REPEATED MEASURES ANOVA
(incomplete)

Repeated measures ANOVA (RM) is a specific type of MANOVA. When the within group covariance matrix has a special form, then the RM analysis usually gives more powerful hypothesis tests than does MANOVA. Mathematically, the within group covariance matrix is assumed to be a type H matrix (SAS terminology) or to meet Huynh-Feldt conditions. These mathematical conditions are given in the appendix. Most software for RM prints out both the MANOVA results and the RM results along with a test of RM assumption about the within group covariance matrix. Consequently, if the assumption is violated, one can interpret the MANOVA results. In practice, the MANOVA and RM results are usually similar.

There are certain stock situations when RM is used. The first occurs when the dependent variables all measure the same construct. Examples include a time series design of growth curves of an organism or the analysis of number of errors in discrete time blocks in an experimental condition. A second use of RM occurs when all the dependent variables are all measured on the same scale (e.g., the DVs are all Likert scale responses). For example, outcome from therapy might be measured on Likert scales reflecting different types of outcome (e.g., symptom amelioration, increase in social functioning, etc.). A third situation is for internal consistency analysis of a set of items or scales that purport to measure the same construct.. Internal consistency analysis consists in fitting a linear model to a set of items that are hypothesized to measure the same construct. For example, suppose that you have written ten items that you think measure the construct of empathy. Internal consistency analysis will provide measures of the extent to which these items "hang together" statistically and measure a single construct. (NOTE WELL: as in most stats, good internal consistency is just an index; *you* must be the judge of whether the single construct is empathy or something else.) If you are engaged in this type of scale construction, you should also use the SPSS subroutine RELIABILITY or SAS PROC CORR with the ALPHA option.

The MANOVA output from a repeated measures analysis is similar to output from traditional MANOVA procedures. The RM output is usually expressed as a "univariate" analysis, despite the fact that there is more than one dependent variable. The univariate RM has a jargon all its own which we will now examine by looking at a specific example.

An example of an RM design[1]

Suppose that 60 students studying a novel foreign language are randomly assigned to two conditions: a control condition in which the foreign language is taught in the traditional manner and a experimental condition in which the language is taught in an "immersed" manner where the instructor speaks only in the foreign language. Over the course of a semester, five tests of language mastery are given. The structure of the data is given in Table 1.

Table 1. Structure of the data for a repeated measures analysis of language instruction.

Student	Group	Test 1	Test 2`	Test 3	Test 4	Test 5
Abernathy	Control	17	22	26	28	31
.	.	.	.	.	.	.
Zelda	Control	18	24	25	30	29
Anasthasia	Experimental	16	23	28	29	34
.	.	.	.	.	.	.
Zepherinus	Experimental	23	25	29	38	47

The purpose of such an experiment is to examine which of the two instructional techniques is better. One very simple way of doing this is to create a new variable that is the sum of the 5 test scores and perform a *t*-test. The SAS code would be

```
DATA rmex1;
   INFILE 'c:\sas\p7291dir\repeated.simple.dat';
   LENGTH group $12.;
   INPUT subjnum group test1-test5;
   testtot = sum(of test1-test5);
RUN;

PROC TTEST;
   CLASS group;
   VAR testtot;
RUN;
```

The output from this procedure would be:

```
TTEST PROCEDURE
Variable: TESTTOT
```

[1] The SAS code for analyzing this example may be found on ~carey/p7291dir/repeated.simple.2.sas.

```
GROUP              N              Mean           Std Dev         Std Error
------------------------------------------------------------------------
Control           30      154.43333333       27.96510551        5.10570637
Experimental      30      170.50000000       32.91237059        6.00894926

Variances        T       DF    Prob>|T|
---------------------------------------
Unequal    -2.0376     56.5      0.0463
Equal      -2.0376     58.0      0.0462

For H0: Variances are equal, F' = 1.39    DF = (29,29)    Prob>F' = 0.3855
```

The mean total test score for the experimental group (170.5) is greater than the mean total test score for controls (154.4). The difference is significant (t = -2.04, df = 58, $p < .05$) so we should conclude that the experimental language instruction is overall superior to the traditional language instruction.

There is nothing the matter with this analysis. It gives an answer to the major question posed by the research design and suggests that in the future, the experimental method should be adopted for foreign language instruction.

But the expense of time in generating the design of the experiment and collecting the data merit much more than this simple analysis. One very interesting question to ask is whether the means for the two groups change over time. Perhaps the experimental instruction is good initially but poor at later stages of foreign language acquisition. Or maybe the two techniques start out equally but diverge as the semester goes on. A simple way to answer these questions is to perform separate *t*-tests for each of the five tests. The code here would be:

```
PROC TTEST DATA=rmex1;
   CLASS GROUP;
   VAR test1-test5;
RUN;
```

Table 2. Group means and *t*-test results for the 5 tests.

