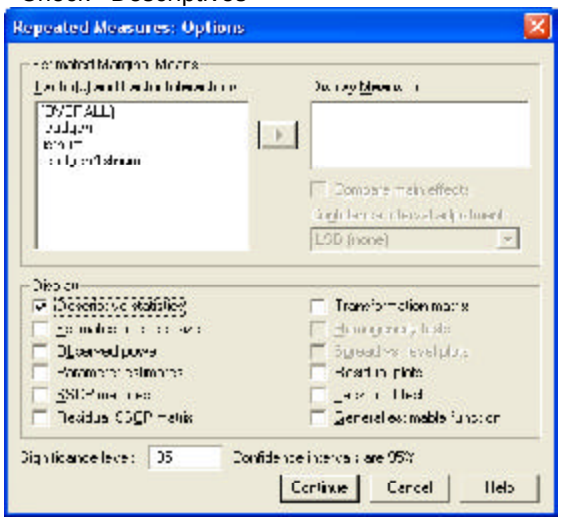
Research Hypothesis: The researcher hypothesized that there would be an interaction between Type of Practice and List Number. Specifically, the expected pattern was that the generation condition would always lead to better scores, but that this effect would be stronger on the earlier lists. The researcher also hypothesized that there would be main effects for List Number, such that scores would go up on each successive list.

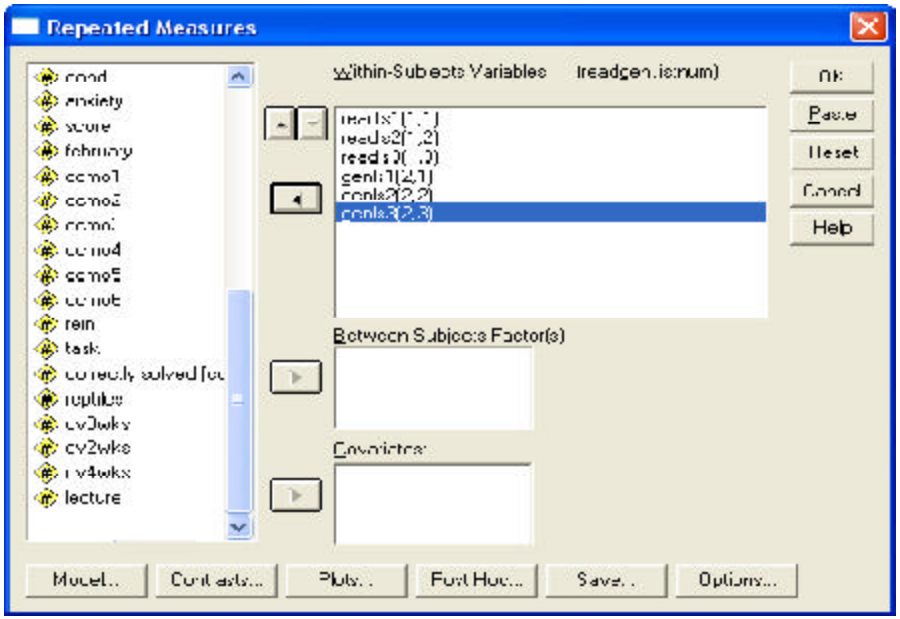
A bit of explanation: In this study of paired comparison learning, subjects were presented with three lists to learn. One-half of the items on each list were "read" (both words were presented, e.g., GUN BUN) and one-half were "generated" (the first word was presented and only the first letter of the second word - subjects "generated" the second words, based on the rule that the two words rhyme, e.g., CAT H_ _).

Analyze ➔ General Linear Model ➔ Repeated Measures

- Specify 1st IV
  - Type name of 1st IV in "Within-Subject Factor Name " window (e.g., readgen)
  - Type number of conditions of 1st IV (e.g., 2)
  - Press "Add" button
- Specify 2nd IV
  - Type name of 2nd IV in "Within-Subject Factor Name " window (e.g., listnum)
  - Type number of conditions of 2nd IV (e.g., 3)
  - Press "Add" button
- Press "Define" button
  - Highlight the variable holding the DV score in each combination of IV conditions and press the arrow, in turn (e.g., readls1 into (1,1), readls2 into (1,2), readls3 into (1,3), genls1 into (2,1), genls2 into (2,2), and genls3 into (2,3)
  - Be sure that you assign the correct "DV" to the correct "cell" of the WG design -- be sure to double check !!!
- Click "Options" button
  - Check "Descriptives"







Descriptive Statistics

	Mean	Std. Deviation	N
score for read words - list 1	7.6250	1.3025	8
score for read words - list 2	10.5000	1.6036	8
score for read words - list 3	14.3750	1.9226	8
score for gened words - list 1	13.1250	1.5526	8
score for gened words - list 2	15.1250	1.3562	8
score for gened words - list 3	17.5000	1.5119	8

The table below was constructed from the SPSS output -- unfortunately SPSS doesn't provide the magrinal means.

