# The Single-Factor Within-Subject Design

ERSH 8310

Keppel and Wickens Chapter 16

# Today's Class

- The simplest within-subject design with a single within-subject factor is discussed, including:
  - The basic analysis of variance
  - Tests of contrasts
  - Effect-size calculations
  - Power determinations

# The Analysis of Variance – Design and Notation

- Table 16.1 shows the basic structure of the scores in a single-factor within-subject design

- The scores in the data table are denoted by $Y_{ij}$, where the first subscript i indicates the subject, and the second subscript j indicates the particular level of factor A

- There are a levels of factor A and n subjects

- The subjects are nested within the groups, and this fact is symbolized by denoting the subject factor S/A

- It may be convenient and accurate to refer to this design as an A×S design

# The Analysis of Variance – Design and Notation

*Table 16.1: The notation system for an $A \times S$ design with $a = 4$ conditions and $n = 3$ subjects.*

| | Levels of factor $A$ | | | | |
|---|---|---|---|---|---|
| Subjects | $a_1$ | $a_2$ | $a_3$ | $a_4$ | Sum |
| $s_1$ | $Y_{11}$ | $Y_{12}$ | $Y_{13}$ | $Y_{14}$ | $S_1$ |
| $s_2$ | $Y_{21}$ | $Y_{22}$ | $Y_{23}$ | $Y_{24}$ | $S_2$ |
| $s_3$ | $Y_{31}$ | $Y_{32}$ | $Y_{33}$ | $Y_{34}$ | $S_3$ |
| Sum | $A_1$ | $A_2$ | $A_3$ | $A_4$ | $T$ |

# Formulae

- From Table 16.1 we have…
- The mean for each factor level: $\bar{Y}_j = \frac{A_j}{n}$
- The mean for each subject: $\bar{Y}_{S_i} = \frac{S_i}{a}$
- And the grand mean: $\bar{Y}_T = \frac{T}{an}$

# Partitioning the Variability

- The amount of variability among the scores in a one-way between-subject design is measured by the total variability, as expressed by the total sum of squares:

$$SS_T = \sum_{ij} \left( Y_{ij} - \bar{Y}_T \right)^2$$

- Just as the between-subjects variability in the two-factor A×B design with crossed factors is divided into two effects and an interaction, the total variability in an A×S design is divided into $SS_A$, $SS_S$, and $SS_{A \times S}$

- There is no within-cell term comparable to $SS_{S/AB}$ and $SS_{A \times S}$ will be treated as error
  - Only have one observation per subject per condition

# Computational Formulas

- The computational formulas for the A ×S design are presented in Table 16.2
  - Note that $df_T$ = an - 1 that is one less than the total number of observations

**Table 16.2: Computational formulas for the $A \times S$ design.**

Formulas for the analysis of variance

| Source | $df$ | $SS$ | $MS$ | $F$ |
|---|---|---|---|---|
| $A$ | $a-1$ | $[A]-[T]$ | $\frac{SS_A}{df_A}$ | $\frac{MS_A}{MS_{A\times S}}$ |
| $S$ | $n-1$ | $[S]-[T]$ | $\frac{SS_S}{df_S}$ | |
| $A\times S$ | $(a-1)(n-1)$ | $[Y]-[A]-[S]+[T]$ | $\frac{SS_{A\times S}}{df_{A\times S}}$ | |
| Total | $an-1$ | $[Y]-[T]$ | | |

Formulas for the bracket terms

$$[A]=\frac{\sum A_j^2}{n}=n\sum \bar{Y}_A^2 \qquad [S]=\frac{\sum S_i^2}{a}=a\sum \bar{Y}_S^2$$

$$[Y]=\sum Y_{ij}^2 \qquad [T]=\frac{T^2}{an}=an\bar{Y}_T^2$$

# NUMERICAL EXAMPLE

# Searching for Letters

- Consider an experiment in which college students search for a particular letter in a string of letters on a computer screen
  - Half of the time the letter occurs in the string, and half of the time it does not
- On one third of the trials the letter string is a word (condition $a_1$), on one third it is a pronounceable nonword ($a_2$), and on one third it is an unpronounceable set of random letters ($a_3$)
- The response measure is the average speed with which subjects correctly detect the target letter, measured in milliseconds
- The experiment is a single-factor AxS design with a=3 types of letter strings