	Means:			
Test	Control	Experimental	*t*	*p* <
1	20.17	18.80	.70	.49
2	27.8	29.10	-.59	.56
3	30.7	36.10	-2.66	.01
4	36.83	40.43	-1.90	.07
5	38.93	46.07	-3.35	.002

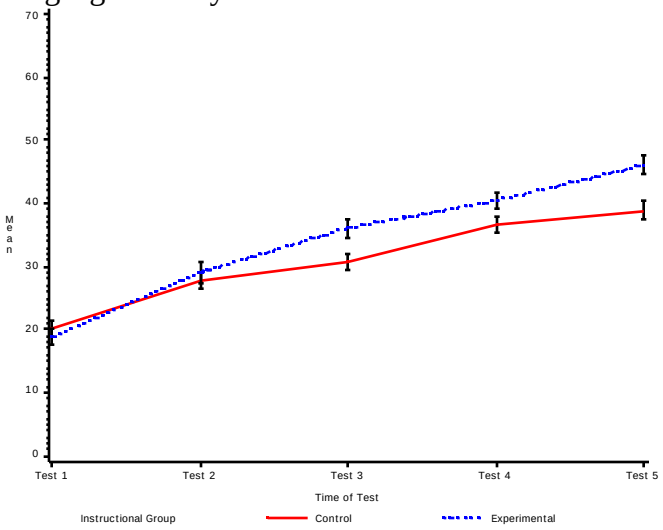
A summary of these results is given in Table 2 and a lot of the group means over time is given in Figure 1. Both the plot of the means and the *t*-test results suggest that the two groups start out fairly similar at times 1 and times 2, diverge significantly at time 3, are almost significantly different at time 4, and diverge again at time 5.

This analysis gives more insight, but leads to its own set of problems. We have performed 5 different significance tests. If these tests were independent—and they are clearly nor independent because they were performed on the same set of individuals—then we should adjust the α level by a Bonferroni formula

$$\alpha_{\text{adjusted}} = 1-.95^{\frac{1}{\text{number of tests}}} = 1-.95^{.2} = .01 .$$

Using this criterion, we would conclude that the differences in test 3 are barely significant, those in test 4 are not significant, while the means for the last test are indeed different. Perhaps the major reason why the two groups differ is only on the last exam in the course.

Figure 1. Means and standard errors for five tests of language mastery.



A repeated measures example can help to clarify the situation. The advantage to RM is that it will control for the correlations among the tests and come up with an overall test for each of the hypotheses given above. The RM design divides ANOVA factors into two types: *between subjects factors* (or effects) and *within subject factors* (or effects). If you think of the raw data matrix, you should have little trouble distinguishing the two. A single between subjects factor has one and only one value per observation. Thus, *group* is a between subjects factor because each observation in the data matrix has only one value for this variable (*Control* or *Experimental*).

Within subjects factors have more than one value per observation. Thus, *time of test* is a within subject factor because each subject has five difference values--Test1 through Test5.

Another way to look at the distinction is that between subject factors are all the independent variables. Within subject factors involve the dependent variables. All interaction terms that involve a within subject factor are included in within subject effects. Thus, interactions of *time of test* with *group* is a within subject effect. Interactions that include *only* between subject factors are included in between subjects effects.

The effects in a RM ANOVA are the same as those in any other ANOVA. In the present example, there would be a main effect for *group*, a main effect for *time* and an interaction between *group* and *time.* RM differs only in the mathematics used to compute these effects.

At the expense of putting the cart before the horse, the SAS commands to perform the repeated measures for this example are:

```
TITLE Repeated Measures Example 1;
PROC GLM DATA=rmex1;
  CLASS group;
  MODEL test1-test5 = group;
  MEANS group;
  REPEATED time 5 polynomial / PRINTM PRINTE SUMMARY;
RUN;
```

As in an ANOVA or MANOVA, the CLASS statement specifies the classification variable which is *group* in this case. The MODEL statement lists the dependent variables (on the left hand side of the equals sign) and the independent variables (on the right hand side). The MEANS statement asks that the sample size, means, and standard deviations be output for the two groups.

The novel statement in this example is the REPEATED statement. Later, this statement will be discussed in detail. The current REPEATED statement gives a name for the repeated measures or within subjects factor—*time* in this case—and the number of levels of that factor—5 in this example because the test was given over 5 time periods. The word *polynomial* instructs SAS to perform a polynomial transform of the 5 dependent variables. Essentially this creates 4 "new" variables from the original 5 dependent variables. (Any transformation of k dependent variables will result in k - 1 new transformed variables.) The PRINTM option prints the transformation matrix, the PRINTE option prints the error correlation matrix and some other important output, and the SUMMARY option prints ANOVA results for each of the four transformed variables.

Usually it is the transformation of the dependent variables that gives the RM analysis additional insight into the data. Transformations will be discussed at length later. Here we just note that the *polynomial* transformation literally creates 4 new variables from the 5 original dependent variables. The first of the new variables is the linear effect of time; it tests whether the means of the language mastery tests increase or decrease over time. The second new variable

is the quadratic effect of time. This new variable tests whether the means have a single "bend" to them over time. The third new variable is the cubic effect over time; this tests for two "bends" in the plot of means over time. Finally the fourth new variable is the quartic effect over time, and it tests for three bends in the means over time.

