DOCUMENT RESUME

ED 426 091 TM 029 320

AUTHOR Wells, Robert D.
TITLE Conducting Repeated Measures Analyses Using Regression: The General Linear Model Lives.
PUB DATE 1998-11-05
NOTE 33p.; Paper presented at the Annual Meeting of the Mid-South Educational Research Association (New Orleans, LA, November 4-6, 1998).
PUB TYPE Numerical/Quantitative Data (110) -- Reports - Evaluative (142) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Heuristics; *Multivariate Analysis; *Regression (Statistics); *Research Design; Tables (Data)
IDENTIFIERS *General Linear Model; *Repeated Measures Design; Univariate Analysis

ABSTRACT

The use of repeated measures research designs is explored. Repeated measures designs are often advantageous and can be implemented in a variety of research settings. One of the main advantages in repeated measures designs is the control of subject variability. Other advantages are the reduction of error variance and economy in subject recruitment. A disadvantage is the carry-over practice, effect of the repetition. A heuristic example is presented to illustrate the different statistical and conceptual properties of univariate and multivariate approaches when using repeated measures designs. It is demonstrated that repeated measures designs can be analyzed using analysis of variance, linear regression, and multivariate analysis of variance. In repeated measures designs, certain assumptions have to be met in order for the test statistics to be accurate. When assumptions are not met, a wide array of analytic choices is provided. The choice between univariate and multivariate analyses within repeated measures designs depends largely on the characteristics of the data. Appendixes present the Statistical Package for the Social Sciences (SPSS) program and SPSS data. (Contains 7 tables and 13 references.) (SLD)

Running head: REPEATED MEASURES

Conducting Repeated Measures Analyses Using Regression:

The General Linear Model Lives

Robert D. Wells

Texas A&M University 77843-4225

Paper presented at the annual meeting of the Mid-South Educational Research Association, New Orleans, November 5, 1998.

Abstract

The purpose of the present paper is to review the basics of repeated measures designs. It is demonstrated that repeated measures data can be analyzed using ANOVA, linear regression, and MANOVA. Discussion is provided detailing the assumptions that must be met in order for the test statistics to be accurate. A wide array of analytic choices is provided when the assumptions are not met. A small heuristic data set is utilized to make the discussion more concrete.

Experimental designs employ random assignment in an attempt to equate groups with respect to the dependent variable prior to the experiment. This equality prior to the experiment is desired so any posttest differences can be attributed solely to the independent variables (i.e., treatment effects) that were introduced (Girden, 1992). However, random assignment is not always possible or even the most desirable experimental design; repeated measures designs can prove more advantageous.

An example of the simplest repeated measures design is a pretest, posttest experimental design with intervening treatment (Minke, 1997). In this design a single group of subjects are measured under various treatments, also called conditions (Stevens, 1996). Repeated measures designs are also advantageous when the research is focusing on performance trends over time. For example, participants might be randomly assigned to two experimental groups, and then at the conclusion of the study all be posttested at three times on a monthly basis, so as to explore decay or enhancement of the treatment effects over time.

Repeated measures designs are implemented in a variety of research settings, thus researchers use many different designs, called by many different names (Minke, 1997). She provided a list of these different names, including a one-way repeated measures ANOVA, also called a one factor within-subjects ANOVA, a treatment-by-subjects ANOVA, or a randomized blocks ANOVA. This list includes a two-way repeated measures ANOVA, also called a two-way within-subjects ANOVA, a two-way ANOVA with repeated measures on both factors, a multiple treatments-by-subjects ANOVA, or treatments-by-treatments-by-subjects ANOVA (Huck & Cormier, 1996).

Repeated measures designs can include both between and within variables. Such designs are termed "mixed-model designs" (Barcikowski & Robey, 1984). Stevens (1996) wrote that a between variable is merely a grouping or classification variable (e.g., sex, age, social class) for which subjects will appear in only one level. For example, gender is inherently a between variable, since it is not reasonable to alter everyone's gender during the course of the experiment. A within variable is one in which the subjects will appear in each level (i.e., subjects participate in more than one experimental condition), although at varying times.

There are several advantages of implementing a repeated measures design. Lewis (1993) indicated that the main advantage is the control of subject variability (i.e., individual differences). He wrote that by selecting only one group of subjects and having that group participate in all the treatment conditions it would seem to guarantee that any observed differences among the treatment conditions will be due solely to the effects of the treatment. That is, in a between subjects design, even though random assignment is used, a highly unusual person or outlier may substantially influence the results in the group to which that person was assigned, notwithstanding the fact that assignment was random. Of course, the effects of such individual differences will be particularly noteworthy when group sizes are small.

However, Lewis (1993) noted that controlling an outlier's influence by assigning all subjects to all groups might not fully control individual difference, because in fact subjects do not remain perfectly constant. For example, a subject's interest level, attention, motivation, and knowledge of a task may change at differing levels of the treatment. However, these types of changes are probably not going to be as great as

differences produced by the random assignment of subjects to the different experimental conditions (Lewis, 1993). Therefore, if everything else remains constant, the error sum of squares, SS error, associated with the experimental conditions should be lower in a repeated measures design than in an experimental design in which randomly assigned groups of subjects are utilized.

This is associated with another advantage of repeated measures designs, the reduction of error variance. The lower sum of squares error represents an increase in statistical power and economy (Lewis, 1993). Stevens (1996) stated that because the subjects remain the same, the variance due to subjects can be partitioned out of the error variance term, thereby making any statistical tests more sensitive (powerful).