	List			
Condition	**1st**	**2nd**	**3rd**	
Read	**7.63**	**10.50**	**14.38**	**10.84**
Generate	**13.13**	**15.13**	**17.13**	**15.13**
	10.38	**12.82**	**15.76**	

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
READGEN	Sphericity Assumed	234.083	1	234.083	312.111	.000
	Greenhouse-Geisser	234.083	1.000	234.083	312.111	.000
	Huynh-Feldt	234.083	1.000	234.083	312.111	.000
	Lower-bound	234.083	1.000	234.083	312.111	.000
Error(READGEN)	Sphericity Assumed	5.250	7	.750		
	Greenhouse-Geisser	5.250	7.000	.750		
	Huynh-Feldt	5.250	7.000	.750		
	Lower-bound	5.250	7.000	.750		
LISTNUM	Sphericity Assumed	248.792	2	124.396	114.512	.000
	Greenhouse-Geisser	248.792	1.669	149.025	114.512	.000
	Huynh-Feldt	248.792	2.000	124.396	114.512	.000
	Lower-bound	248.792	1.000	248.792	114.512	.000
Error(LISTNUM)	Sphericity Assumed	15.208	14	1.086		
	Greenhouse-Geisser	15.208	11.686	1.301		
	Huynh-Feldt	15.208	14.000	1.086		
	Lower-bound	15.208	7.000	2.173		
READGEN * LISTNUM	Sphericity Assumed	11.542	2	5.771	3.824	.047
	Greenhouse-Geisser	11.542	1.691	6.826	3.824	.058
	Huynh-Feldt	11.542	2.000	5.771	3.824	.047
	Lower-bound	11.542	1.000	11.542	3.824	.091
Error(READGEN*LISTNUM)	Sphericity Assumed	21.125	14	1.509		
	Greenhouse-Geisser	21.125	11.836	1.785		
	Huynh-Feldt	21.125	14.000	1.509		
	Lower-bound	21.125	7.000	3.018		

SPSS provides different "versions" of the ANOVA output. We will focus on the "traditional" analysis, which SPSS labels as "Sphericity Assumed"

df(cond), F and p-values for Type of Practice main effect

df(error), MSe for the Type of Practice main effect

df(cond), F and p-values for List Number main effect

df(error), MSe for the List Number main effect

df(cond), F and p-values for Type of Practice x List Number interaction

df(error), MSe for the Type of Practice x List Number interaction

We will use LSD minimum mean differences to further analyze the data. There are three significant effects (main effect of Practice Type, main effect ofList Number and the interaction), so we might need as many as three d_{LSD} values. However, since the main effect of Practice Type has only two conditions (Read vs. Generate), we will not need any type of follow-up analyses to compare the marginal means -- we need only compare the direction of the significant mean difference with the RH:.

For the Interaction

based on df(error) = 14, **t = 2.15** also **n = 8** **MS(error) = 1.51**

$$d_{LSD} = \frac{t * \sqrt{(2 * MS_{Error})}}{\sqrt{n}} = \frac{2.15 * \sqrt{(2 * 1.51)}}{\sqrt{8}} = \mathbf{1.32}$$

Remember: n is based on the average number of data points making up each mean -- N = 48 and there are 6 conditions in the design, so n = N/k = 48/6 = 8

Applying this d_{LSD} to the cell means ...

SE of Type of Practice:

For List 1	Generate (13.13)	>	Read (7.63)
For List 2	Generate (15.13)	>	Read (10.50)
For List 3	Generate (17.13)	>	Read (14.38)

SE of List Number

For Generate	1(13.13)	>	2 (15.13)	2 >	3 (17.13)	1 > 3
For Read	1 (7.63)	>	2(10.50)	2 >	3 (14.38)	1 > 3

Remember, we need only one set of SEs to describe the pattern of the interaction, but we need each set to evaluate the descriptiveness of the corresponding main effect.

For the main effect of List Number

based on df(error) = 14, **t = 2.15** also **n = 16** **MS(error) = 1.09**

$$d_{LSD} = \frac{t * \sqrt{(2 * MS_{Error})}}{\sqrt{n}} = \frac{2.15 * \sqrt{(2 * 1.09)}}{\sqrt{16}} = \mathbf{.79}$$

Remember: n is based on the average number of data points making up each mean -- N = 48 and there are 3 conditions of the List Number IV, so n = N/k = 48/3 = 16

Applying this d_{LSD} to the marginal means ...

List 1(10.38) > List 2 (12.82) List 2 >List 3 (15.76) List 1 > List 3

We need to compare this pattern to those of the simple effects for Type of Reinforcement, to determine if the main effect is descriptive or potentially misleading.

As you can see, the main effect pattern corresponds with both the simple effect of List for the Generate and the Read conditions, and so is descriptive as a general statement.

Reporting the Results:

A within-groups factorial ANOVA with follow-up analyses using the LSD procedure (p = .05) was performed to examine the effects of Type of Practice and List Number upon performance on a paired-associate learning task. Table 1 shows the means for each condition of the design.

There was an interaction of Type or Practice and List Number as they relate to perormance (F(2, 14) = 5.771, p = .047, Mse = 1.51). As hypothesized, the pattern of this interaction was that while performance was consistently better in the generate condition than in the read condition, this effect was smaller for each successive list (LSD minimum mean difference = 1.32).

There was a main effect of Type or Practice (F(1, 7) = 312.11, p = .001, Mse = .750), with better overall performance in the generate than in the read condition, as hypothesized. There was also a main effect of List Number (F(2,24) = 114.512, p = .001). The pattern of the mean differences was that, as hypotheized, performance got better with each successive list (LSD minimum mean difference = .93).