The first few pages of output from this procedure give the results from the univariate ANOVAs for test1 through test5. Because there are only two groups, the *F* statistics for these analysis are equal to the square of the *t* statistics in Table 2 and the *p* values for the ANOVAs will be the same as those in Table 2. Hence these results are not presented. The rest of the output begins with the MEANS statement.

```
---<PAGE>----------------------------------------------------------------------
Repeated Measures Example 1                                                7
General Linear Models Procedure

Level of          -----------TEST1----------   -----------TEST2----------
GROUP          N     Mean           SD            Mean           SD
Control       30   20.1666667    7.65679024     27.8000000    7.76996871
Experimental  30   18.8000000    7.46208809     29.1000000    9.12499410

Level of          -----------TEST3----------   -----------TEST4----------
GROUP          N     Mean           SD            Mean           SD
Control       30   30.7000000    6.88902175     36.8333333    7.06659618
Experimental  30   36.1000000    8.70334854     40.4333333    7.57347155

Level of          ------------TEST5------------
GROUP          N     Mean              SD
Control       30   38.9333333       8.65401375
Experimental  30   46.0666667       7.80332968

---<PAGE>----------------------------------------------------------------------
Repeated Measures Example 1                                                8
General Linear Models Procedure
Repeated Measures Analysis of Variance
Repeated Measures Level Information
```

The following section of output shows the design of the repeated measure factor. It is quite simple in this case. You should always check this to make certain that design specified on the REPEATED statement is correct.

```
Dependent Variable      TEST1    TEST2    TEST3    TEST4    TEST5
     Level of TIME          1        2        3        4        5

---<PAGE>----------------------------------------------------------------------
Repeated Measures Example 1                                                9
General Linear Models Procedure
Repeated Measures Analysis of Variance
```

These are the partial correlations for the dependent variables controlling for all the independent variables in the model. They are printed because the PRINTE option was specified and they answer the question, "To what extent is Test1 correlated with Test2 within each of the two groups?" Many of these correlations should be significant. Otherwise, there is really no sense in do repeated measures analysis—just do separate ANOVA on each dependent variable.

Partial Correlation Coefficients from the Error SS&CP Matrix / Prob > |r|

DF = 58	TEST1	TEST2	TEST3	TEST4	TEST5
TEST1	1.000000	0.451725	0.417572	0.510155	0.477928
	0.0001	0.0003	0.0010	0.0001	0.0001
TEST2	0.451725	1.000000	0.445295	0.599058	0.493430
	0.0003	0.0001	0.0004	0.0001	0.0001
TEST3	0.417572	0.445295	1.000000	0.650573	0.465271
	0.0010	0.0004	0.0001	0.0001	0.0002
TEST4	0.510155	0.599058	0.650573	1.000000	0.493323
	0.0001	0.0001	0.0001	0.0001	0.0001
TEST5	0.477928	0.493430	0.465271	0.493323	1.000000
	0.0001	0.0001	0.0002	0.0001	0.0001

Below is the transformation matrix. It is printed here because the PRINTM option was specified in the REPEATED statement. Because we specified a POLYNOMIAL transformation, this matrix gives coefficients for what are called *orthogonal polynomials*. They are analogous but not identical to contrast codes for independent variables. The first new variable, TIME.1, gives the linear effect over time, the second, TIME.2 is the quadratic effect, etc.

TIME.N represents the nth degree polynomial contrast for TIME

M Matrix Describing Transformed Variables

	TEST1	TEST2	TEST3	TEST4	TEST5
TIME.1	-.6324555320	-.3162277660	0.0000000000	0.3162277660	0.6324555320
TIME.2	0.5345224838	-.2672612419	-.5345224838	-.2672612419	0.5345224838
TIME.3	-.3162277660	0.6324555320	-.0000000000	-.6324555320	0.3162277660
TIME.4	0.1195228609	-.4780914437	0.7171371656	-.4780914437	0.1195228609

E = Error SS&CP Matrix

TIME.N represents the nth degree polynomial contrast for TIME

	TIME.1	TIME.2	TIME.3	TIME.4
TIME.1	1723.483333	-2.521377	236.550000	313.306091
TIME.2	-2.521377	2187.726190	98.502728	35.622692
TIME.3	236.550000	98.502728	1657.950000	-468.112779

```
TIME.4       313.306091        35.622692      -468.112779      1715.000476

---<PAGE>----------------------------------------------------------------------
Repeated Measures Example 1                                              10
General Linear Models Procedure
Repeated Measures Analysis of Variance
```

The within group partial correlation matrix, but this time for the transformed variables. If the transformation resulted in orthogonal variables, then the correlations in this matrix should be around 0. SAS checks on this by performing a *sphericity test* on the transformed variables. The sphericity test tests whether the error covariance matrix for the transformed variables is diagonal (i.e., all off diagonal elements are within sampling error of 0).

```
Partial Correlation Coefficients from the Error SS&CP Matrix
of the Variables Defined by the Specified Transformation / Prob > |r|

DF = 58       TIME.1    TIME.2    TIME.3    TIME.4
TIME.1      1.000000 -0.001298  0.139937  0.182236
              0.0001    0.9922    0.2905    0.1671

TIME.2     -0.001298  1.000000  0.051721  0.018391
              0.9922    0.0001    0.6972    0.8900

TIME.3      0.139937  0.051721  1.000000 -0.277608
              0.2905    0.6972    0.0001    0.0333

TIME.4      0.182236  0.018391 -0.277608  1.000000
              0.1671    0.8900    0.0333    0.0001
```