Repeated measures designs are also more economical in terms of subject recruitment. Because repeated measures designs employ the same subjects throughout different treatments, they require fewer subjects overall (Minke, 1997). Girden (1992) provided a simplistic example: An investigation of the effect of three different drugs (each a level) at two different doses (two levels) would require six different groups of individuals. Supposing a minimum of 10 individuals per group, a total of 60 individuals, need to be recruited. However, if a repeated measures design is utilized then the same subjects can take part in all conditions of the study, thereby requiring fewer participants. In this scenario, the same study might employ only 10 or 20 subjects each measured six times. These designs are particularly beneficial when the subject pool is limited.

However, there are also disadvantages associated with repeated measures designs. One such disadvantage is a carryover or practice effect, which Lewis (1993) stated concerns the fact that subjects may change systematically during the course of multiple

testing. He wrote that practice effects can be both positive (i.e., subjects show improvement) or negative (i.e., boredom or fatigue results due to successive testings). For example, in an experiment to teach 10 vocabulary words using two instructional methods, once subjects learn some or all of the words in their first experimental condition, they will carry this knowledge into their second experimental exposure, and this carryover in this example would distort posttest scores in the second condition by inflating estimated intervention effects. Girden (1992) also noted that latency effects might occur, which involve a delayed effect of treatment that is not evident until a second treatment is introduced. She stated that often these order effects can be controlled through counterbalancing; however, whereas increasing the time between treatments can minimize carryover effects, latency effects are more difficult to control.

Counterbalancing is a procedure of presenting the different levels of treatment so that each level a) occurs equally often at each stage of practice (i.e., each is presented first, then second, etc.) and b) precedes equally as many times as it follows each level (Girden, 1992). She provided the following strategy to achieve counterbalancing:

> With an even number of levels and a number of individuals that is some multiple of it, these two requirements can be met applying the following guideline: 1, 2, n, 3, n-1, 4, n-2, etc. Each number refers to a treatment level. If four levels are to be tested, then the guideline reduces to 1, 2, 4, 3, the first order of the levels. Each subsequent order is derived by adding "1" to the numbers of the preceding order. The second order would be 2 (1 + 1), 3 (2 + 1), 1 (4 + 1 does not apply), 4 (3 + 1).... Each fixed ratio occurs once at

each stage of practice, and each precedes and follows every other ratio the same number of times. (Girden, 1992, pp. 3-4)

An example of this procedure is provided in Table 1. Counterbalancing, while important and necessary in conducting repeated measures designs, does not remedy all the problems experienced in these analyses (Kieffer, 1998). Carryover effects, even when counterbalancing techniques are implemented, may raise issues involving external validity (Minke, 1997).

Insert Table 1 About Here

Univariate Approaches to Repeated Measures Analyses

A heuristic example is presented to illustrate the different statistical and conceptual properties of univariate and multivariate approaches when using repeated measures designs. Suppose a pharmaceutical company has developed a new supplemental vitamin that it believes increases restful sleep patterns for insomniacs. The vitamin was tested in each of four dosage levels (0 mg, 15 mg, 30 mg, and 50 mg) with each of four clinically diagnosed insomniacs. The differing levels of vitamin dosage were counterbalanced across the participants, and between each change in dosage a three-day no vitamin condition was imposed to allow any traces of the previous vitamin to exit the subject's system. After each vitamin dosage level had taken effect, the researchers collected data from the participants to determine how much REM sleep (i.e., the dependent variable) had increased in minutes. Once the final vitamin dosage treatment was concluded, a single within-factor (vitamin treatment) repeated measures analysis was conducted. The data for this hypothetical example are presented in Table 2.

Insert Table 2 About Here

Repeated Measures ANOVA

Table 3 indicates that the repeated measures ANOVA results in this hypothetical example are quite favorable. The various vitamin dosages (treatment conditions) accounted for 69.94% of the total variation on the increased REM sleep for the insomniacs. These results were statistically significant at the alpha = .05 level. Based on these results (large effect size and statistical significance), the pharmaceutical company can report that different vitamin dosages produced an increase in the REM sleep of clinically diagnosed insomniacs.

Insert Table 3 About Here

Table 4 indicates that the between-subjects ANOVA results in this example are not as favorable. In using a between-subjects ANOVA design with this data, each individual would be randomly assigned to only one treatment condition, as opposed to being measured on all four treatment conditions. Between-subjects ANOVA examines the effects between subjects, therefore, 16 subjects would be required to complete this same study (Keiffer, 1998). In the between-subjects ANOVA design it is only possible to study the differences between the vitamin dosages (levels of treatments) and not the differences between the subjects. This means that this ANOVA partitions the variance on the dependent variable into one less component than the repeated measures design.

Insert Table 4 About Here

As Tables 3 and 4 indicate, partitioning the variance into one more component in the repeated measures design proved beneficial in this example. The sum of squares (SS) residual was smaller in the repeated measures design than in the between-subjects ANOVA since there was one more partition. In this example 734.75 SS units were accounted for by the inclusion of this "person individual differences" (i.e., the differences across people as individuals) source of variation in the repeated measures analysis. Thus, utilizing the repeated measures design in this example proved beneficial to the outcome of the study in terms of statistical significance (Keiffer, 1988). While both results were statistically significant at the alpha = .05 level, the Fcalc value was larger in the repeated measures analysis. If the researcher were interested in statistically significant results, a larger Fcalc would prove beneficial.

