Chapter 14 – Repeated-Measures Designs

[As in previous chapters, there will be substantial rounding in these answers. I have attempted to make the answers fit with the correct values, rather than the exact results of the specific calculations shown here. Thus I may round cell means to two decimals, but calculation is carried out with many more decimals.]

14.1 Does taking the GRE repeatedly lead to higher scores?

a. Statistical model:

$$X_{ij} = \mu + \pi_i + \tau_j + \pi\tau_{ij} + e_{ij} \text{ or } X_{ij} = \mu + \pi_i + \tau_j + e_{ij}'$$

b. Analysis:

Subject	Mean
1	566.67
2	450.00
3	616.67
4	663.33
5	436.67
6	696.67
7	503.33
8	573.33
Mean	563.33

Test Session	Mean
1	552.50
2	563.75
3	573.75

$$SS_{total} = \Sigma X^2 - \frac{(\Sigma X)^2}{N}$$

$$= 7811200 - \frac{(13520)^2}{24} = 194933.33$$

$$SS_{subj} = t\Sigma\left(\bar{X}_{i.} - \bar{X}_{..}\right)^2$$

$$= 3[(566.67 - 563.33)^2 + ... + (573.33 - 563.33)^2] = 3(63222.22) = 189,666.67$$

$$SS_{test} = n\Sigma\left(\bar{X}_{.j} - \bar{X}_{..}\right)^2 = 8[(552.50 - 563.33)^2 + (563.75 - 563.33)^2 + (573.75 - 563.33)^2]$$

$$= 8[226.04] = 1808.33$$

$$SS_{error} = SS_{total} - SS_{subj} - SS_{test}$$

$$= 194,933.33 - 189,666.67 - 1808.33 = 3458.33$$

Source	*df*	SS	MS	*F*
Subjects	7	189,666.66		
Within subj	16	5266.67		
Test session	2	1808.33	904.17	3.66 ns

Error	14	3458.33	247.02	
Total	23	194,933.33		

14.3 Teaching of self-care skills to severely retarded children:

Cell means:

		Phase		
		Baseline	Training	Mean
Group:	Exp	4.80	7.00	5.90
	Control	4.70	6.40	5.55
	Mean	4.75	6.70	5.72

Subject means:

		S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}
Grp	Exp	8.5	6.0	2.5	6.0	5.5	6.5	6.5	5.5	5.5	6.5
	Control	4.0	5.0	9.0	3.5	4.0	8.0	7.5	4.5	5.0	5.5

$\Sigma X^2 = 1501$ $\Sigma X = 229$ $N = 40$ $n = 10$ $g = 2$ $p = 2$

$$SS_{total} = \Sigma X^2 - \frac{(\Sigma X)^2}{N} = 1501 - \frac{(229)^2}{40} = 189.975$$

$$SS_{subj} = p\Sigma\left(\bar{X}_{ij.} - \bar{X}_{...}\right)^2$$
$$= 2[(8.5 - 5.72)^2 + ... + (5.5 - 5.72)^2] = 106.475$$

$$SS_{group} = pn\Sigma\left(\bar{X}_{..k} - \bar{X}_{...}\right)^2$$
$$= 2(8)[(5.90 - 5.72)^2 + (5.55 - 5.72)^2] = 1.225$$

$$SS_{phase} = gn\Sigma\left(\bar{X}_{.j.} - \bar{X}_{...}\right)^2$$
$$= 2(10)[(4.75 - 5.72)^2 + (6.70 - 5.72)^2] = 38.025$$

$$SS_{cells} = n\Sigma\left(\bar{X}_{.jk} - \bar{X}_{...}\right)^2$$
$$= 10\left[(4.80 - 5.72)^2 + ... + (6.40 - 5.72)^2\right] = 39.875$$

$$SS_{PG} = SS_{cells} - SS_{phase} - SS_{group} = 39.875 - 38.025 - 1.225 = 0.925$$

Source	*df*	SS	MS	*F*
Between Subj	19	106.475		
Groups	1	1.125	1.125	0.19
Ss w/in Grps	18	105.250	5.847	
Within Subj	20	83.500		
Phase	1	38.025	38.025	15.26*
P x G	1	0.625	0.625	0.25
P x Ss w/in Grps	18	44.850	2.492	
Total	39	189.975		

$* p < .05 \quad [F_{.05(1,18)} = 4.41]$

There is a significant difference between baseline and training, but there are no group differences nor a group x phase interaction.