# The Old Style of Analysis

- Prior to running the repeated measures analysis, let’s imagine we tried using what we already know about ANOVA:
  - Let’s run the analysis as if we have subjects nested within each factor (not crossed)

# Old Analysis – No Repeated Measures

**Between-Subjects Factors**

| | | N |
|---|---|---|
| condition | 1.00 | 6 |
| | 2.00 | 6 |
| | 3.00 | 6 |

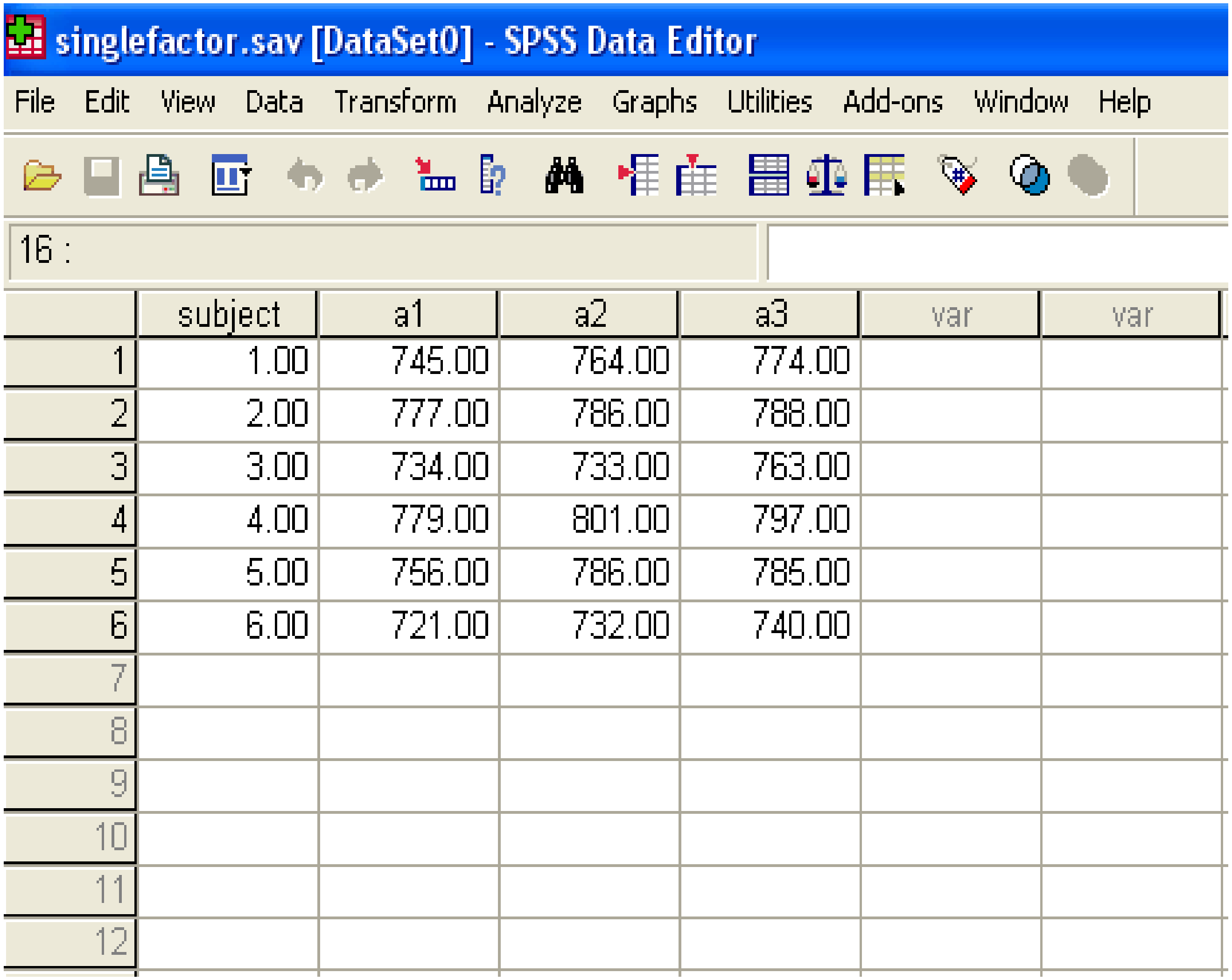
**Tests of Between-Subjects Effects**

Dependent Variable: time

| Source | Type III Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|
| Corrected Model | 1575.000[a] | 2 | 787.500 | 1.299 | .302 |
| Intercept | 10520284.5 | 1 | 10520284.50 | 17353.524 | .000 |
| condition | 1575.000 | 2 | 787.500 | 1.299 | .302 |
| Error | 9093.500 | 15 | 606.233 | | |
| Total | 10530953.0 | 18 | | | |
| Corrected Total | 10668.500 | 17 | | | |

a. R Squared = .148 (Adjusted R Squared = .034)