While comparing the two analyses, notice that the variance accounted for by the treatment component (69.94%) in the present example remained identical. Keiffer (1998) stated that in a real-world setting with actual data, the distinctions between these two approaches would be more pronounced, and the power of employing a repeated measures design could be illustrated more convincingly. Keep in mind that similar results were observed in both analyses, even with the repeated measures design requiring only four subjects compared to the 16 required in the between-subjects ANOVA design.

Repeated Measures Analysis Using Regression

Another univariate approach that could be employed to examine the repeated measures data is multiple regression analysis. The first step in utilizing this approach is to define k contrast variables where k depicts the number of treatment levels. In this

example there are 4 treatment levels, so three mutually orthogonal (i.e., uncorrelated) contrast variables can be created. Each contrast variable illustrates a separate effect. Here the three contrast variables represent linear, quadratic, and cubic trends in the data, respectively; any orthogonal trends could be employed, but polynomial trends have been arbitrarily chosen in this example (Kieffer, 1998). In repeated measures regression a fourth orthogonal (i.e., uncorrelated) contrast variable is created as a sum vector that adds the responses of each subject over all four treatment conditions. Table 5 illustrates the resulting matrix. For example, person one had scores in the four dosages of 0, 18, 20, and 22, so that person's sum vector score was 60 (0 + 18 + 20 + 22).

Insert Table 5 About Here

The Pearson r_{XY} equals $COV_{XY} / (SD_X * SD_Y)$. The covariance, COV_{XY}, equals $(\Sigma (X_i - M_X) (Y_i - M_Y)) / n-1$; the numerator of this expression is called "the sum of the cross products (of the deviations from the mean)." If the cross products sum to zero, both the covariance and r will equal zero. Table 5 also presents two sets of cross products (CV1 with CV2, and CV3 with CV4), to illustrate that these contrast variables are uncorrelated (i.e., "orthogonal"), since the cross products both sum to zero. The same is true for all the pairwise combinations of the four contrast variables presented in Table 5.

After this matrix is generated, each contrast variable is entered into the multiple regression equation. The sequence in which the contrast variables are entered into the equation will not affect the variance accounted for by each contrast variable, since both the contrast variables and the variance they explain are uncorrelated. The results of the multiple regression repeated measures analysis are presented in Table 6. Table 7 shows

how the Table 6 results from the regression analysis are converted into a conventional repeated measures ANOVA format. Notice from Table 6 that the SS for the subtotal of the three contrasts testing linear, quadratic, and cubic trends in the dependent variable scores across the four dosages sum to the same value (SS = 2045.25) presented in Table 3; the cumulative eta^2 value (69.9%) is also the same across the two analytic approaches. Finally, note that the within effect due to variations in individual differences in drug reactions across the four people were the same (SS = 734.75, eta^2 = 13.7%) across these analyses.

Insert Tables 6 and 7 About Here

These comparisons illustrate that both ANOVA and regression analyses of the same data yield identical results. (However, the regression approach does facilitate further partitioning the dosage effects into three different trend effects.) The comparability of the results across the two analyses illustrates that all analyses are correlational, and that all analyses are part of a single general linear model family (cf. Cohen, 1968; Thompson; 1991, 1998).

Assumptions of Univariate Repeated Measures Analyses

Sphericity (or Circularity)

In conducting a repeated measures design, certain assumptions have to be met in order for the test statistics to be accurate. Sphericity (which subsumes the compound symmetry assumption as a special case) is deemed a necessary and sufficient condition for employing a repeated measures analysis (Huynh & Feldt, 1970). Sphericity is met if

all the differences between pairs of treatment condition scores are equally variable. If a variance-covariance matrix was constructed, it should demonstrate equal variances between these pairs of scores (Girden, 1992). Girden (1992) offered a formula to check for sphericity:

$$\sigma^2_{(y-y)} = \sigma^2_1 + \sigma^2_2 - 2\sigma_{1\,2}$$

where σ^2_1 is the variance of one set of scores, σ^2_2 is the variance of the paired set of scores, and σ_{12} is the covariance of the two sets of scores. She further stated that sphericity requires that variances of differences for all treatment combinations are homogenous. This indicates that the variance of level 1 and 2 has to equal the variance of level 2 and 3, and so forth.

Violations of the sphericity assumption

Violating the sphericity assumption can create incorrect statistical decisions. The resulting Fcalc will be too small (by some epsilon value) and the Fcrit value will be too high, when the variances of the differences in levels of the treatments are not equal. Stevens (1996) wrote that if sphericity is not examined, the unwary researcher will be inclined to reject the null hypothesis more often than it should be rejected which leads to higher Type I error rates.

However, there are methods created to correct for violations of the sphericity violation, which use a statistical correction factor called "epsilon." Epsilon is multiplied times the degrees of freedom to obtain corrected degrees of freedom. Box (1954) found the upper limit of epsilon to be 1.0, when the sphericity assumption is perfectly met in the data (i.e., no correction is invoked). When this occurs, the researcher may use the same degrees of freedom as calculated when constructing the summary table. However, the

sphericity assumption is not always met perfectly. (Geisser & Greenhouse, 1958) found the lower limit of epsilon is 1/ (k-1), where k is the number of treatment conditions in the study. Thus, the specific Geisser-Greenhouse correction that is computed by SPSS and other statistics packages ranges bewteen1/(k-1) and 1.0, but the exact value takes into account the degree to which the sphericity assumption is violated in a particular data set, and adjusts accordingly. However, the Greenhouse-Geisser epsilon produces a very conservative estimate for the degrees of freedom employed to test for statistical significance (i.e., the Fcrit value will tend to be too large which may result in the researcher not rejecting the null hypothesis as often as the data indicates).

