Experimental Design

Part II
Richard S. Balkin, Ph.D, LPC-S, NCC

Models for experimental design
- Models for experimental design may be classified as two types:
  - *between groups*, in which outcomes are compared between two or more groups
  - *within groups*, in which a single group is measured across time using two or more different treatments.

The models discussed assume utilization of a true experimental design—utilization of random assignment and comparison groups.

However, random assignment is not always feasible, due to the categorical nature of the factors being studied (i.e. ethnicity, sex, etc.).

When random assignment is compromised, the design is referred to as *quasi-experimental*.

When a quasi-experimental design is utilized, the researcher needs to address the potential threats to internal validity, which is often controlled when random assignment is used.

Between groups design
- In a between groups design, the effect of the independent variable on the dependent variable is based upon the examination of group differences.
- In a true experimental design, one group receives a treatment or intervention (known as the treatment group), and the comparison group typically experiences no treatment (the control group).

However, studies can be conducted in which two separate treatments are being compared.
- For example, a counseling agency may wish to determine if counselors employing a person-centered approach have better outcomes than counselors using a cognitive-behavioral approach when working with clients diagnosed with PTSD.
- In this case, what are the two treatments being compared?

A person-centered approach is being compared to a cognitive-behavioral approach

We will discuss four different types of between group designs

Post-test only
Consider the following scenario:

- A counselor wishes to know whether a peer mentoring program would be effective in assisting students who are at-risk for academic failure. The counselor utilizes the Youth Outcome Questionnaire (Y-OQ-SR-2.0) as a measure of program effectiveness. The Y-OQ-SR-2.0 is a youth survey designed to be repeatedly administered to adolescents to assess their ongoing progress in counseling. It has 30-items on a 5 point Likert scale. Internal consistency was assessed at .91, and is noted to have adequate validity (Wells, Burlingame, Rose, 1999).

Post-test only

- Participants are randomly assigned to a treatment group and a comparison/control group.
- The treatment group receives some type of manipulation or intervention while a control group would receive none.
- A quantitative measure is then used to determine the effect of the intervention.
- In this case, the quantitative measure is the dependent variable and the presence or absence of the treatment is the independent variable.

Post-test only—Example

- Using our example for a post-test only design, the counselor would randomly assign students identified as at-risk for academic failure to receive peer mentoring over the next three months or to receive peer mentoring three months later after the initial group has received the treatment.
- The effect of peer mentoring is being evaluated for change in well-being, as evidenced by the score on the Y-OQ-SR-2.0, the dependent variable.

Post-test only—Model

- The treatment group receives peer mentoring; the control group initially does not, and both groups are evaluated by observing their scores on the Y-OQ-SR-2.0.

Post-test only—Strengths

- The post-test only design is easy to implement.
- Random assignment is very important to help ensure equality between groups at the onset of the study.

Post-test only—Weaknesses

- Without a pre-test, however, the researcher cannot be certain that scores from the Y-OQ-SR-2.0 are the same between the treatment and control groups at the beginning of the study.
- Even with random assignment, equal groups are not guaranteed.
- When random assignment is not used, any change in scores may not necessarily be attributed to the independent variable.
- A quasi-experimental study will not protect against threats to internal validity.

Pretest-posttest control group

- Participants are randomly assigned to a treatment group and a comparison/control group.
A quantitative measure is then used to determine the effect of the intervention.

Both the treatment group and control group receive a pretest.

The treatment group receives some type of manipulation or intervention while a control group would receive none.

After the intervention, a posttest is administered. In this case, the groups are still compared based on posttest scores, but the researcher can be certain of the degree of equal groups at the onset of the study.

16 Pretest-posttest control group--Example

Using our example for a pretest posttest control group design, the counselor would randomly assign students identified as at-risk for academic failure to receive peer mentoring over the next three months or to receive peer mentoring three months later after the initial group has received the treatment.

17 Pretest-posttest control group--Model

Each group is administered the Y-OQ-SR-2.0 as a pretest in order to ensure equality of groups.

The treatment group receives peer mentoring; the control group initially does not, and both groups are evaluated by observing their scores on the Y-OQ-SR-2.0 posttest.

18 Pretest-posttest control group--Strengths

The primary advantage of the pretest posttest control group design is the assurance of equality at the onset of the study between the treatment group and control group.

The use of a pretest makes this design more appropriate when the design is quasi-experimental.

19 Pretest-posttest control group--Weaknesses

However, pretest posttest control group designs are susceptible to testing effects.

A limitation of this design is that participants who are exposed to a pretest may have an idea of how to answer on the posttest in order to appear that they made progress or lack thereof.

20 Solomon four group

Participants are randomly assigned to one of four groups:
- (a) a treatment group that receives both a pretest and a posttest
- (b) a treatment group that receives a posttest only
- (c) a control group that receives both a pretest and a posttest
- (d) a control group that receives a posttest only.

Thus, only one treatment group and one control group are administered a pretest.

21 Solomon four group

Both treatment groups receive some type of manipulation or intervention while both control groups would receive none.

After the intervention, a posttest is administered to all four groups.

In this case, the groups are still compared based on posttest scores, but the researcher can be certain of the equivalence of the groups at the onset of the study and assess the impact of the pretest.
to ascertain whether or not a testing effect exists.

22 Solomon four group--Example
- Using our example for a Solomon four group, the counselor would randomly assign students identified as at-risk for academic failure to receive peer mentoring over the next three months or to receive peer mentoring three months later after the initial group has received the treatment.
- From the treatment group, those students are randomly assigned to be administered a pretest or no pretest.
- From the control group, those students are randomly assigned to be administered a pretest or no pretest.
- One treatment group and one control are administered the Y-OQ-SR-2.0 as a pretest in order to ensure equality of groups.