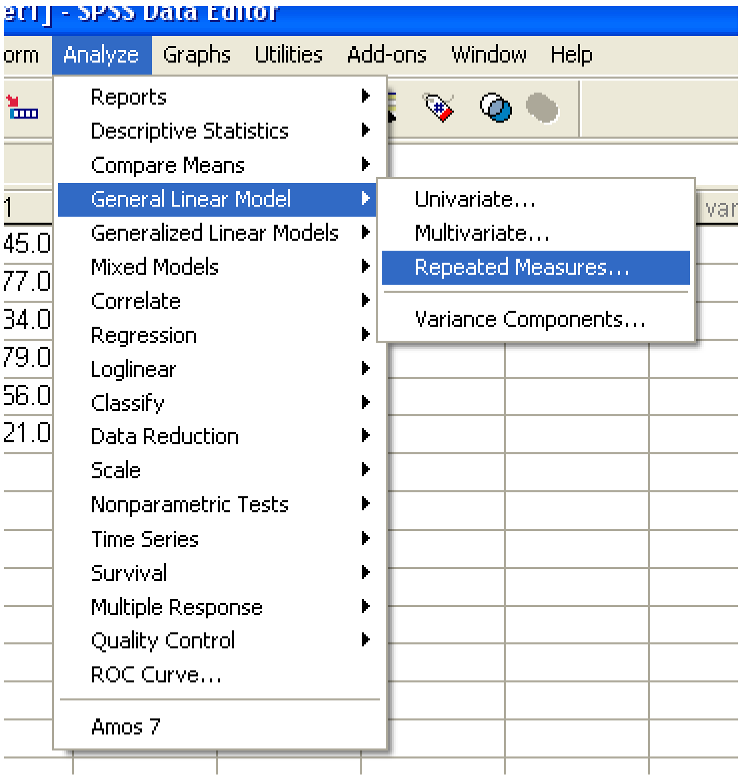
# Repeated Measures: Data File in SPSS

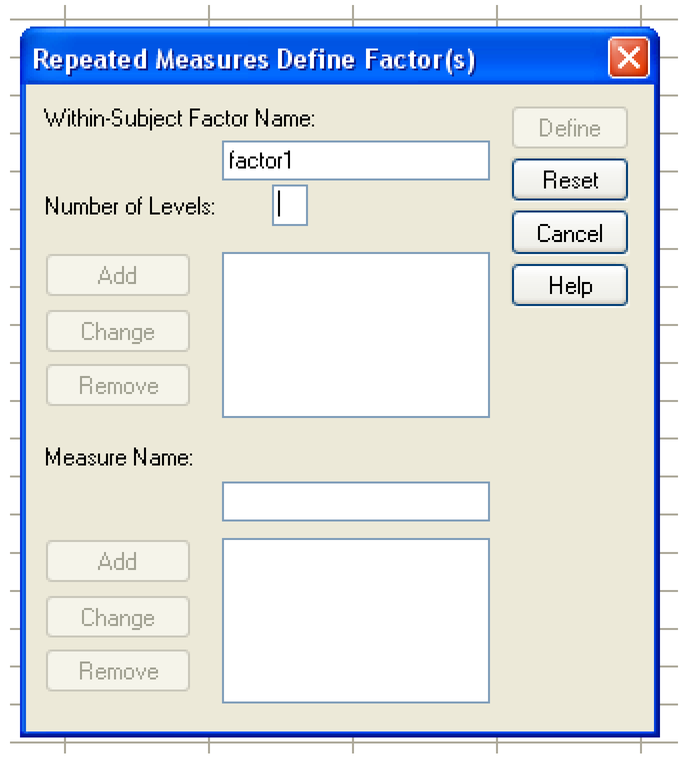



| | subject | a1 | a2 | a3 | var | var |
|---|---|---|---|---|---|---|
| 1 | 1.00 | 745.00 | 764.00 | 774.00 | | |
| 2 | 2.00 | 777.00 | 786.00 | 788.00 | | |
| 3 | 3.00 | 734.00 | 733.00 | 763.00 | | |