Huynh and Feldt (1970) developed an epsilon correction index that is not as conservative. The Huynh-Feldt epsilon produces an overestimation of the true epsilon value. This may result in a Fcrit value that is too small, which may lead to Type I errors. Keiffer (1998) stated that since these two epsilon correction indices may produce biased results, several authors recommend averaging the two indices to produce a more accurate epsilon estimate (Barcikowski & Robey, 1983; Girden, 1992; Stevens, 1996). If the researcher is to choose one over the other, the Greenhouse-Geisser epsilon should be chosen because it is a more conservative estimate.

Multivariate Approach to Repeated Measures Analysis

Thus far the paper has discussed univariate designs to repeated measures analyses. Now a multivariate approach for a single factor repeated measures design will be considered as yet another analytic alternative. Multivariate Analysis of Variance (MANOVA) requires a conceptual shift. Lewis (1993) wrote that instead of thinking of the design (univariate) having subjects as a random factor crossed with the fixed repeated

measures factor, responses to different levels of the repeated measures factor are viewed as different dependent variables, and subjects are considered to provide replications in a single-cell design. He stated that unlike most multivariate approaches, for repeated measures all dependent variables are measured on the same scale, and the primary interest is in differences among the means of the different variables or levels. To achieve this a new set of dependent variables, based on the original ones, needs to be defined. For our heuristic data set, the multivariate approach would conceptually treat the outcome scores for the four drug dosages as separate dependent variables.

There are two approaches that can be utilized to define the new dependent variable set. Minke (1997) stated that the most common approach is to transform the k dependent variables into k-1 linearly independent pairwise difference scores and then perform the analysis on these new dependent variables. In this analysis the null hypothesis usually tested is that the difference scores have population means of zero. The analysis is then tested by transforming an F transformation of Hotelling's T^2 (Stevens, 1996). The second approach creates a matrix of orthonormal coefficients, weights each score by the corresponding coefficient, and then analyzes the new matrix. Keiffer (1998) stated that either of the two approaches will generate exactly the same results (Girden, 1992; Stevens, 1996) and there is no advantage to utilizing either approach.

However, as with many statistical designs, there are both advantages and disadvantages to utilizing a univariate repeated measures analysis within a multivariate framework. One advantage to the multivariate approach is the sphericity assumption does not have to be met (Keiffer, 1998). Keiffer argued that this may more closely honor a researcher's reality, provided that the researcher believes measurements are correlated in

a real world situation. Whereas in the multivariate design, sphericity is not an issue since each measurement is considered a separate and unique variable (Keiffer, 1998).

The present paper discussed the advantages and disadvantages of utilizing univariate and multivariate analyses within univariate repeated measures designs. When the sphericity assumption is met, the repeated measures ANOVA would prove the most effective design, because that analysis has the most statistical power against Type II error (Keiffer, 1998). However, when the sphericity assumption is severely violated a multivariate approach would be more beneficial. Therefore, the choice to use univariate or multivariate analyses within repeated measures designs depends largely on the characteristics of the data.

References

Barcikowski, R.S., & Robey, R.R. (1983). Decisions in single group repeated measures analysis: Statistical tests and three computer packages. American Statistician, 38, 148-150.

Box, G.E.P. (1954). Some theorems on quadratic forms in the study of analysis of variance problems: II. Effects of inequality of variance and of correlation between errors in the two-way classification. Annals of Mathematical Statistics, 25, 484-498.

Cohen, J. (1968). Multiple regression as a general data-analytic system. Psychological Bulletin, 70, 426-443.

Geisser, S., & Greenhouse, S.W. (1958). An extension of Box's results on the use of the F distribution in multivariate analysis. Annals of Mathematical Statistics, 29, 885-891.

Girden, E.R. (1992). ANOVA: Repeated measures. Newbury Park, CA: Sage.

Huck, S.W., & Cormier, W.H. (1996). Reading statistics and research (2nd ed.). New York: Harper Collins.

Huynh, H., & Feldt, L.S. (1970). Conditions under which mean square ratios in repeated measurements designs have exact F distributions. Journal of the American Statistical Association, 65, 1582-1589.

Keiffer, K.M. (1998, February). Repeated measures: Univariate tests, multivariate tests, or both? Paper presented at the third annual Education Research Exchange, College Station, TX.

Lewis, C. (1993). Analyzing means from repeated measures data. In G. Keren & C. Lewis (Eds.), A handbook for data analysis in the behavioral sciences: Statistical issues (pp. 73-94). Hillsdale, NJ: Erlbaum.

Minke, A. (1997, January). Conducting repeated measures analyses: Experimental design considerations. Paper presented at the annual meeting of the Southwest Educational Research Association, Austin, TX. (ERIC Document Reproduction Service No. ED 407 415)

Stevens, J. (1996). Applied multivariate statistics for the social sciences (3rd ed.). Mahwah, NJ: Lawrence Erlbaum.

Thompson, B. (1991). A primer on the logic and use of canonical correlation analysis. Measurement and Evaluation in Counseling and Development, 24, 80-95.

