Chapter 6 – Selecting and Interpreting Statistics
Chapter Outline

I. General Design Classifications for Difference Questions
   A. Labeling difference question designs.
      1. State overall type of design (e.g. between groups, within subjects).
      2. State the number of independent variables.
      3. State the number of levels within each independent variable.
   B. Between groups designs: each participant in the research is in only one condition or group.
   C. Within subjects or repeated measures designs
      1. Within subjects designs.
         a. Each participant in the research receives or experiences all of the conditions or levels of the independent variable.
         b. Also includes designs where participants are matched (e.g. parent & child; husband & wife).
      2. Repeated measures designs: each participant is assessed more than once (e.g. pretest & posttest).
   D. Single factor (one-way) design
      1. Has only one independent variable.
      2. Factor and way are other terms for group difference independent variables.
   E. Between groups factorial design
      1. When there is more than one group difference independent variable.
      2. Each level of one factor (independent variable) is possible in combination with each level of the other factor(s).
         a. The number of levels of each factor is used in the description of the design.
         b. For example: a design that includes gender (2 levels) and ethnicity (4 levels) would be labeled as a 2 x 3 between groups factorial design.
   F. Mixed factorial design: Has both a between groups independent variable and a within subjects independent variable.
   G. Describing designs
      1. Each independent variable is described using one number that represents the number of levels for that variable.
      2. Example: 3 x 4 between groups factorial design would have 2 independent variables, one with 3 levels and one with 4 levels.

II. Selection of Inferential Statistics
   A. Types of research questions.
      1. Difference questions: compare groups and utilize difference inferential statistics. (Tables 6.1 & 6.3)
         a. Basic (bivariate) statistics: one independent and one dependent variable.

b. Complex statistics: three or more variables.

2. Associational questions: examine the association or relationship between two or more variables and utilize associational inferential statistics (Tables 6.2 & 6.4).

B. Using Tables 6.1 and 6.4 to Select Inferential Statistics

1. Decide the number of variables.
   a. 2 variables = Tables 6.1 or 6.2
   b. 3 or more variables = Tables 6.3, 6.4 or 6.5

(Basic 2 variable Questions and Statistics)

2. If there are two variables and the independent variable is nominal or has 2-4 levels = Table 6.1.
   a. Identify number of levels of IV.
   b. Identify type of research design (between or within).
   c. Determine the type of measurement for the DV.
3. If there are 2 variables and both are nominal use the bottom rows of Table 6.1 (difference question) or Table 6.2 (associational question).
4. If there are 2 variables and both variables have 5 or more ordered levels use Table 6.2 (associational question).

(Complex Questions and Statistics-3 or more variables)

5. If there is one normal/scale DV and the IV's (2 or more) are nominal or have a few ordered levels use Table 6.3.
6. If there is one normal/scale DV and the IV's/predictors (2 or more) are normal/scale or dichotomous use the top row of Table 6.4 (complex associational question).
7. If there is one DV that is nominal or dichotomous and there are 2 or more IV's use the bottom row of Table 6.4 (or 6.3).
8. If there are 2 are more normal (scale) DV's use the general linear model to do MANOVA.

III. The General Linear Model (GLM)

A. Difference between associational and difference questions.

1. Mathematically, the distinction between associational and difference questions is artificial.
2. Both associational and difference inferential statistics serve the purpose of exploring and describing relationships (Fig. 6.2).
   a. The GLM subsumes both associational and difference inferential statistics.
   b. The relationship between the IV and DV can be expressed by an equation with weights for each of the independent/predictor variables plus an error term.

IV. Interpreting the Results of a Statistical Test

A. Statistical Significance

1. The SPSS calculated value is compared to a critical value found in a statistics table.
2. Statistically significant: probability (*p*) is less than the preset alpha (usually .05).

a. Sig.: SPSS label for the *p* value.
b. Usually, if the calculated value (*t*, *F*, etc.) is large, the probability (*p*) is small.
c. This Sig. is also the probability of committing a Type I error (rejecting the null hypothesis when it is actually true).