Here is the sphericity test. If the χ^2 is low and the *p* value is high, then the transformed variables are "spherical" or uncorrelated. If the χ^2 is high and the *p* value is low, then the transformed variables are not "spherical." They are uncorrelated.

```
Test for Sphericity: Mauchly's Criterion = 0.8310494
Chisquare Approximation = 10.440811 with 9 df   Prob > Chisquare = 0.3160
```

A sphericity test is also performed on the orthogonal components of the error covariance matrix. It is very important to pay attention to this test because it determines whether the statistical assumptions required for the univariate repeated measures are justified. (The univariate tests near the end of the output.) If this test is *not* significant (i.e., the χ^2 is low and the *p* value is high), then the assumptions are met. But if the test *is* significant (i.e., the χ^2 is high and the *p* value is low), then the assumptions have not been met. In this case, the univariate results at the end *may* be interpreted, but one should interpret the *adjusted* significance levels (i.e., the G-G *p* value and the H-Y *p* value). If the sphericity test fails badly (e.g., $p < .0001$), then even the adjusted significance levels are suspect.

When the transformation is mathematically orthogonal, then the two sphericity tests—that performed on the transformed variables and that done on the orthogonal components of the transformed variables give the same result. The polynomial transformation is mathematically orthogonal, so the two sphericity tests are the same in this example.

```
Applied to Orthogonal Components:
Test for Sphericity: Mauchly's Criterion = 0.8310494
Chisquare Approximation = 10.440811 with 9 df   Prob > Chisquare = 0.3160
```

There are two different tests for the within subjects (i.e., repeated measures) effects. The first of these is the straight forward MANOVA tests. The assumption about sphericity are not needed for the MANOVA results, so they can *always* be interpreted. The MANOVA for the TIME effect tests whether the means for the 5 time periods are the same. Wilk's l is very low, and its *p* value is highly significant, so we can conclude that the means for the language mastery test do indeed change over time.

```
Manova Test Criteria and Exact F Statistics for
the Hypothesis of no TIME Effect
H = Type III SS&CP Matrix for TIME   E = Error SS&CP Matrix

S=1    M=1    N=26.5

Statistic                     Value          F      Num DF    Den DF  Pr > F
Wilks' Lambda              0.07669737   165.5260         4        55  0.0001
Pillai's Trace             0.92330263   165.5260         4        55  0.0001
Hotelling-Lawley Trace    12.03825630   165.5260         4        55  0.0001
Roy's Greatest Root       12.03825630   165.5260         4        55  0.0001

---<PAGE>----------------------------------------------------------------------
Repeated Measures Example 1                                            11
General Linear Models Procedure
Repeated Measures Analysis of Variance
```

The TIME*GROUP effect tests whether the means for the two instructional groups are the same over time.

```
Manova Test Criteria and Exact F Statistics for
the Hypothesis of no TIME*GROUP Effect
H = Type III SS&CP Matrix for TIME*GROUP   E = Error SS&CP Matrix

S=1    M=1    N=26.5

Statistic                     Value          F      Num DF    Den DF  Pr > F
Wilks' Lambda              0.74044314     4.8200         4        55  0.0021
Pillai's Trace             0.25955686     4.8200         4        55  0.0021
Hotelling-Lawley Trace     0.35054260     4.8200         4        55  0.0021
Roy's Greatest Root        0.35054260     4.8200         4        55  0.0021
Repeated Measures Example 1                                            12
```

The test for between group effects asks whether the overall mean for the control group differs from the overall mean of the experimental group. Compare the *p* value for this test with the *p* value for the *t*-test done above for the variable TESTTOT.

```
General Linear Models Procedure
Repeated Measures Analysis of Variance
Tests of Hypotheses for Between Subjects Effects

Source                   DF    Type III SS   Mean Square   F Value     Pr > F
GROUP                     1       774.4133      774.4133      4.15     0.0462
Error                    58     10818.5733      186.5271

---<PAGE>----------------------------------------------------------------------
```

The following output is the heart and soul of repeated measures. It gives the univariate tests for the within subjects effects. For each within subject effect, there are three different *p* values. The first is the traditional *p* value that is valid when the assumptions of the repeated measures design have been met. The second and third are adjusted *p* values, the second (G – G) using the Greenhouse-Geisser adjustment and the third *H – H), the Huynh-Feldt adjustment. The adjustments are made to account for small discrepancies in the assumptions. Once again, these results may be untrustworthy when the test for the sphericity of the orthogonal components fails badly.

```
Repeated Measures Example 1                                           13
General Linear Models Procedure
Repeated Measures Analysis of Variance
Univariate Tests of Hypotheses for Within Subject Effects

Source: TIME
                                                          Adj  Pr > F
     DF     Type III SS     Mean Square   F Value   Pr > F    G - G    H - F
      4  19455.82000000   4863.95500000    154.92   0.0001   0.0001   0.0001

Source: TIME*GROUP
                                                          Adj  Pr > F
     DF     Type III SS     Mean Square   F Value   Pr > F    G - G    H - F
      4    674.02000000    168.50500000      5.37   0.0004   0.0005   0.0004

Source: Error(TIME)
     DF     Type III SS     Mean Square
    232   7284.16000000     31.39724138

Greenhouse-Geisser Epsilon = 0.9332
       Huynh-Feldt Epsilon = 1.0225

---<PAGE>----------------------------------------------------------------------
```