Thompson, B. (1998, April). Five methodology errors in educational research: The pantheon of statistical significance and other faux pas. Invited address presented at the annual meeting of the American Educational Research Association, San Diego. (ERIC Document Reproduction Service No. ED forthcoming)

Table 1

Counterbalancing Order for Design with Four Treatment Conditions

	Trial Number			
Person	One	Two	Three	Four
1	1	2	4	3
2	2	3	1	4
3	3	4	2	1
4	4	1	3	2

Table 2

Heuristic Data for Vitamin Dosage Study

	Treatment Level					
Subject	0mg	15mg	30mg	50mg	Sum	Mean (Yi)
1	0	18	20	22	60	15.0
2	3	15	24	26	68	17.0
3	7	25	38	40	110	27.5
4	3	36	40	45	125	31.0
ΣY	13	94	122	133	362	- - -
Mean (Y_j)	3.25	23.5	30.5	33.25	- - -	22.63

Table 3

Summary Table for Repeated Measures ANOVA with Table 2 Data

Source	SS	df	MS	Fcalc	eta^2
Subjects	734.75	n-1 = 3	244.92	10.36	13.70%
TIME	2204.25	k-1 = 3	734.75	31.08	69.94%
Res	212.75	(n-1)(k-1) = 9	23.64		
Total	3151.75	(n)(k)-1 = 15			

Note. n = number of subjects=4. k = number of treatment conditions. Here "subjects" represents the influence of the individual differences in the people as individuals.

Table 4

Summary Table for Classical ANOVA (Four Level One-way) with Table 2 Data

Source	SS	df	MS	Fcalc	eta^2
Between	2204.25	k-1=3	734.75	9.31	69.94%
Residual	947.50	(n-1)(k-1)=12	78.96		
Total	3151.75	n-1=15			

Note. **n=16.** This result is statistically significant, since Fcrit (3, 12) = 3.49 at the alpha = .05 level.

Table 5

Orthogonal Contrasts for Regression Analysis with Table 2 Data

	Contrasts						
Subject	CV1	CV2	CV3	CV4	DV	$CP_{1 \times 2}$	$CP_{3 \times 4}$
1	-3	1	-1	60	0	-3	-60
2	-3	1	-1	68	3	-3	-68
3	-3	1	-1	110	7	-3	-110
4	-3	1	-1	124	3	-3	-124
1	-1	-1	3	60	18	1	180
2	-1	-1	3	68	15	1	204
3	-1	-1	3	110	25	1	330
4	-1	-1	3	124	36	1	372
1	1	-1	-3	60	20	-1	-180
2	1	-1	-3	68	24	-1	-204
3	1	-1	-3	110	38	-1	-330
4	1	-1	-3	124	40	-1	-372
1	3	1	1	60	22	3	60
2	3	1	1	68	26	3	68
3	3	1	1	110	40	3	110
4	3	1	1	124	45	3	124
Sum	0	0	0	1148		0	0
Mean	0	0	0	90.5		0	0

Note. CV1 = contrast variable 1 (linear trend), CV2 = contrast variable 2 (quadratic trend), CV3 = contrast variable 3 (cubic trend), CV4 = sum vector, and DV = the dependent variable score.

Table 6

Results of Multiple Regression Analysis with Table 2 Data

Source	SS	df	MS	Fcalc	R^2
CV3	16.20	1	16.20	.072	.51%
Residual	3135.55	14	3135.55		
CV2	322.45	2	161.23	.741	10.23%
Residual	2829.30	13	217.64		
CV1	2204.25	3	734.75	9.306	69.94%
Residual	947.50	12	78.96		
SUM	2939.00	4	734.75	37.989	93.25%
Residual	212.75	11	19.34		

Note. In the regression analysis, the contrast variables were arbitrarily entered in The order CV3, CV2, CV1, and CV4 (i.e., the sum of given individuals' four DV Scores).

Table 7

Re-expression of the Table 6 Results in an ANOVA Format

Source	SS	df	MS	Fcalc	eta^2
Between					
Cubic					
DV trend (CV3)	16.20 – 0.00 = 16.20	1	16.20	0.69	0.3%
Quadratic					
DV trend (CV2)	322.45 – 16.20 = 306.25	1	306.25	12.96	5.7%
Linear					
DV trend (CV1)	2204.25 – 322.45 = 1881.80	1	1881.80	79.61	35.1%
(Subtotal	2204.25	3	734.75	31.08	41.2%)
Within					
Individual differences (CV4)	2939.00 – 2204.25 = 734.75	3	244.92	10.36	13.7%
Residual	212.75	9	23.64		
Total	5356.00	15			

Note. The Table 6 results present *cumulative* SS values associated with the use of one (CV3), two (CV3 and CV2), three (CV3, CV2, and CV1), and four (CV3, CV2, CV1, and CV4) contrast variables, respectively. The unique SS due to each contrast variable or hypothesis is computed by subtracting the cumulative value for a given prior cumulative total from the current cumulative total (e.g., 322.45 – 16.20 = 306.25).