| 4 | 4.00 | 779.00 | 801.00 | 797.00 | | |
| 5 | 5.00 | 756.00 | 786.00 | 785.00 | | |
| 6 | 6.00 | 721.00 | 732.00 | 740.00 | | |
| 7 | | | | | | |
| 8 | | | | | | |
| 9 | | | | | | |
| 10 | | | | | | |
| 11 | | | | | | |
| 12 | | | | | | |

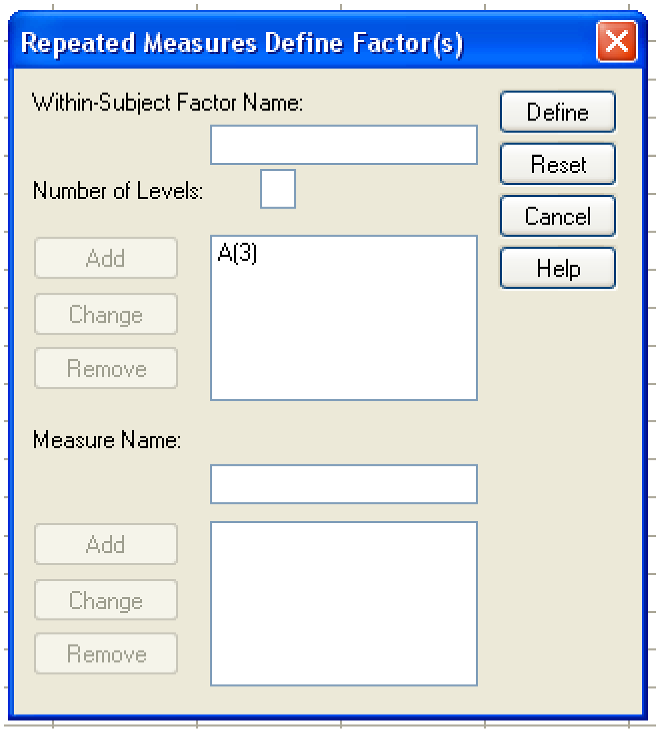
The data are arranged in a wide-format:
- One variable per column.