14.5 Adding a No Attention control group to the study in Exercise 14.3:

Cell means:

		Phase		
		Baseline	Training	Total
Group	Exp	4.8	7.0	5.90
	Att Cont	4.7	6.4	5.55
	No Att Cont	5.1	4.6	4.85
	Total	4.87	6.00	5.43

Subject means:		S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}
Group:	Exp	8.5	6.0	2.5	6.0	5.5	6.5	6.5	5.5	5.5	6.5
	Att Cont	4.0	5.0	9.0	3.5	4.0	8.0	7.5	4.5	5.0	5.0
	No Att Cont	3.5	5.0	7.0	5.5	4.5	6.5	6.5	4.5	2.5	3.0

$\Sigma X^2 = 2026 \qquad \Sigma X = 326 \qquad N = 60 \qquad n = 10 \qquad g = 3 \qquad p = 2$

$$SS_{Total} = \Sigma X^2 - \frac{(\Sigma X)^2}{N} = 2026 - \frac{(326)^2}{40} = 254.7333$$

$$SS_{subj} = p\Sigma\left(\bar{X}_{ij.} - \bar{X}_{...}\right)^2$$

$$= 2[(8.5 - 5.43)^2 + ... + (3.0 - 5.43)^2] = 159.733$$

$$SS_{group} = pn\Sigma\left(\bar{X}_{..k} - \bar{X}_{...}\right)^2$$

$$= 2(8)[(5.90-5.43)^2 + (5.55-5.43)^2 + (4.85-5.43)^2] = 11.433$$

$$SS_{phase} = gn\Sigma\left(\bar{X}_{.j.} - \bar{X}_{...}\right)^2$$

$$= 3(10)[(4.87-5.43)^2 + (6.00-5.43)^2] = 19.267$$

$$SS_{cells} = n\Sigma\left(\bar{X}_{.jk} - \bar{X}_{...}\right)^2$$

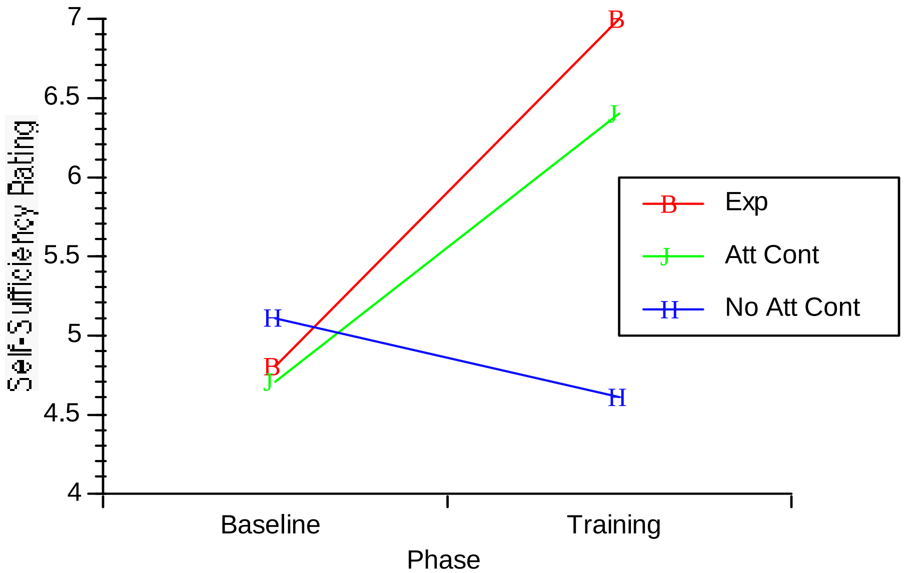
$$= 10\left[(4.80-5.43)^2 + \ldots + (4.60-5.43)^2\right] = 52.333$$

$$SS_{PG} = SS_{cells} - SS_{phase} - SS_{group} = 51.333 - 19.267 - 11.433 = 20.633$$