Appendix A

SPSS Program

```
TITLE 'EPSY 690W repeated measures#1'.
COMMENT ************************************************.
COMMENT ROBERT D. WELLS
COMMENT 9/8/1997.
COMMENT ************************************************.
SET BLANKS=SYSMIS UNDEFINED=WARN printback=listing.
subtitle 'One-way Repeated ANOVA & Multivariate ##########'.
DATA LIST
  File='a:repeated.DAT' FIXED RECORDS=4 TABLE
  /1 id 1 dv1 9-10
  /2 dv2 9-10
  /3 dv3 9-10
  /4 dv4 9-10 .
list variables=all/cases=99/format=numbered .
manova dv1 to dv4/
  wsfactors=time(4)/
  contrast(time)=polynomial/
  rename=cons, linear, quad, cubic/
  print=transform param(estim) signif(averf hf gg)/
  wsdesign=time .


SUBTITLE 'ANOVA THRU REGRESSION USING ORTHO CONTRASTS *****'.
execute .
DATA LIST
  File='a:repeated.DAT' FIXED RECORDS=1 TABLE
  /1 id 1 time 3 sum 5-7 dv 9-10.
EXECUTE.
list variables=all/cases=99/format=numbered .
IF (TIME EQ 1) CV1=-3.
IF (TIME EQ 2) CV1=-1.
IF (TIME EQ 3) CV1=1.
IF (TIME EQ 4) CV1=3.
IF (TIME EQ 1) CV2=1.
IF (TIME EQ 2) CV2=-1.
IF (TIME EQ 3) CV2=-1.
IF (TIME EQ 4) CV2=1.
IF (TIME EQ 1) CV3=-1.
IF (TIME EQ 2) CV3=3.
IF (TIME EQ 3) CV3=-3.
IF (TIME EQ 4) CV3=1.
REGRESSION VARIABLES=DV CV1 TO CV3 sum/DESCRIPTIVES=ALL/
  CRITERIA=PIN(.95)POUT(.999)TOLERANCE(.00001)/DEPENDENT=DV/
  ENTER CV3/ENTER CV2/ENTER CV1/ENTER sum/.
```

Appendix B

SPSS Data

repeated.lst 9/8/98

```
-> TITLE 'EPSY 690W repeated measures#1'.
-> COMMENT ************************************************.
-> COMMENT ROBERT D. WELLS
-> COMMENT 9/8/1997.
-> COMMENT ************************************************.
-> SET BLANKS=SYSMIS UNDEFINED=WARN printback=listing.
-> subtitle 'One-way Repeated ANOVA & Multivariate ##########'.
-> DATA LIST
->   File='a:repeated.DAT' FIXED RECORDS=4 TABLE
->   /1 id 1 dv1 9-10
->   /2 dv2 9-10
->   /3 dv3 9-10
->   /4 dv4 9-10 .
-> list variables=all/cases=99/format=numbered .


      ID DV1 DV2 DV3 DV4
    1  1   0  18  20  22
    2  2   3  15  24  26
    3  3   7  25  38  40
    4  4   3  36  40  45


Number of cases read:  4    Number of cases listed:  4


-> manova dv1 to dv4/
->   wsfactors=time(4)/
->   contrast(time)=polynomial/
->   rename=cons, linear, quad, cubic/
->   print=transform param(estim) signif(averf hf gg)/
->   wsdesign=time .

 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
The default error term in MANOVA has been changed from WITHIN CELLS to
WITHIN+RESIDUAL.  Note that these are the same for all full factorial
designs.


* * * * * * A n a l y s i s   o f   V a r i a n c e * * * * * *


        4 cases accepted.
        0 cases rejected because of out-of-range factor values.
        0 cases rejected because of missing data.
        1 non-empty cell.

        1 design will be processed.

 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
```

```
* * * * * * Analysis of Variance -- design 1 * * * * * *

Orthonormalized Transformation Matrix (Transposed)

                CONS      LINEAR       QUAD      CUBIC

DV1             .500       -.671        .500       -.224
DV2             .500       -.224       -.500        .671
DV3             .500        .224       -.500       -.671
DV4             .500        .671        .500        .224

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -


* * * * * * Analysis of Variance -- design 1 * * * * * *

Order of Variables for Analysis

  Variates     Covariates

   CONS

   1 Dependent Variable
   0 Covariates

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
Note.. TRANSFORMED variables are in the variates column.
       These TRANSFORMED variables correspond to the
       Between-subject effects.

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -


* * * * * * Analysis of Variance -- design 1 * * * * * *

Tests of Between-Subjects Effects.

Tests of Significance for CONS using UNIQUE sums of squares
Source of Variation          SS      DF        MS         F  Sig of F

WITHIN+RESIDUAL          734.75       3    244.92
CONSTANT                8190.25       1   8190.25     33.44      .010

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
Estimates for CONS
--- Individual univariate .9500 confidence intervals

CONSTANT

 Parameter      Coeff.  Std. Err.    t-Value     Sig. t Lower -95% CL- Upper

       1   45.2500000    7.82491    5.78282     .01028   20.34765   70.15235

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
```

```
* * * * * * Analysis of Variance -- design 1 * * * * * *

Order of Variables for Analysis

  Variates     Covariates

   LINEAR
   QUAD
   CUBIC

   3 Dependent Variables
   0 Covariates

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
Note.. TRANSFORMED variables are in the variates column.
       These TRANSFORMED variables correspond to the
       'TIME' WITHIN-SUBJECT effect.
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

* * * * * * Analysis of Variance -- design 1 * * * * * *

Tests involving 'TIME' Within-Subject Effect.

Mauchly sphericity test, W =      .00687
Chi-square approx. =          8.57765 with 5 D. F.
Significance =                   .127

Greenhouse-Geisser Epsilon =      .50054
Huynh-Feldt Epsilon =             .89129
Lower-bound Epsilon =             .33333

AVERAGED Tests of Significance that follow multivariate tests are equivalent to
univariate or split-plot or mixed-model approach to repeated measures.
Epsilons may be used to adjust d.f. for the AVERAGED results.

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

* * * * * * Analysis of Variance -- design 1 * * * * * *

EFFECT .. TIME
Multivariate Tests of Significance (S = 1, M = 1/2, N = -1/2)

Test Name         Value    Exact F Hypoth. DF   Error DF  Sig. of F

Pillais          .94556    5.78972       3.00       1.00       .294
Hotellings     17.36915    5.78972       3.00       1.00       .294
Wilks            .05444    5.78972       3.00       1.00       .294
Roys             .94556
Note.. F statistics are exact.

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
```

```
* * * * * * Analysis of Variance -- design 1 * * * * * *

Tests involving 'TIME' Within-Subject Effect.

 AVERAGED Tests of Significance for DV using UNIQUE sums of squares
 Source of Variation          SS      DF        MS         F  Sig of F

 WITHIN+RESIDUAL          212.75       9     23.64
    (Greenhouse-Geisser)              4.50
    (Huynh-Feldt)                     8.02
    (Lower bound)                     3.00
 TIME                    2204.25       3    734.75     31.08      .000
    (Greenhouse-Geisser)              1.50               31.08      .003
    (Huynh-Feldt)                     2.67               31.08      .000
    (Lower bound)                     1.00               31.08      .011

 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
 Estimates for LINEAR
 --- Individual univariate .9500 confidence intervals

 TIME

  Parameter     Coeff.  Std. Err.    t-Value     Sig. t Lower -95%  CL- Upper

        1   21.6898594    3.23780    6.69895     .00679   11.38574   31.99398

 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
 Estimates for QUAD
 --- Individual univariate .9500 confidence intervals

 TIME

  Parameter     Coeff.  Std. Err.    t-Value     Sig. t Lower -95%  CL- Upper

        1   -8.7500000    1.88746   -4.63586     .01891  -14.75674   -2.74326

 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
 Estimates for CUBIC
 --- Individual univariate .9500 confidence intervals

 TIME

  Parameter     Coeff.  Std. Err.    t-Value     Sig. t Lower -95%  CL- Upper

        1   2.01246118    1.91920    1.04859     .37139   -4.09529    8.12022

 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -


->
->
-> SUBTITLE 'ANOVA THRU REGRESSION with ORTHO POLY CONTRASTS *****'.
-> execute .
-> DATA LIST
->   File='a:repeated.DAT' FIXED RECORDS=1 TABLE
->   /1 id 1 time 3 sum 5-7 dv 9-10.
-> EXECUTE.
-> list variables=all/cases=99/format=numbered .
```