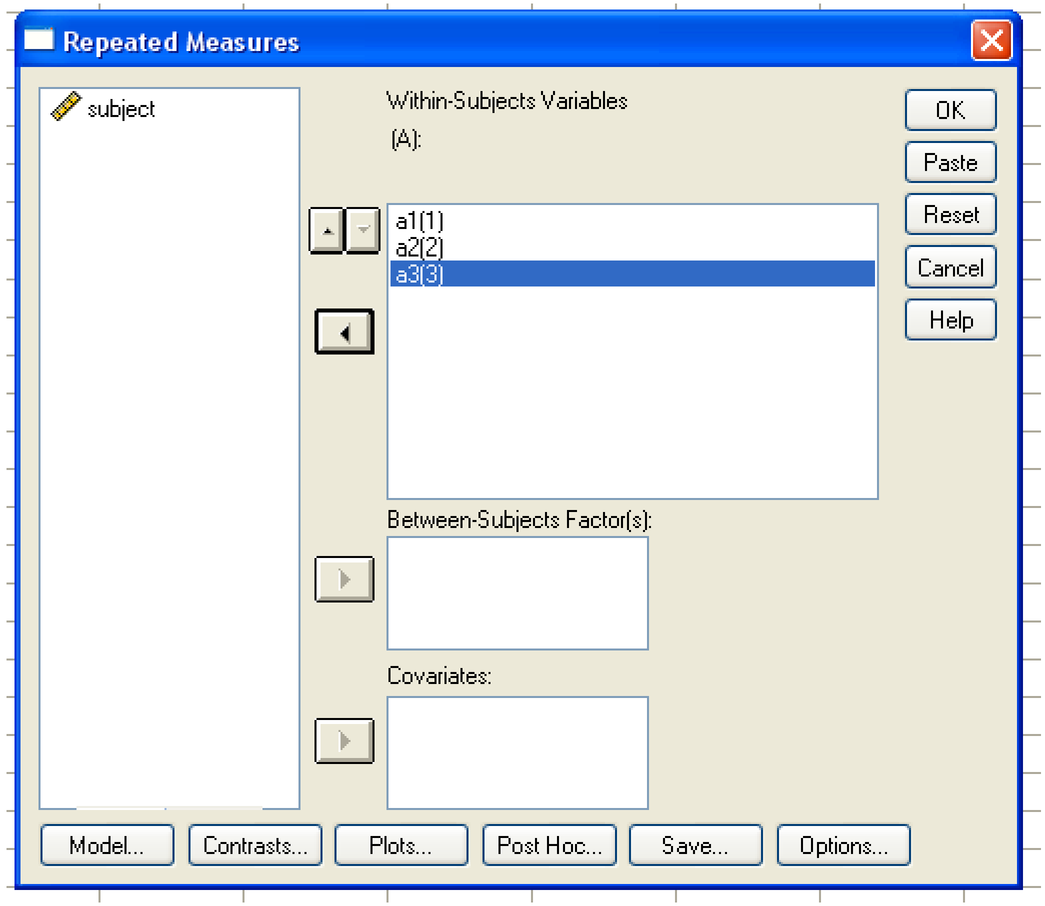
# Running the Analysis



# Running the Analysis



# Analysis Output: SPSS

- There are two relevant parts to SPSS- a test for the factor

**Tests of Within-Subjects Effects**

Measure:MEASURE_1

| Source | | Type III Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|---|
| factor1 | Sphericity Assumed | 1575.000 | 2 | 787.500 | 14.432 | .001 |
| | Greenhouse-Geisser | 1575.000 | 1.610 | 978.061 | 14.432 | .003 |
| | Huynh-Feldt | 1575.000 | 2.000 | 787.500 | 14.432 | .001 |
| | Lower-bound | 1575.000 | 1.000 | 1575.000 | 14.432 | .013 |
| Error(factor1) | Sphericity Assumed | 545.667 | 10 | 54.567 | | |
| | Greenhouse-Geisser | 545.667 | 8.052 | 67.771 | | |
| | Huynh-Feldt | 545.667 | 10.000 | 54.567 | | |
| | Lower-bound | 545.667 | 5.000 | 109.133 | | |

- And:

**Tests of Between-Subjects Effects**

Measure:MEASURE_1
Transformed Variable:Average

| Source | Type III Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|
| Intercept | 1.052E7 | 1 | 1.052E7 | 6153.773 | .000 |
| Error | 8547.833 | 5 | 1709.567 | | |

# Analysis Output: Table 16.3

The $AS$ data table

| Subjects | $a_1$ | $a_2$ | $a_3$ | Sum |
|---|---|---|---|---|
| | Types of strings | | | |
| $s_1$ | 745 | 764 | 774 | 2,283 |
| $s_2$ | 777 | 786 | 788 | 2,351 |
| $s_3$ | 734 | 733 | 763 | 2,230 |
| $s_4$ | 779 | 801 | 797 | 2,377 |
| $s_5$ | 756 | 786 | 785 | 2,327 |
| $s_6$ | 721 | 732 | 740 | 2,193 |
| Sum | 4,512 | 4,602 | 4,647 | 13,761 |
| $\sum_i Y_{ij}^2$ | 3,395,728 | 3,534,002 | 3,601,223 | |
| Mean | 752.00 | 767.00 | 774.50 | |
| $s_j$ | 23.26 | 29.22 | 20.60 | |
| $s_{M_j}$ | 9.50 | 11.93 | 8.41 | |