Source	*df*	SS	MS	*F*
Between subj	29	159.7333		
Groups	2	11.4333	5.7166	1.04
Ss w/ Grps	27	148.300	5.4926	
Within subj	30	95.0000		
Phase	1	19.2667	19.2667	9.44*
P * G	2	20.6333	10.3165	5.06*
P * Ss w/Grps	27	55.1000	2.0407	
Total	59	254.733		

$* p < .05$ $[F_{.05(1,27)} = 4.22;\ F_{.05(2,27)} = 3.36]$

b. Plot:



c. There seems to be no difference between the Experimental and Attention groups, but both show significantly more improvement than the No Attention group.

14.7 For the data in Exercise 14.6:

a. Variance-covariance matrices:

$$\hat{\Sigma}_{owners} = \begin{bmatrix} 1.30 & 1.50 & 0.75 \\ 1.50 & 2.00 & 1.00 \\ 0.75 & 1.00 & 1.00 \end{bmatrix} \qquad \hat{\Sigma}_{non-owners} = \begin{bmatrix} 2.70 & 1.20 & 1.85 \\ 1.20 & 0.70 & 0.60 \\ 1.85 & 0.60 & 3.30 \end{bmatrix}$$

$$\hat{\Sigma}_{pooled} = \begin{bmatrix} 2.00 & 1.35 & 1.30 \\ 1.35 & 1.35 & 0.80 \\ 1.30 & 0.80 & 2.15 \end{bmatrix} \begin{matrix} \bar{s}_j \\ 1.550 \\ 1.167 \\ 1.417 \end{matrix} \qquad \hat{\Sigma}_{between} = \begin{bmatrix} 0.18 & 0.36 & 1.38 \\ 0.36 & 0.72 & 2.76 \\ 1.38 & 2.76 & 10.58 \end{bmatrix}$$

b. $\hat{e}$

$$\bar{s}_{jj} = \frac{2.00 + 1.35 + 2.15}{3} = 1.833$$

$$\bar{s} = \frac{2.00 + ... + 2.15}{9} = 1.378$$

$$\Sigma s_{jk}^2 = 2.00^2 + \ldots + 2.15^2 = 18.750$$

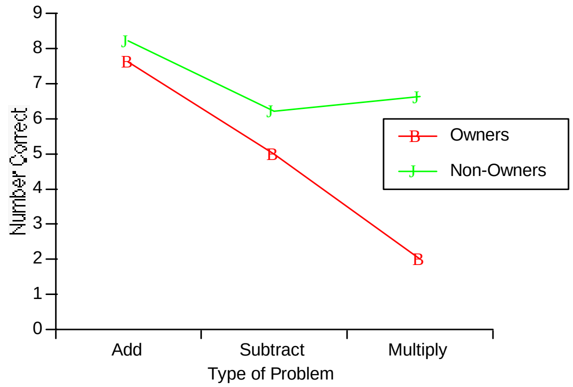
$$\Sigma \overline{s}_j^2 = 1.550^2 + 1.167^2 1.417^2 = 18.750$$

$$\hat{e} = \frac{b^2\left(\overline{s}_{jj} - \overline{s}\right)^2}{(b-1)\left(\Sigma s_{jk}^2 - 2b\Sigma\overline{s}_j^2 + b^2\overline{s}^2\right)} = \frac{9(1.833-1.378)^2}{2\left[18.75-6(5.772)+9(1.378^2)\right]}$$

$$= \frac{1.863}{2.406} = 0.771$$

14.9 Back to the calculator usage data in Exercise 14.6:

a. Plot



b. Analysis of simple effects:

Cell means:

	Add	Subt	Mult	Mean
	Problems			
Owners	7.6	5.0	2.0	4.87
Non-Owners	8.2	6.2	6.6	7.00
Mean	7.9	5.6	4.3	5.935