	ID	TIME	SUM	DV
1	1	1	60	0
2	1	2	60	18
3	1	3	60	20
4	1	4	60	22
5	2	1	68	3
6	2	2	68	15
7	2	3	68	24
8	2	4	68	26
9	3	1	110	7
10	3	2	110	25
11	3	3	110	38
12	3	4	110	40
13	4	1	124	3
14	4	2	124	36
15	4	3	124	40
16	4	4	124	45

Number of cases read: 16 Number of cases listed: 16

```
-> IF (TIME EQ 1) CV1=-3.
-> IF (TIME EQ 2) CV1=-1.
-> IF (TIME EQ 3) CV1=1.
-> IF (TIME EQ 4) CV1=3.
-> IF (TIME EQ 1) CV2=1.
-> IF (TIME EQ 2) CV2=-1.
-> IF (TIME EQ 3) CV2=-1.
-> IF (TIME EQ 4) CV2=1.
-> IF (TIME EQ 1) CV3=-1.
-> IF (TIME EQ 2) CV3=3.
-> IF (TIME EQ 3) CV3=-3.
-> IF (TIME EQ 4) CV3=1.
-> REGRESSION VARIABLES=DV CV1 TO CV3 sum/DESCRIPTIVES=ALL/
->   CRITERIA=PIN(.95)POUT(.999)TOLERANCE(.00001)/DEPENDENT=DV/
->   ENTER CV3/ENTER CV2/ENTER CV1/ENTER sum/.
```

* * * * M U L T I P L E R E G R E S S I O N * * * *

Listwise Deletion of Missing Data

	Mean	Std Dev	Variance	Label
DV	22.625	14.495	210.117	
CV1	.000	2.309	5.333	
CV2	.000	1.033	1.067	
CV3	.000	2.309	5.333	
SUM	90.500	27.995	783.733	

N of Cases = 16

Correlation, Covariance, 1-tailed Sig, Cross-Product:

	DV	CV1	CV2	CV3	SUM
DV	1.000	.773	-.312	.072	.483
	210.117	25.867	-4.667	2.400	195.933
	.	.000	.120	.396	.029
	3151.750	388.000	-70.000	36.000	2939.000
CV1	.773	1.000	.000	.000	.000
	25.867	5.333	.000	.000	.000
	.000	.	.500	.500	.500
	388.000	80.000	.000	.000	.000
CV2	-.312	.000	1.000	.000	.000
	-4.667	.000	1.067	.000	.000
	.120	.500	.	.500	.500
	-70.000	.000	16.000	.000	.000
CV3	.072	.000	.000	1.000	.000
	2.400	.000	.000	5.333	.000
	.396	.500	.500	.	.500
	36.000	.000	.000	80.000	.000
SUM	.483	.000	.000	.000	1.000
	195.933	.000	.000	.000	783.733
	.029	.500	.500	.500	.
	2939.000	.000	.000	.000	11756.000