The final part of the output gives univariate ANOVAs for the transformed variables. SAS labels these as "Contrast Variables" in the output. Make certain that you do not confuse this with any contrast codes in the data. They are really the transformed variables. The first variable is TIME.1 and because a polynomial transformation was requested, it gives the linear effect of time. The row for MEAN whether the means for the five mastery exams change linearly over time. This effect is significant, so we know that there is an overall tendency for the overall mean to go up (or down) over time. The row for GROUP tests whether the linear increase (or decrease) in the control group differs from that in the experimental group. This is also significant, so the slope of the line over time differs in the two groups.

The next transformed variable is the quadratic effect of time. The effect for MEAN test whether the five means over time have a significant "bend" or "curve" to them. The result is significant, so there is an important curve to the plot of means over time. The degree of curvature however is the same for the two groups because the GROUP effect is not significant.

```
Repeated Measures Example 1                                          14
General Linear Models Procedure
Repeated Measures Analysis of Variance
Analysis of Variance of Contrast Variables

TIME.N represents the nth degree polynomial contrast for TIME

Contrast Variable: TIME.1

Source                  DF    Type III SS   Mean Square   F Value     Pr > F
MEAN                     1    18961.88167   18961.88167    638.12     0.0001
GROUP                    1      558.73500     558.73500     18.80     0.0001
Error                   58     1723.48333      29.71523

Contrast Variable: TIME.2
Source                  DF    Type III SS   Mean Square   F Value     Pr > F
MEAN                     1    421.4583333   421.4583333     11.17     0.0015
GROUP                    1     18.6011905    18.6011905      0.49     0.4853
Error                   58   2187.7261905    37.7194171
```

The next two variables are the cubic and the quartic effect over time testing for, respectively, two and three bends or curves in the plot of means by time. None of the effects here are significant.

```
Contrast Variable: TIME.3
Source                  DF    Type III SS   Mean Square   F Value     Pr > F
MEAN                     1    42.13500000   42.13500000      1.47     0.2296
GROUP                    1    22.81500000   22.81500000      0.80     0.3753
Error                   58   1657.9500000   28.58534483

Contrast Variable: TIME.4
Source                  DF    Type III SS   Mean Square   F Value     Pr > F
MEAN                     1    30.34500000   30.34500000      1.03     0.3152
GROUP                    1    73.86880952   73.86880952      2.50     0.1194
Error                   58   1715.0004762   29.56897373
```

Factorial Within Subjects Effects

The example given above had only a single within subjects factor, the time of the test. Other designs, however, may have more than a single within subjects effect. As an example, consider a study aimed at testing whether an experimental drug improves the memories of Alzheimer patients[2]. Thirty patients were randomly assigned to three treatment groups: (1) placebo or 0 milligrams (mg) of active drug, (2) 5 mg of drug, and (3) 10 mg. The memory task consisted of two modes of memory (Recall vs Recognition) on two types of things remembered (Names versus Objects) with two frequencies (Rare vs Common). That is, the patients were gives a list consisting of Rare Names (Waldo), Rare Objects (14th century map), Common Names (Bill), and Common Objects (fork) to memorize. The ordering of these four categories within a list was randomized from patient to patient. Half the patients within each drug group were asked to recall as many things on the list as they could. The other half were asked to recognize as many items as possible from a larger list. A new list was presented to the patients, and the opposite task was performed. That is, those who had the recall task were given the recognition task, and those initially given the recognition task were required to recall. Assume that previous research with this paradigm has shown that there are no order effects for the presentation. (This assumption is for simplicity only--it could easily be modeled using more advanced techniques than repeated measured.) The dependent variables are "memory scores." High scores indicate more items remembered. Table 3 gives the design.

Table 3. Design of a study evaluating the influence of an experimental drug on memory in patients with Alzheimer's disease.

		Recall				**Recognition**			
		Names		**Objects**		**Names**		**Objects**	
Obs.	**Drug**	**Rare**	**Comm**	**Rare**	**Comm**	**Rare**	**Comm**	**Rare**	**Comm**
1	1	63	70	73	76	68	81	77	90
2	1	72	74	68	70	76	87	82	94
.	.	.	.	.	.	.	.	.	.
30	3	80	85	78	87	88	92	94	103

There is one between subject's factor, the dose of drug. It has three levels—no active drug, 5 mg

[2] The data are fictitious. The program for analysis may be found in the file ~carey/p7291dir/alzheimr.sas

of drug, and 10 mg of drug. There are three within subjects factors. The first is memory mode with two levels, recall and recognition; the second is type of item recalled with the two levels of names and objects; and the third is the frequency of the item recalled with the two levels of rare or common.