For the first three simple effects we're breaking down a combination of SS_{Groups} and SS_{PG} from the overall analysis in Exercise 14.6, so for the error term we will need to combine the error terms which were used to test MS_{Groups} and MS_{PG} in that analysis.

$$MS_{w/in Cell} = \frac{MS_{Ss\ w/in\ grps} + MS_{P*Ss\ w/in\ grps}}{df_{Ss\ w/in\ grps} + df_{P*Ss\ w/in\ grps}} = \frac{33.066 + 10.934}{8 + 16} = 1.833$$

The *df* against which to evaluate *F* also must be adjusted for these simple effects. F_{obt} will be evaluated against $F_{.05(g-1, f')}$ but we must first calculate f'.

$$f' = \frac{(u+v)^2}{\frac{u^2}{df_u} + \frac{v^2}{df_v}} = \frac{(33.066 + 10.934)^2}{\frac{33.066^2}{8} + \frac{10.934^2}{16}} = 13.431 \qquad u = SS_{Ss\ w/in\ grps}$$

$$v = SS_{P*Ss\ w/in\ grps}$$

With the error terms and degrees of freedom ready, we go ahead with calculating the sums of squares and testing them:

$$SS_{group\ \text{at}\ Add} = n\Sigma\left(\bar{X}_{i1} - \bar{X}_{.1}\right)^2$$
$$= 5[(7.6-7.9)^2 + (8.2-7.9)^2] = 0.9$$
$$MS_{group\ \text{at}\ Add} = \frac{SS_{group\ \text{at}\ Add}}{df_{group\ \text{at}\ Add}} = \frac{0.9}{1} = 0.9$$
$$F_{group\ \text{at}\ Add} = \frac{MS_{group\ \text{at}\ Add}}{MS_{w/in\ cells}} = \frac{0.9}{1.833} = < 1$$

$$SS_{group\ at\ Subt} = n\Sigma\left(\bar{X}_{i2} - \bar{X}_{.2}\right)^2$$
$$= 5[(5.0-5.6)^2 + (6.0-5.6)^2] = 3.6$$
$$MS_{group\ at\ Subt} = \frac{SS_{group\ at\ Subt}}{df_{group\ at\ Subt}} = \frac{3.6}{1} = 3.6$$
$$F_{group\ at\ Subt} = \frac{MS_{group\ at\ Subt}}{MS_{w/in\ cells}} = \frac{3.6}{1.833} = 1.96 ns$$

$$SS_{group\ at\ Mult} = n\Sigma\left(\bar{X}_{i3} - \bar{X}_{.3}\right)^2$$
$$= 5[(2.0-4.3)^2 + (6.6-4.3)^2] = 52.9$$
$$MS_{group\ at\ Mult} = \frac{SS_{group\ at\ Mult}}{df_{group\ at\ Mult}} = \frac{52.9}{1} = 52.9$$
$$F_{group\ at\ Mult} = \frac{MS_{group\ at\ Mult}}{MS_{w/in\ cells}} = \frac{52.9}{1.833} = 28.86*$$

$$*p < .05 \quad [F_{.05(g-1,f')} = F_{.05(1,13)} = 4.67]$$

$SS_{\text{prob at calc}}$ = 78.53 F = 112.19*

$SS_{\text{prob at noncalc}}$ = 11.20 F = 5.51*

For the last two simple effects we're breaking down a combination of SS_{Problems} and SS_{PG} from the overall analysis in Exercise 14.6. Since MS_{Problems} and MS_{PG} were both tested by the same error term in that analysis ($MS_{\text{P*Ss w/in grps}}$) we could use that error term to test these simple effects. However, as pointed out in the chapter, violations of sphericity create serious problems when testing simple effects, and for that reason we will use separate error terms for the two analyses. The easiest way to do this is to run separate repeated measures analysis of variance for each group. This will produce the same sums of squares for the simple effect, as well as the appropriate error term.

The following results were produced by SPSS. Notice that the form of the printout is quite different from what we usually have. The corrected *df* (assuming a lack of sphericity) are given, as is the resulting significance level. (SPSS adjusts the two mean squares as the *df* are adjusted, but that does not alter the resulting *F*.)