23 Solomon four group--Example
- The other treatment and control groups do not receive a pretest.
- The treatment groups receive peer mentoring; the control groups initially do not.
- The posttest scores from the two treatment groups and the posttest scores from the two control groups can be compared in order to ascertain whether the pretest contributed to differences in the scores.

24 Solomon four group--Model
- If the posttest scores for both treatment groups and the posttest scores for both control groups are similar, then the administration of the pretest had no effect.
- Assuming there is no testing effect, the posttest scores between the treatment groups and control groups can be compared to determine the effect of the intervention.

25 Solomon four group--Strengths
- The Solomon four group is a strong design because it assures equality of groups, even in a quasi-experimental design.
- Additionally, the researcher can determine whether or not the administration of a pretest is affecting the change in the dependent variable as opposed to the actual intervention.

26 Solomon four group--Weaknesses
- The utilization of this design is limited due to the need for a larger sample size and cost. For example, sample size needs to be doubled in order to have equal representation in each group.
- Additionally, more instruments are required and this adds to the cost of the study.
- Additional time needs to be spent conducting analyses on posttest scores between both treatment groups, posttest scores on both control groups, and posttest scores between treatment and control groups.
- Analyses can be quite complex, especially if the pretest is having an unintended effect on the dependent variable.

27 Factorial designs
- The purpose of a factorial design is to study change in the dependent variable across two or more independent variables.
- For example, instead of simply examining the effect of a peer mentoring program, the counselor wishes to know whether sex plays a role.
  - Is the degree of change in well-being different across a peer mentoring intervention for males and females?

28 Factorial designs
In other words, two analyses will be conducted:

- (a) differences in Y-OQ-SR-2.0 scores across the treatment and control groups
- (b) differences in Y-OQ-SR-2.0 scores across males and females.

29 Factorial designs

- When more than one independent variable is studied simultaneously, the statistical analysis can become quite complex because results may not be able to be generalized across both independent variables.
- This is known as an interaction effect, when the change in the dependent variable is not the same across both independent variables.
- The best way to demonstrate this concept is to graph it.

30 Factorial designs—Nonsignificant interaction

- When an interaction effect is not present, the results can be generalized across both independent variables.

31 Factorial designs—Nonsignificant interaction

- Notice that the same pattern of scores exists for males and females across both treatment and control groups. In other words, males and females in the control group scored lower than males and females in the treatment group.

32 Factorial designs—Nonsignificant interaction

- Thus, the researcher can generalize the findings to both independent variables.
- The researcher could determine whether statistically significant differences exist between males and females and whether statistically significant differences exist between treatment and control groups.

33 Factorial designs—Significant interaction

- When an interaction effect is present, the results cannot be generalized across both independent variables.

34 Factorial designs—Significant interaction

- Notice the effect of peer mentoring on well-being for females showed higher scores for the treatment group when compared to the control group.
- In contrast, scores for males were not affected by the peer mentoring program.
- There were no differences in the treatment group and control group for males.
- In this case the researcher would need to investigate males and females separately.

35 Factorial designs—Strengths

- A factorial design can be applied to any of the experimental designs and has the added benefit of gaining more information because more than one independent variable is being examined.

36 Factorial designs—Weaknesses

- However, the addition of another independent variable creates problems related to sample size.
- Instead of needing a representative sample for a treatment group and a control group,
a representative sample is necessary for males in the treatment group, males in the control group, females in the treatment group, and females in the control group. 

- If a pretest is also added to the analysis or a Solomon four group design is utilized, sample size may need to increase dramatically.

### 37 Within Group Design

- So far, we have examined how participants are affected across a dependent variable when either exposed or not exposed to an independent variable.
- Then, the change in the dependent variable can be compared between the groups.
- Change, however does not occur only because of exposure to an independent variable.

### 38 Within Group Design

- Change can also occur across time.
- A within group design is utilized when a change in the dependent variable in a group is measured across time.
- Random assignment may be employed in some within-group designs.

### 39 Within Group Design

- In the between group design, pretests were often used to assure equality of groups an the onset of a study.
- In a within group design, the pretest also serves as a baseline in which to compare subsequent tests.

### 40 Within Group Design--Example

- For example, the Y-OQ-SR-2.0 is administered at the onset of the peer mentoring study to get a baseline measure.
- Then the Y-OQ-SR-2.0 is administered four additional times on a monthly basis in order to compare progress to the initial administration.

### 41 Within Group Design

- Within Group designs can become much more complex, with the addition of group comparisons and the implementation of treatment conditions.
- Different sequences of a treatment or series of treatments may be highlighted in a within-group design

### 42 Applications to Counseling

- True experimental designs are less common in social science research.
- Many studies in counseling literature utilize intact groups and are quasi-experimental in nature.
- When reviewing a research study, particularly one that did not utilize a true experimental design, counselors should evaluate the degree to which internal and external validity were not substantiated.

### 43 Applications to Counseling

- Threats to internal validity are most common when random assignment is not present.
- However, methods are available to evaluate group equivalence when random assignment is not feasible (i.e. pretests, matching groups, etc.). Counselors should assess the extent to which a study is meaningful to populations outside of the experimental setting.
- Particular attention should be paid to descriptive data, such as age, sex, ethnicity, etc. to evaluate the
extensions of the research to other populations and settings.