In terms of ANOVA factors, a repeated measures design has the same logical structure as regular ANOVA. Hence, we can look on this design as being a 3 x 2 x 2 x 2 ANOVA. The ANOVA factors are drug, memory mode, type, and frequency. The ANOVA will fit main effects for all four of these factors, all two way interactions (e.g., drug*memory mode, drug*type, etc.), all of the three way interactions (e.g., drug*memory mode*type), and the four way interaction. Every ANOVA effect that contains at least one within subjects factor is considered a within subject's effect. Hence, the interaction of drug*type is a within subjects effect as well as the interaction of drug*type*frequency.

Transformations: Why Do Them?

The primary reason for transforming the dependent variables in RM is to generate variables that are more informative than the original variables. Consider the RM example. The dependent variables are tests of language mastery. Scores on a mastery test should be low early in the course and increase over the course. At some point, we might expect them to asymptote, depending upon the difficulty of the mastery test. We can then reformulate the research questions into the following set of questions: Do some types of instruction increase mastery at a faster rate than other types? Does computer based instruction increase mastery at a faster rate than classroom instruction? To answer these questions, we want to compare the independent variables on the *linear* increase over time in mastery test scores. If the mastery test is constructed so that scores asymptote at some time point during the course, we could ask the following questions: Do some types of instructions asymptote faster than other types? Or, does computer based instruction asymptote faster than classroom instruction? To answer these questions, we want to test the differences in the *quadratic* effect over time for the independent variables. In short, we want to transform the test scores so that the first transformed variable is the linear effect over time and the second variable is the quadratic effect over time. We can then do a MANOVA or RM on the transformed variables.

Transformations: How to Do Them

Both SAS and SPSS recognize that constructing transformation matrices is a real pain in

the gluteus to the max. Thus, they do it for you. As a user of a RM design, your major obligation is to choose the transformation *that makes the most sense for your data.* Be very careful, however, is using automatic transformations in statistical packages. There is no consensus terminology, so what is called a "contrast" transformation in one package may not be the same as a "contrast" transformation in another package. ***Always RTFM[3]!*** SAS will automatically perform the following transformations for you:

CONTRAST. A CONTRAST transformation compares each level of the repeated measures with the first level. It is useful when the first level represents a control or baseline level of response and you want to compare the subsequent levels to the baseline. For our example, the first transformed variable will be the (Test1 - Test2), the second transformed variable will be (Test1 - Test3), the third (Test1 - Test4) and the last (Test1 - Test5). CONTRAST is the default transformation in SAS--the one you get if you do not specify a transformation.

MEAN. A MEAN transformation compares a level with the mean of all the other levels. It is mostly useful if you haven't the vaguest idea of how to transform the repeated measures variables. For the example, the first transformed variable will be (Test1 - mean of [Test2 + Test3 + Test4 + Test5]), the second variable will be (Test2 - mean of [Test1 + Test3 + Test4 + Test5]), etc. Note that there is always one less transformation than the number of variables. Hence, if you use a MEAN transform in SAS, you will not get the last level contrasted with the mean of the other levels. If you have a burning passion to do this, see the PROC GLM documentation in the SAS manual.

PROFILE. A PROFILE transformation compares a level against the next level. It is sometimes useful in testing responses that are not expected to increase or decrease regularly over time. For the example, the first transformed variable is (Test1 - Test2), the second is (Test2 - Test3), the third is (Test3 - Test4), etc.

HELMERT. A HELMERT transform compares a level to the mean of all subsequent levels. This is a very useful transformation when one wants to pinpoint when a response changes over time. For our example, the first transformed variable would be (Test1 - mean of [Test2 + Test3 + Test4 + Test5]), the second would be (Test2 - mean of [Test3 + Test4 + Test5]), the third would be (Test3 - mean of [Test4 + Test5]), and the last would simply be (Test4 - Test5). If the

[3] For those who have not encountered the acronym RTFM, R stands for Read, T stands for The, and M stands for Manual.

univariate *F* statistics were significant for the first and second transformed variables but not significant for the third and fourth, then we would conclude that language mastery was achieved by the time of the third test.

POLYNOMIAL. A POLYNOMIAL transform fits orthogonal polynomials. Like a Helmert transform, this is useful to pinpoint changes in response over time. It is also useful when the repeated measures are ordered values of a quantity, say the dose of a drug. The first transformed variable represents the linear effect over time (or dose). The second transformed variable denotes the quadratic effect, the third the cubic effect, etc. If you are familiar enough with polynomials to interpret the observed means in light of linear, quadratic, cubic, etc. effects, this is an exceptionally useful transformation. Often, one can predict beforehand the order of the polynomial but not the exact time period where the response might be maximized (or minimized).