3. The *p* and the null hypothesis
   a. $p > .05$: don't reject the null hypothesis; results are not statistically significant and could be due to chance.
   b. $p \leq .05$: reject the null hypothesis; results are statistically significant and are not likely due to chance.

B. Practical Significance versus Statistical Significance
   1. Statistical significance does not necessarily insure that the results have practical significance or are important.
   2. Effect size and/or confidence intervals must be examined to determine the strength of association.
      a. It is possible, with a large sample, to have a statistically significant result that is weak (small effect size).
      b. Small effect size may indicate that the difference or association is of little practical importance.

C. Confidence Intervals
   1. An alternative to null hypothesis significance testing (NHST).
   2. May provide more practical information than NHST.
   3. Confidence intervals allow us to determine the interval that contains population mean difference 95% of the time.

D. Effect Size
   1. The strength of the relationship between the independent variable and the dependent variable.
   2. *r* family of effect size measures
      a. Pearson correlation coefficient (*r*): values range from –1.0 to +1.0 (0 = no effect and +1/-1 =maximum effect).
      b. Also includes other associational statistics such as rho, phi, eta and the multiple correlation (*R*).
      c. Can be reported as a squared or unsquared value.
         i. Squared values (r^2) indicate the percent of variance of the DV that can be predicted from the IV, but give small numbers that give an underestimated impression of the strength or importance of the effect.
         ii. Unsquared values (*r*) give a larger value and are recommended for *r* family indices.
   3. *d* family of effect size measures
      a. Focuses on the magnitude of the difference rather than the strength of the association.
      b. Computed by subtracting the mean of the second group from the mean of the first group and dividing by the pooled standard deviation of both groups.

c. All *d* family effect sizes express effect sizes in standard deviation units.
d. Values usually vary from 0 to +/- 1.0, but can be > 1.0.

4. Issues about effect size measures.
   a. *d* is not available on SPSS outputs but can be calculated from information provided on SPSS outputs.
   b. *r* and *R* are available on SPSS outputs.
   c. Most journals now expect authors to discuss the effect size as well as statistical significance.

E. Interpreting Effect Sizes

1. Table 6.5 provides guidelines for the interpretation of effect sizes based upon the effect sizes usually found in the behavioral sciences and education.
2. The absolute meaning of large, medium, and small are relative to findings in these disciplines. Suggest using the following terms instead:
   a. Minimal in place of small.
   b. Typical in place of medium.
   c. Substantial in place of large.
3. Cohen's (1998) examples of effect size:
   a. Small = "difficult to detect".
   b. Medium = "visible to the naked eye".
   c. Large = "grossly perceptible".
4. Effect size is not the same as practical significance.
   a. Effect size indicates the strength of the relationship and is more relevant to practical significance than statistical significance.
   b. However, effect size measures are not direct indexes of the importance of a finding.

V. An Example of How to Select and Interpret Inferential Statistics

A. Steps in the process:

1. Identify the research problem.
2. Identify the variables and their level of measurement.
3. State the research question(s).
4. Identify the type of each research question.
5. Select an appropriate statistic.
6. Interpret the results of the statistic.
   a. Determine if the results were statistically significant.
   b. If the results are statistically significant:
      i. Determine the direction of the effect.
      ii. Calculate and interpret the effect size.
      iii. If necessary, calculate and interpret confidence intervals to evaluate practical significance.

VI. Writing About Your Outputs

A. Methods Chapter

1. Update methods to include descriptive statistics about the demographics of the participants.
2. Add literature based evidence about the reliability and validity of measures/instruments.
3. Discuss if statistical assumptions were violated or not.

B. Results Chapter
    1. Includes a description of the findings.
    2. Include figures and tables to illustrate the findings.
    3. Do not include a discussion of the findings in this section.
    4. Results of statistics should include:
        a. The value of the statistic (e.g. $t = 2.05$)
        b. The degrees of freedom (and N for chi-square)
        c. The p or Sig. Value (e.g. $p = .048$)

C. Discussion Chapter
    1. Puts the findings in context to research literature, theory and the purposes of the study.
    2. Explain why the results turned out the way they did.