Calculator owners:

Tests of Within-Subjects Effects[a]

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
PROBLEM	Sphericity Assumed	78.533	2	39.267	112.190	.000
	Greenhouse-Geisser	78.533	1.477	53.157	112.190	.000
	Huynh-Feldt	78.533	2.000	39.267	112.190	.000
	Lower-bound	78.533	1.000	78.533	112.190	.000
Error(PROBLEM)	Sphericity Assumed	2.800	8	.350		
	Greenhouse-Geisser	2.800	5.910	.474		
	Huynh-Feldt	2.800	8.000	.350		
	Lower-bound	2.800	4.000	.700		

a. Calculator Owner = Owner

Non-owners

Tests of Within-Subjects Effects[a]

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
PROBLEM	Sphericity Assumed	11.200	2	5.600	5.508	.031
	Greenhouse-Geisser	11.200	1.564	7.159	5.508	.047
	Huynh-Feldt	11.200	2.000	5.600	5.508	.031
	Lower-bound	11.200	1.000	11.200	5.508	.079
Error(PROBLEM)	Sphericity Assumed	8.133	8	1.017		
	Greenhouse-Geisser	8.133	6.258	1.300		
	Huynh-Feldt	8.133	8.000	1.017		
	Lower-bound	8.133	4.000	2.033		

a. Calculator Owner = NonOwner

The *F* is significant in both cases, indicating that Task made a difference. If we had used a pooled error term, MS_{error} would have been 0.683, which, because of the fact that the sample sizes were equal, is the average of the two error terms we used. But notice that the pooled error term would have been about 60% of what it was when we treat the groups separately.

14.11 From Exercise 14.10:

a. Simple effect of reading ability for children:

$$SS_{R\,at\,C} = in\Sigma\left(\bar{X}_{R\,at\,C} - \bar{X}_{C}\right)^2$$

$$= 3(5)[(4.80 - 3.50)^2 + (2.20 - 3.50)^2] = 50.70$$

$$MS_{R\,at\,C} = \frac{SS_{R\,at\,C}}{df_{R\,at\,C}} = \frac{50.70}{1} = 50.70$$

Because we are using only the data from Children, it would be wise not to use a pooled error term. The following is the relevant printout from SPSS for the Between-subject effect of Reader.

Tests of Between-Subjects Effects[a]

Measure: MEASURE_1

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	367.500	1	367.500	84.483	.000
READERS	50.700	1	50.700	11.655	.009
Error	34.800	8	4.350		

a. AGE = Children

b. Simple effect of items for adult good readers:

$$SS_{I\,at\,AG} = n\Sigma\left(\bar{X}_{I\,at\,AG} - \bar{X}_{AG}\right)^2$$

$$= 5[(6.20 - 5.73)^2 + (6.00 - 5.73)^2 + (5.00 - 5.73)^2] = 4.133$$

Again, we do not want to pool error terms. The following is the relevant printout from SPSS for Adult Good readers. The difference is not significant, nor would it be for any decrease in the *df* if we used a correction factor.

Tests of Within-Subjects Effects

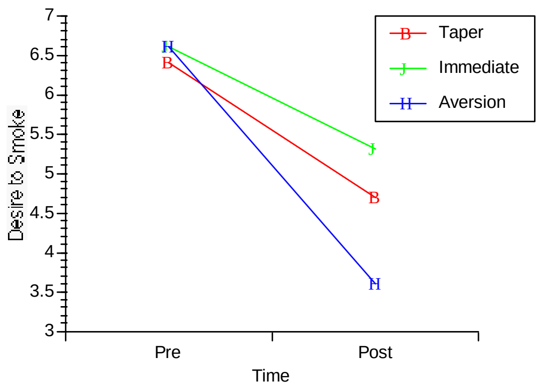
Measure: MEASURE_1

Sphericity Assumed

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
ITEMS	4.133	2	2.067	3.647	.075
Error(ITEMS)	4.533	8	.567		

14.13 It would certainly affect the covariances because we would force a high level of covariance among items. As the number of responses classified at one level of Item went up, another item would have to go down.