SPSS will also transform repeated measures variables using the CONTRAST subcommand to the MANOVA procedure. Note, however, that terminology and procedures differ greatly between SAS and SPSS. Be particularly careful of the CONTRAST statement. In SAS, the CONTRAST *statement* refers to a transformation of only the independent variables. The CONTRAST *option* on the REPEATED statement allows for a specific type of transformation for repeated measures variables. SPSSx views a CONTRAST as a transformation of either independent variables or dependent variables, depending upon the context. To make the issue more confusing, SPSSx has another subcommand, TRANSFORM, that applies only to the dependent variables. Also, both packages will do a "profile" transformation, but actually do different transformations. *You should always consult the appropriate manual before ever transforming the repeated measures variables.* ***Always RTFM!***

Quick and Dirty Approach to Repeated Measures

Background: There are probably as many different ways to perform repeated measures analysis as there are roads that lead to Rome. Furthermore, there are just as many differences in terminology. Here the term "repeated measures" is used synonymously with "within subjects." Thus, within subjects factors are the same as repeated measures factors. Also note that the SAS use of a "contrast" transformation for repeated measures is not the same as contrast coding as taught by Chick and Gary. Here, the term "transformation" is used to refer to the creation of new dependent variables from the old dependent variables. [Sorry about all this but I did not make up the rules.] The following is one quick and dirty way to perform a repeated measures ANOVA (or

regression). There are several other ways to accomplish the same task, so there is no "right" or "wrong" way as long as the correct model is entered and the correct statistics interpreted.

Setting up the data and the SAS commands

1. Make certain the data are entered so that each row of the data matrix is an independent observation. That is, if Abernathy is the first person, belongs to group 1, and has three scores over time (11, 12, and 13). Then enter

```
Abernathy    1    11    12    13
```

and not

```
Abernathy    1    1    11
Abernathy    1    2    12
Abernathy    1    3    13
```

It is possible to do a repeated measures analysis with the same person entered as many times as there are repeats of the measures, but that type of analysis will not be explicated here.

2. Use GLM and use the model statement as if you were doing a MANOVA. All repeated measures variables are the dependent variables. Suppose the three scores are called SCORE1, SCORE2, and SCORE3 in the SAS data set and GROUP is the independent variable. Then use

```
PROC GLM; CLASSES GROUP;
MODEL SCORE1 SCORE2 SCORE3 = GROUP;
```

3. Use the REPEATED statement to indicate that the dependent variables are repeated measures of the same construct or, if you prefer the other terminology, within subjects factors. A recommended statement is

```
REPEATED <name> <number of levels> <transformation> / PRINTM
    PRINTE SUMMARY;
```

where <*name*> is a name for the measures (or within subject factors), <*number of levels*> gives the number of levels for the factor, and <*transformation*> is the type of multivariate transformation. For our example,

REPEATED TIME 3 POLYNOMIAL / PRINTM PRINTE SUMMARY;

will work just fine.

When there is more than a single repeated measures factor, then you must specify them in the correct order. For example, suppose the design called for a comparison of recall versus recognition memory for phrases that are syntactically easy, moderate, and hard to remember. Each subject has 2 x 3 = 6 scores. Suppose Abernathy's scores are arranged in the following way:

		Recall			Recognition		
		Easy	Mod	Hard	Easy	Mod	Hard
	Group	Y1	Y2	Y3	Y4	Y5	Y6
Abernathy	1	12	8	3	21	16	14

The SAS statements should be:

```
PROC GLM; CLASSES GROUP;
  MODEL Y1-Y6 = GROUP;
  REPEATED MEMTYPE 2, DIFFCLTY 3 POLYNOMIAL / PRINTM PRINTE
SUMMARY;
```

There we specify two repeated measures factors (or within subjects factors). The first is MEMTYPE for recall versus recognition memory, and the second is DIFFCLTY and to denote the difficulty level of the phrases, . **Note that the factor that changes least rapidly always comes first**. Had we specified DIFFCLTY 3, MEMTYPE 2, then SAS would have interpreted Y1 as Recall-Easy, Y2 as Recognition-Easy, Y3 as Recall-Moderate, etc.

4. Remember that using a REPEATED statement will always generate a transformation of the variables. Always choose the type of transformation that will reveal the most meaningful information about your data.

5. That is all there is to doing a repeated measures ANOVA or Regression. You can use the CONTRAST statement if you wish to contrast code categorical independent variables. Just make certain that you place the CONTRAST statement before the REPEATED statement.

Interpreting the Output

This is a synopsis of the handout on *Repeated Measures*. You should follow these steps to interpret the output.

1. The first thing SAS writes in the output is the design for the repeated measures. Always check this to make certain that correctly specified the levels of the repeated measures. This is particularly important when there is more than a single within subjects factor.

2. The second thing to check is whether error covariance matrix can be orthogonally transformed. The tests of sphericity will tell you that. Some transformations in SAS are deliberately set up to be orthogonal (e.g., POLYNOMIAL with no further qualifiers); other transformations are not orthogonal (e.g., CONTRAST). If a transformation is orthogonal, then SAS will print out one test of sphericity. If a transformation is not orthogonal, then SAS spits out two tests of sphericity. The first test is for the straight transformation. The second test is for the orthogonal components of the transformation. In this case, it is the second test--the one for the orthogonal components--that you want to interpret.

3. If the χ^2 test for sphericity is not significant, then ignore all the MANOVA output and interpret the RM ANOVA results for the within subjects effects. These are labelled in the SAS output as "Univariate Tests of Hypotheses for Within Subjects Effects."

4. If the χ^2 test for sphericity was very significant, then you can interpret the MANOVA results or the adjusted probability levels from Greenhouse-Geisser and the Huynh-Feldt corrections for the within subjects effects. If is often a good idea to compare the MANOVA significance with the Greenhouse-Geisser and the Huynh-Feldt adjusted significance levels to make certain there is agreement between them.