Calculation of the bracket terms

$$[A] = \frac{\sum A_j^2}{n} = \frac{4{,}512^2 + 4{,}602^2 + 4{,}647^2}{6} = 10{,}521{,}859.5$$

$$[S] = \frac{\sum S_i^2}{a} = \frac{2{,}283^2 + \cdots + 2{,}193^2}{3} = 10{,}528{,}832.3$$

$$[Y] = \sum Y_{ij}^2 = 745^2 + 764^2 + \cdots + 732^2 + 740^2 = 10{,}530{,}953.0$$

$$[T] = \frac{T^2}{an} = \frac{13{,}761^2}{(3)(6)} = 10{,}520{,}284.5$$

Summary of the analysis of variance

| Source | $SS$ | | $df$ | $MS$ | $F$ |
|---|---|---|---|---|---|
| $A$ | $[A] - [T] =$ | 1,575.0 | 2 | 787.50 | 14.43* |
| $S$ | $[S] - [T] =$ | 8,547.8 | 5 | 1,709.56 | |
| $A \times S$ | $[Y] - [A] - [S] + [T] =$ | 545.7 | 10 | 54.57 | |
| Total | $[Y] - [T] =$ | 10,668.5 | 17 | | |

* $p < .05$

# Matching Output to Interpretation

**Tests of Within-Subjects Effects**

Measure:MEASURE_1

| Source | | Type III Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|---|
| factor1 | Sphericity Assumed | 1575.000 | 2 | 787.500 | 14.432 | .001 |
| | Greenhouse-Geisser | 1575.000 | 1.610 | 978.061 | 14.432 | .003 |
| | Huynh-Feldt | 1575.000 | 2.000 | 787.500 | 14.432 | .001 |
| | Lower-bound | 1575.000 | 1.000 | 1575.000 | 14.432 | .013 |
| Error(factor1) | Sphericity Assumed | 545.667 | 10 | 54.567 | | |
| | Greenhouse-Geisser | 545.667 | 8.052 | 67.771 | | |
| | Huynh-Feldt | 545.667 | 10.000 | 54.567 | | |
| | Lower-bound | 545.667 | 5.000 | 109.133 | | |

Source: A

Source: SxA

**Tests of Between-Subjects Effects**

Measure:MEASURE_1
Transformed Variable:Average

| Source | Type III Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|
| Intercept | 1.052E7 | 1 | 1.052E7 | 6153.773 | .000 |
| Error | 8547.833 | 5 | 1709.567 | | |

Source: S

# EFFECT SIZE AND POWER

# Estimating Treatment Effects

- The estimate of the partial omega squared is:

$$\hat{\omega}^2_{<A>} = \frac{(a-1)(F_A - 1)}{(a-1)(F_A - 1) + an}$$

- Or, perhaps more readily obtainable is the partial squared correlation ratio ($\eta^2$ or $R^2$) is obtained by:

$$\hat{\eta}^2_{<A>} = \frac{SS_A}{SS_A + SS_{AxS}}$$

# In Our Data…

**Tests of Within-Subjects Effects**

Measure:MEASURE_1

| Source | | Type III Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|---|
| factor1 | Sphericity Assumed | 1575.000 | 2 | 787.500 | 14.432 | .001 |
| | Greenhouse-Geisser | 1575.000 | 1.610 | 978.061 | 14.432 | .003 |
| | Huynh-Feldt | 1575.000 | 2.000 | 787.500 | 14.432 | .001 |
| | Lower-bound | 1575.000 | 1.000 | 1575.000 | 14.432 | .013 |
| Error(factor1) | Sphericity Assumed | 545.667 | 10 | 54.567 | | |
| | Greenhouse-Geisser | 545.667 | 8.052 | 67.771 | | |
| | Huynh-Feldt | 545.667 | 10.000 | 54.567 | | |
| | Lower-bound | 545.667 | 5.000 | 109.133 | | |

Source: A

Source: SxA

$$\hat{\omega}^2_{<A>} = \frac{(a-1)(F_A-1)}{(a-1)(F_A-1)+an} = \frac{(3-1)(14.332-1)}{(3-1)(14.332-1)+3*6} = .597$$

**Tests of Between-Subjects Effects**

Measure:MEASURE_1
Transformed Variable:Average

| Source | Type III Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|
| Intercept | 1.052E7 | 1 | 1.052E7 | 6153.773 | .000 |
| Error | 8547.833 | 5 | 1709.567 | | |

Source: S

$$\hat{\eta}^2_{<A>} = \frac{SS_A}{SS_A + SS_{AxS}} = \frac{1575}{1575+545.667} = .743$$

# Power and Sample Size

- Power and sample-size calculations in the within-subject design are based on the measures of effect size
- The example data (and the summary of the analysis of variance) contained in Table 16.7 can be used to illustrate the calculation of a required sample size to achieve a given power