14.15 Plot of results in Exercise 14.14:



14.17 Analysis of data in Exercise 14.5 by BMDP:

a. Comparison with results obtained by hand in Exercise 14.5.

b. The *F* for Mean is a test on H_0: $\mu = 0$.

c. $MS_{w/in\ \text{Cell}}$ is the average of the cell variances.

14.19 Source column of summary table for 4-way ANOVA with repeated measures on A & B and independent measures on C & D.

Source
Between *S*s
C
D
CD
*S*s w/in groups
Within Ss
A
AC
AD
ACD
A x *Ss* w/in groups
B
BC
BD
BCD
B x *S*s w/in groups

AB
ABC
ABD
ABCD
AB x *S*s w/in groups

Total

14.21 Using the mixed models procedure on data from Exercise 14.20

If we assume that sphericity is a reasonable assumption, we could run the analysis with covtype(cs). That will give us the following, and we can see that the *F*'s are the same as they were in our analysis above.

Fixed Effects

Type III Tests of Fixed Effects[a]

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	42.000	450.019	.000
Group	2	42.000	3.749	.032
Time	2	84	73.534	.000
Group * Time	4	84	4.058	.005

a. Dependent Variable: dv.

However, the correlation matrix below would make us concerned about the reasonableness of a sphericity assumption. (This matrix is collapsed over groups, but reflects the separate matrices well.) Therefore we will assume an autoregressive model for our correlations.

Correlations

		Pre	Post	Followup
Pre	Pearson Correlation	1.000	.585**	.282
Post	Pearson Correlation	.585**	1.000	.616**
Followup	Pearson Correlation	.282	.616**	1.000

**. Correlation is significant at the 0.01 level (2-tailed).

Fixed Effects

Type III Tests of Fixed Effects[a]

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	43.256	422.680	.000
Group	2	43.256	3.521	.038
Time	2	81.710	71.356	.000
Group * Time	4	81.710	5.578	.001

a. Dependent Variable: dv.

These *F* values are reasonably close, but certainly not the same.

14.23 Mixed model analysis with unequal size example.

Fixed Effects

Type III Tests of Fixed Effects[a]

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	41.724	393.118	.000
Group	2	41.724	2.877	.068
Time	2	70.480	64.760	.000
Group * Time	4	70.459	5.266	.001

a. Dependent Variable: dv.

Notice that we have a substantial change in the *F* for Time, though it is still large.

14.25 Everitt's study of anorexia:

a. SPSS printout on gain scores:

Tests of Between-Subjects Effects

Dependent Variable: GAIN

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	614.644[a]	2	307.322	5.422	.006
Intercept	732.075	1	732.075	12.917	.001
TREAT	614.644	2	307.322	5.422	.006
Error	3910.742	69	56.677		
Total	5075.400	72			
Corrected Total	4525.386	71			

a. R Squared = .136 (Adjusted R Squared = .111)

b. SPSS printout using pretest and posttest:

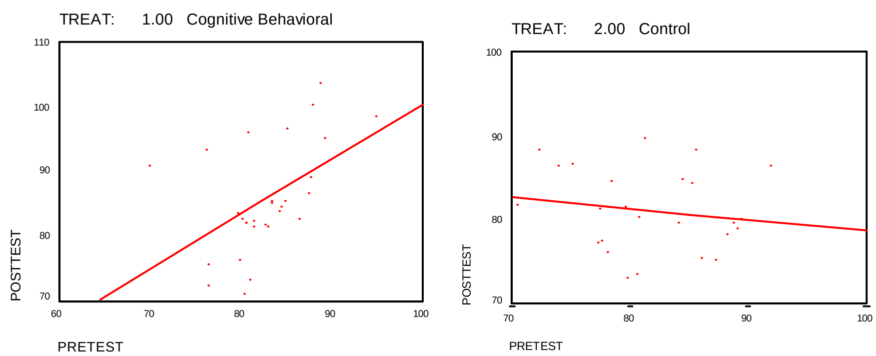
Tests of Within-Subjects Effects

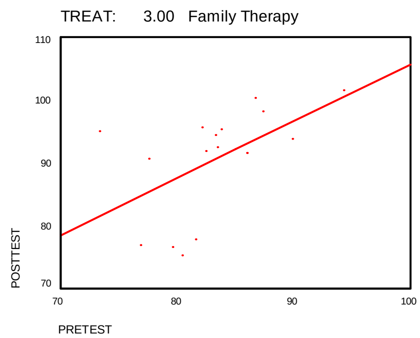
Measure: MEASURE_1
Sphericity Assumed

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
TIME	366.037	1	366.037	12.917	.001
TIME * TREAT	307.322	2	153.661	5.422	.006
Error(TIME)	1955.371	69	28.339		

c. The F comparing groups on gain scores is exactly the same as the F for the interaction in the repeated measures design.

d.





The plots show that there is quite a different relationship between the variables in the different groups.

e. Treatment Group = Control

One-Sample Statistics[a]

	N	Mean	Std. Deviation	Std. Error Mean
GAIN	26	-.4500	7.9887	1.5667

a. Treatment Group = Control

One-Sample Test[a]

	Test Value = 0					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
GAIN	-.287	25	.776	-.4500	-3.6767	2.7767

a. Treatment Group = Control

This group did not gain significantly over the course of the study. This suggests that any gain we see in the other groups can not be attributed to normal gains seen as a function of age.

f. Without the control group we could not separate gains due to therapy from gains due to maturation.

14.27 t = -0.555. There is no difference in Time 1 scores between those who did, and did not, have a score at Time 2.

b. If there had been differences, I would worried that people did not drop out at random.
to answer.

14.29 Differences due to Judges play an important role.

14.31 If I were particularly interested in differences between subjects, and recognized that judges probably didn't have a good anchoring point, and if this lack was not meaningful, I would not be interested in considering it.

14.33 Strayer et al. (2006)

Tests of Between-Subjects Effects

Measure:MEASURE_1
Transformed Variable:Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	7.711E7	1	7.711E7	724.691	.000
Error	4149966.533	39	106409.398		

Tests of Within-Subjects Effects

Measure:MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Condition	Sphericity Assumed	134696.067	2	67348.033	4.131	.020
	Greenhouse-Geisser	134696.067	1.992	67619.134	4.131	.020
	Huynh-Feldt	134696.067	2.000	67348.033	4.131	.020
	Lower-bound	134696.067	1.000	134696.067	4.131	.049
Error(Condition)	Sphericity Assumed	1271689.267	78	16303.709		
	Greenhouse-Geisser	1271689.267	77.687	16369.337		
	Huynh-Feldt	1271689.267	78.000	16303.709		
	Lower-bound	1271689.267	39.000	32607.417		

b) Contrasts on means:

Because the variances within each condition are so similar, I have used $MS_{error(within)}$ as my error term. The means are 776.95, 778.95, and 849.00 for Baseline, Alcohol, and Cell phone conditions, respectively..

$$t = \frac{\hat{\psi}}{\sqrt{\frac{\Sigma a_i^2 MS_{error}}{n}}}$$

$$\hat{\psi}_{1vs2} = 776.95 - 778.95 = 2$$

$$\hat{\psi}_{1vs3} = 776.95 - 849.00 = 72.05$$

$$\hat{\psi}_{2vs3} = 778.95 - 849.00 = 70.5$$

$$den = \sqrt{\frac{\Sigma a_i^2 MS_{error}}{n}} = \sqrt{\frac{2 \times 16303.709}{40}} = 28.551$$

$$t_{1vs2} = 2 / 28.551 = 0.07$$

$$t_{1vs3} = 72.05 / 28.551 = 2.52^{*}$$

$$t_{2vs3} = 70.05 / 28.551 = 2.45^{*}$$

Both Baseline and Alcohol conditions show poorer performance than the cell phone condition, but, interestingly, the Baseline and Alcohol conditions do not differ from each other.