5. The between subjects effects are not affected by the results of the sphericity test. Hence, SAS output with the heading "Tests of Hypotheses for Between Subjects Effects" will always be correct.

6. Always interpret the output for the transformed variables. It can often tell you something important about the data. Exactly what it tells you will depend upon the type of transformation you used in the REPEATED statement.

7. Always make certain that the raw means and standard deviations are printed. If you have not gotten them in the GLM procedure with the MEANS statement, then get them by using PROC MEANS, PROC UNIVARIATE, or PROC SUMMARY. Repeated measures or within subjects designs are useless when the results are not interpreted with respect to the raw data.

Appendix 1: Mathematical Assumptions of RM analysis

There are two major assumptions required for RM analysis. The first of these is that the within group covariance matrices are homogeneous. That means that the covariance matrix for group 1 is within sampling error of the covariance matrix of group 2 which is within sampling error of the covariance matrix of group 3, etc. for all groups in the analysis. Programs such as SPSS permit a direct test of homogeneity of covariance matrices. Much to the dismay of many SAS enthusiasts, testing for homogeneity of covariance must be done in a roundabout way. To perform such an analysis, create a new variable, say, *group*, that has a unique value for each group in the analysis. For example, suppose that you have a 2 (sex) by 3 (treatment) factorial design. The data are in a SAS data set called *wacko* where sex is coded as 1 = male, 2 = female and the 3 categories of treatment are numerically coded as 1, 2, and 3. Then a new variable called group may be generated using the following code

```
DATA wacko2;
   SET wacko;
   group = 10*sex + treatmnt;
RUN;
```

The second step is to perform a discriminant analysis using *group* as the classification variable, the RM variables as the discriminating variables, and the METHOD=NORMAL and POOL=TEST options in the PROC DISCRIM statement. If the RM variables in data set *wacko* were *vara*, *varb*, and *varc*, then the SAS code would be

```
PROC DISCRIM DATA=wacko2 METHOD=NORMAL POOL=TEST;
    CLASS group;
    VAR vara varb varc;
RUN;
```

Pay attention only to the results of the test for pooling the within group covariance matrices and ignore all the other output.

The second major assumption for RM is that the pooled within group covariance matrix has a certain mathematical form, called a type H matrix or, synonymously, a matrix meeting Huyhn-Feldt conditions. A covariance matrix, Σ, that is a type H matrix or, in other words, satisfies the Huynh-Feldt conditions is defined as a matrix which equals

$$\Sigma = \alpha\mathbf{I} + \beta\mathbf{1}^t + \mathbf{1}\beta^t$$

where α is a constant, $\mathbf{I}$ is an identity matrix, β is a vector, and $\mathbf{1}$ is a vector of 1s. For example., if α = 10 and β = (1, 2, 3), then

$$\Sigma = 10\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix} + \begin{pmatrix}1\\ 2\\ 3\end{pmatrix}\begin{pmatrix}1 & 1 & 1\end{pmatrix} + \begin{pmatrix}1\\ 1\\ 1\end{pmatrix}\begin{pmatrix}1 & 2 & 3\end{pmatrix} =$$

$$\begin{pmatrix}10 & 0 & 0\\ 0 & 10 & 0\\ 0 & 0 & 10\end{pmatrix} + \begin{pmatrix}1 & 1 & 1\\ 2 & 2 & 2\\ 3 & 3 & 3\end{pmatrix} + \begin{pmatrix}1 & 2 & 3\\ 1 & 2 & 3\\ 1 & 2 & 3\end{pmatrix} =$$

$$\begin{pmatrix}11 & 13 & 14\\ 13 & 14 & 15\\ 14 & 15 & 16\end{pmatrix}.$$

One classic type of matrix that satisfies this condition is a matrix where the all of the variances are the same and all of the covariances are the same. Such a matrix has the properties of *homogeneity of variance* and *homogeneity of covariance*. In this case all the elements of vector β are the same and equal .5*covariance and the constant α equals the variance minus the covariance. For example, if the variance were 8 and the covariance were 3, then α = 5 and β = (1.5, 1.5, 1.5). You should verify that

$$\begin{pmatrix}8 & 3 & 3\\ 3 & 8 & 3\\ 3 & 3 & 8\end{pmatrix} = 5\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix} + \begin{pmatrix}1.5\\ 1.5\\ 1.5\end{pmatrix}\begin{pmatrix}1 & 1 & 1\end{pmatrix} + \begin{pmatrix}1\\ 1\\ 1\end{pmatrix}\begin{pmatrix}1.5 & 1.5 & 1.5\end{pmatrix}.$$

A common mistake among many statistics texts is that RM assumes homogeneity of variance and homogeneity of covariance. If there is homogeneity of variance and homogeneity of covariance, then the RM assumptions are indeed met. But the converse of this statement is not true—the RM assumptions can in fact be met by matrices that conform to type H matrices (i.e., meet the Huynh-Feldt conditions) but do not have the joint properties of homogeneity of

variance and homogeneity of covariance.