# ADVANTAGES AND LIMITATIONS OF RM ANOVA

# Advantages and Limitations

- A study conducted with a within-subject design obtains more data from each subject than one conducted with a between-subjects design, and the analysis has a smaller error term

- Repeated observations of a subject, however, cannot be collected under constant conditions, and any earlier observation has the potential to influence later ones

- The assumptions that underlie the analysis are more complex than those of the between-subjects designs

# Advantages of the Within-Subject Design

- The three principal advantages of a within-subject design are:
  1. More efficient use of subject resources (i.e., the economy of the design)
  2. Greater comparability of the conditions (i.e., increased control of subject variability)
  3. Reduced error variance (i.e., the treatment-by-subject interaction variability is almost always less than the pooled within-group variability)

# Limitations of the Within-Subject Design

- The within-subject design has both statistical and nonstatistical limitations.:
  1. The statistical problems mostly concern the sensitivity of the assumptions of the analysis
     - The scores produced by a single subjects are more alike than are the scores produced by different subjects (i.e., the observations are not independent).
  2. The nonstatistical problems arise from the fact that the repeated observations must necessarily take place under somewhat different conditions, and some aspect of this difference, other than the treatment being investigated, can affect the scores (e.g., incidental effects: practice and fatigue, memory; carryover effect, contrast effect, context effect)

# Limitations of the Within-Subject Design

- A carryover effect occurs when a treatment has a transient effect that carries over to affect whatever condition is administered immediately after it
- A contrast effect is a carryover effect that occurs when two treatments interact in a way that depends on both conditions
- A context effect occurs when a subject's behavior is influenced by the context provided by exposure to other conditions in an experiment

# THE STATISTICAL MODEL

# Statistical Models

- The difference between the models for the between-subjects and within-subject designs lies in the assumption of independence of the scores
  - Two different models have been applied to within-subject data (i.e., univariate and multivariate)

- In the univariate approach, each score $Y_{ij}$ is viewed as a separate random variable made up of systematic and random components, including a component specific to the subject

- In the multivariate approach, all the scores from a single subject are treated as a single statistical entity; fewer assumptions about the data are required

- The authors emphasized the univariate approach (see p. 373)

# The Univariate Model

- A score $Y_{ij}$ is expressed by the equation:

$$Y_{ij} = \mu_T + \alpha_j + S_i + (S\alpha)_{ij} + E_{ij}$$

- Where:
  - $\mu_T$ is the grand mean.
  - $\alpha_j$ is the treatment effect.
  - $S_i$ is the overall ability of the subject i.
  - $(S\alpha)_{ij}$ is the idiosyncratic response of the subject in a particular condition.
  - $E_{ij}$ is the variability of the individual observations.
  - Note that $S_i \sim N(0,\sigma_S^2)$, $(S\alpha)_{ij} \sim N(0,\sigma_{A\times S}^2)$, and $E_{ij} \sim N(0,\sigma_{error}^2)$.

# Expected Mean Squares

- The expected mean squares are

$$E(MS_A) = \frac{1}{n-1}\sum_j \alpha_j^2 + \sigma_{AxS}^2 + \sigma_{error}^2$$

$$E(MS_S) = a\sigma_S^2 + \sigma_{error}^2$$

$$E(MS_{AxS}) = \sigma_{AxS}^2 + \sigma_{error}^2$$

- So, $MS_{AxS}$ is the error term…because we expect F = 1 if zero treatment effects

# Assumptions…

- For the univariate model, the variances of all the treatment conditions are identical (i.e., homogeneity of variance) and the correlations between the scores are identical (i.e., homogeneity of correlation)

- When these restrictions hold, the data are said to show compound symmetry

# The Multivariate Model

- The set of scores has a multivariate normal distribution
  - A multi-variable extension of the normal distribution
- The multivariate model relaxes the assumption of compound symmetry
- When the assumptions of the univariate model hold, however, the multivariate tests have less power
- We will not talk about the Multivariate approach in this class
  - You will have to take ERSH 8350…

# THE SPHERICITY ASSUMPTION

# The Sphericity Assumption

- A slightly weaker assumption is all that is needed than the assumption of compound symmetry
- Compound symmetry need not hold for the scores themselves, but only for the differences between pairs of scores
- This condition is referred to as circularity or sphericity

# The Sphericity Assumption

- There are tests for violations of sphericity of compound symmetry
- The most widely used of these, a likelihood-ratio test statistic W developed by Mauchly (1940), is included in a number of computer programs
- This statistic should not be significant for the analysis to proceed

# Dealing with Violations of Sphericity

- There are four approaches we can take.
  1. Box (1954a) suggested using the values:

     $df_{num} = e(a\text{-}1)$ and $df_{denum} = e(a\text{-}1)(n\text{-}1)$

where e measures the extent to which sphericity is violated.