* * * * M U L T I P L E R E G R E S S I O N * * * *

Equation Number 1 Dependent Variable.. DV

Descriptive Statistics are printed on Page 34

Block Number 1. Method: Enter CV3

Variable(s) Entered on Step Number 1.. CV3

Multiple R	.07169	Analysis of Variance			
R Square	.00514		DF	Sum of Squares	Mean Square
Adjusted R Square	-.06592	Regression	1	16.20000	16.20000
Standard Error	14.96556	Residual	14	3135.55000	223.96786

F = .07233 Signif F = .7919

---------- Variables in the Equation ----------

Variable	B	SE B	Beta	T	Sig T
CV3	.450000	1.673200	.071694	.269	.7919
(Constant)	22.625000	3.741389		6.047	.0000

---------- Variables not in the Equation ----------

Variable	Beta In	Partial	Min Toler	T	Sig T
CV1	.772700	.774693	1.000000	4.417	.0007
CV2	-.311718	-.312522	1.000000	-1.186	.2568
SUM	.482830	.484075	1.000000	1.995	.0675

End Block Number 1 All requested variables entered.

* * * * M U L T I P L E R E G R E S S I O N * * * *

Equation Number 1 Dependent Variable.. DV

Block Number 2. Method: Enter CV2

Variable(s) Entered on Step Number 2.. CV2

Multiple R	.31986	Analysis of Variance			
R Square	.10231		DF	Sum of Squares	Mean Square
Adjusted R Square	-.03580	Regression	2	322.45000	161.22500
Standard Error	14.75257	Residual	13	2829.30000	217.63846

F = .74079 Signif F = .4958

---------- Variables in the Equation ----------

Variable	B	SE B	Beta	T	Sig T
CV2	-4.375000	3.688144	-.311718	-1.186	.2568
CV3	.450000	1.649388	.071694	.273	.7893
(Constant)	22.625000	3.688144		6.135	.0000

---------- Variables not in the Equation ----------

Variable	Beta In	Partial	Min Toler	T	Sig T
CV1	.772700	.815544	1.000000	4.882	.0004
SUM	.482830	.509601	1.000000	2.052	.0627

End Block Number 2 All requested variables entered.

*** * * * M U L T I P L E R E G R E S S I O N * * * ***

Equation Number 1 Dependent Variable.. DV

Block Number 3. Method: Enter CV1

Variable(s) Entered on Step Number 3.. CV1

Multiple R	.83629	Analysis of Variance			
R Square	.69937		DF	Sum of Squares	Mean Square
Adjusted R Square	.62422	Regression	3	2204.25000	734.75000
Standard Error	8.88585	Residual	12	947.50000	78.95833
		F = 9.30554	Signif F = .0019		

---------- Variables in the Equation ----------

Variable	B	SE B	Beta	T	Sig T
CV1	4.850000	.993468	.772700	4.882	.0004
CV2	-4.375000	2.221463	-.311718	-1.969	.0724
CV3	.450000	.993468	.071694	.453	.6587
(Constant)	22.625000	2.221463		10.185	.0000

---------- Variables not in the Equation ----------

Variable	Beta In	Partial	Min Toler	T	Sig T
SUM	.482830	.880603	1.000000	6.164	.0001

End Block Number 3 All requested variables entered.

*** * * * M U L T I P L E R E G R E S S I O N * * * ***

Equation Number 1 Dependent Variable.. DV

Block Number 4. Method: Enter SUM

Variable(s) Entered on Step Number 4.. SUM

Multiple R	.96566	Analysis of Variance			
R Square	.93250		DF	Sum of Squares	Mean Square
Adjusted R Square	.90795	Regression	4	2939.00000	734.75000
Standard Error	4.39783	Residual	11	212.75000	19.34091
		F = 37.98942	Signif F = .0000		

---------- Variables in the Equation ----------

Variable	B	SE B	Beta	T	Sig T
CV1	4.850000	.491692	.772700	9.864	.0000
CV2	-4.375000	1.099458	-.311718	-3.979	.0022
CV3	.450000	.491692	.071694	.915	.3797
SUM	.250000	.040561	.482830	6.164	.0001
(Constant)	.000000	3.831888		.000	1.0000

End Block Number 4 All requested variables entered.

TM029320

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement (OERI)
Educational Resources Information Center (ERIC)

REPRODUCTION RELEASE
(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: CONDUCTING REPEATED MEASURES ANALYSES USING REGRESSION: THE GENERAL LINEAR MODEL	
Author(s): ROBERT D. WELLS	
Corporate Source:	Publication Date: 11/5/98

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic/optical media, and sold through the ERIC Document Reproduction Service (EDRS) or other ERIC vendors. Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce the identified document, please CHECK ONE of the following options and sign the release below.

☒ ← Sample sticker to be affixed to document

Check here
Permitting microfiche (4"x 6" film), paper copy, electronic, and optical media reproduction

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

ROBERT D. WELLS

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Level 1

Sample sticker to be affixed to document →

"PERMISSION TO REPRODUCE THIS MATERIAL IN OTHER THAN PAPER COPY HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Level 2

☐ **or here**
Permitting reproduction in other than paper copy.

Sign Here, Please

Documents will be processed as indicated provided reproduction quality permits. If permission to reproduce is granted, but neither box is checked, documents will be processed at Level 1.

"I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce this document as indicated above. Reproduction from the ERIC microfiche or electronic/optical media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries."

Signature: Robert D. Wells	Position: RES ASSOCIATE
Printed Name: ROBERT D. WELLS	Organization: TEXAS A&M UNIVERSITY
Address: TAMU DEPT EDUC PSYC COLLEGE STATION, TX 77843-4225	Telephone Number: (409) 845-1831
	Date: 10/25/98