Use Geisser and Greenhouse (1958) or Huynh and Feldt (1976), of which the latter has the greater power

# Dealing with Violations of Sphericity

2. The smallest value of $e = 1/(a-1)$ can be used and, hence,

$$df_{num} = 1 \quad \text{and} \quad df_{denum} = n-1.$$

- This is known as the conservative F test suggested by Geisser and Greenhouse (1958) (i.e., Lower-bound in SPSS).

# Dealing with Violations of Sphericity

3. We may use the multivariate approach.

4. We may forget about the omnibus test and use tests of contrasts, which are immune to violations of sphericity.

# SPSS has your back…

**Tests of Within-Subjects Effects**

Measure:MEASURE_1

| Source | | Type III Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|---|
| factor1 | Sphericity Assumed | 1575.000 | 2 | 787.500 | 14.432 | .001 |
| | Greenhouse-Geisser | 1575.000 | 1.610 | 978.061 | 14.432 | .003 |
| | Huynh-Feldt | 1575.000 | 2.000 | 787.500 | 14.432 | .001 |
| | Lower-bound | 1575.000 | 1.000 | 1575.000 | 14.432 | .013 |
| Error(factor1) | Sphericity Assumed | 545.667 | 10 | 54.567 | | |
| | Greenhouse-Geisser | 545.667 | 8.052 | 67.771 | | |
| | Huynh-Feldt | 545.667 | 10.000 | 54.567 | | |
| | Lower-bound | 545.667 | 5.000 | 109.133 | | |

# INCIDENTAL EFFECTS

# Incidental Effects

- Factors such as the position in testing sequence or the type of material are examples of the nuisance variables

- When such a variable becomes an explicit factor in the design, we will refer to it as either a nuisance factor or an incidental factor

# Incidental Effects

- The biases that arise when the treatments are confounded with incidental aspects of the study, such as the order of testing or the materials, can be avoided by breaking up any consistent relationship between them

- There are two ways to do this:
  - In randomization, the relationship between the treatments and the incidental aspects of the study is chosen randomly
  - In counterbalancing, it is constructed in a way that systematically balances the incidental effects across the study (e.g., Latin square, see p. 381)

# Randomization

- The randomization procedures are the easiest to apply, but they cannot assure that the incidental factor is completely balanced across treatment and may have large error term

# Counterbalancing and the Latin Square

- The arrangement of the conditions in Table 17.1 is known as a Latin square
- The key feature of the Latin square arrangement is that every latter appears exactly once in each row and each column

# ANALYZING A COUNTERBALANCED DESIGN

# The Omnibus Analysis

- See the analysis using the numerical example in Table 17.2.
- Two within-subject analyses were performed, one for the treatment conditions (factor A) and the other for the order in which the conditions were administered (factor P).
- The error sum of square is
  - $SS_{residual} = SS_{total} - SS_A - SS_S$

with the degrees of freedom

- $df_{residual} = df_{total} - df_A - df_S$

# The Importance of Interactions in a Latin Square

- The particular configuration of conditions in a Latin square makes it impossible to extract information about any interaction that may be present

- See Table 17.3 for the steps to test an effect after removing the influence of an incidental factor from the individual scores. See Table 17.4 for an example

# Wrapping Up…

- The repeated measures ANOVA partitions variability due to a subject

- Removing such variability aids in the power of the test

- The repeated measures analysis described in this class was an initial first pass at the approach

# Up Next…

- In Lab:
  - How to do repeated measures ANOVA in SPSS

- Homework:
  - Assigned tomorrow morning, due Wednesday, December 2 before class

- Next week:
  - Thanksgiving break
    - Have a good break

- The week after:
  - No reading – intro to mixed models lecture
  - Final exam discussion (bring your questions)
  - Lab on mixed models