Assignment 4: True or Quasi-Experiment

Objectives: After completing this assignment, you will be able to

- Evaluate when you must use an experiment to answer a research question
- Develop statistical hypotheses that can be tested through experiments
- Identify the specific design needed to best answer a research question
- Develop a sampling protocol that will meet the requirements of a specific experimental design;
- Select the appropriate methods of data analysis to answer the research question
- Analyze the strengths and weaknesses of a proposed research design in terms of internal and external validity and explanatory power.

“Rules of the Game” for Assignment 4: True or Quasi-Experiment

The point of this assignment is for you to first create and then critically examine an experimental design that tests an educational intervention. By far the most important part of the process is the critical examination. I am not grading the flow chart per se, but rather your justification for decisions and your overall final “reflective” examination of the design you created in the flow. You should not spend hours and hours working on the flow chart. Therefore, you must each complete a flow chart and have it available in class on November 05. I will give you some time in class to develop your group chart. Failure to complete your individual flow chart will cause big problems in completing the assignment successfully. The vast majority of your time – everything after your discussion in class on November – should focus on the critical examination, not the flow chart. Do not change the flow chart as you answer the questions. I want you to show that you can reflect upon what you developed – identify strengths and weaknesses, suggest changes to improve the design. Every researcher has to “redesign” after the initial conception of a study. Being able to think critically about what you first design is a critical skill.

CRITICAL REQUIREMENT FOR THIS ASSIGNMENT -- READ THIS

I am limiting your theoretical approaches for this assignment in order to save time and effort on your part. I offer three alternatives. Pick one. You do have to understand enough about the theory to figure out a reasonable treatment based on it, what the independent or predictor and dependent or outcome variables would be and what other variables that could affect the outcome (neither independent nor dependent, but intervening) you would need to measure in an experiment. You should spend no more than 4 hours researching the theory well enough to develop a detailed diagram (with SEVERAL constructs and linkages) and a table defining the constructs. I provide one pretty comprehensive reference in each case. That should make the four-hour search possible. Divide the task. Remember, expertise in the theory is NOT a requirement for this assignment. I am trying to make this easy. These theories have well developed constructs with enough detail and specificity to work well for this assignment. You must have more than a general diagram with “three big circles or boxes” as social learning theory or socio-ecological theory are often illustrated. Those simplistic models are a path to failure in this assignment.

Integrated Theory of Health Behavior Change developed by Polly Ryan. It has depth. It has detail. It can be used to understand and explain ANY individual behavioral change. There are good examples of using this theory to develop interventions, which will make it
easy for you to see how to use it to create an experiment. In fact, I used it for the Haiti Extension Experiment. This link from the US National Library of Medicine (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2778019/) gives a detailed explanation of the constructs in the theory. It is by Ryan, the person who developed this theory of behavioral change that incorporates critical constructs and linkages of many other theories. It integrates several theories – another reason it is a good choice for you. It also includes a good (short) review of the many decades of failure to achieve individual behavior change using various approaches. This article by Sheeran et al. is a more recent (2017) review of health related behavioral change theories. It is very thorough. The appendix defines many constructs used in behavioral change theories of all types. These two articles and a few of the references in them should suffice.


**Positive Youth Development** as presented by Benson, Scales, Hamilton & Sesma, Jr. is a contemporary mid-range theory that is closely tied to programs that seek to foster positive youth development and/or reduce risky behavior. To be honest, it remains in my view “underdefined.” The classic presentation is three big overlapping circles that include pretty much every social influence that we experience. However, it is the model for 4-H and many other youth development programs. It, too, is an integrative theory that builds on extensive past work and is used to develop interventions. To use it effectively, you simply must get past three big circles as the constructs in the theory. Think of them more as “realms” or “dimensions” of the many factors that influence youth development. I offer two resources that I think will help. Again, I hope these are sufficient for what you need in this class. While a lot of the work has been done with adolescents, the theory applies through emerging adulthood (18-22 years or so). Like the health behavior model, I selected it because you can use it to try to understand almost any problematic behavior by youth and to understand almost any positive developmental outcome. I offer three resources. “Positive Youth Development: Theory, Research, and Applications” by Benson et al. is a good review and provides details about the theory and includes a review of the existing body of knowledge along with an assessment of where knowledge is well developed and missing. There is an issue of the journal *New Directions for Child and Adolescent Development* published in 2008 that focused solely on PYD theory. The first article by Guerra & Bradshaw will give you an overview of the five competencies that are central to this theory. The other articles in the special issue give examples of its application.


Value-Belief-Norm theory has been widely applied to support for environmentalism and pro-environmental behaviors, but it can be applied to understand a wide variety of behaviors and offers an alternative to purely rational actor theories. It has well defined and substantiated constructs and many instruments to measure the constructs have been developed for the environmental applications. It is unfortunate that this theory has not been applied to a wider range of pro-social behaviors. I am not sure why this is the case. **VBN as it is commonly termed is a theory of that focuses on the roles of an individual’s values and beliefs and the norms of the reference groups that influence the individual in behavioral decision-making.** If you are interested in approaches to getting people to engage in pro-social behaviors (including things like pro-environmental behaviors, funding public services for the disadvantaged, participation in social movements like Black Lives Matter, and volunteerism), this could be a good theory for you. It ties individual behavioral change to larger societal processes and influences, and ultimately to larger societal outcomes. An article by Stern and others provides a detailed presentation of the theory and its use in the environmental arena (cited below). VBN theory can also explain other phenomena like consumers’ expression of values and beliefs through their purchasing behaviors. I provide three items that I think will provide you with enough background to use this theory as the basis for your design. Stern lays out the **basic constructs in VBN theory** and established the relationship to environmental behaviors. He provides very clear definitions of the constructs and relationships among them. **Jansson examines consumer adoption of “high-involvement” (required a lot of effort) eco-innovations.** Chen’s work is a contemporary application of VBN to understand **how people react to climate change.** His model is a good one. Again, you can adapt VBN to any behavior in which values, beliefs or behavioral norms could play a role in behavioral decision-making and change. It just has not yet moved into the general social science sphere for reasons that I do not understand.


**Context for the Experiment. I do not care what context you choose for your experiment – where your intervention would happen, the theoretical population of interest, or what you want to achieve with the intervention. You will greatly reduce your workload and probably do a better job on this assignment if you select a context with which one of you is thoroughly familiar.** If you can rely on one team member to know the literature about the PIN and the specifics of the context in which you would conduct your experiment, the rest of you will not have to spend as much time exploring the literature.

There are three components in this assignment. Submit a separate Word document for each:

A. **Diagram of the theoretical framework you use.** Title this document **LastName1 of Team Members in Alphabetical Order_A4_THEORY,** e.g., Alvarez_Bailey_Nkedi_Zhang_A4_THEORY

Assignment 4: True or Quasi-Experiment, FYC 6800, Page 3
B. A completed flow chart. Title this document LastName1 of Team Members in Alphabetical Order_A4_FLOW, e.g., Alvarez_Bailey_Nkedi_Zhang_A4_FLOW

C. A reflection and self-analysis in which you analyze the strengths and weaknesses of your design from the perspectives of external validity, internal validity, and explanatory power. Title this document LastName1 of Team Members in Alphabetical Order_A4_REFLECT, e.g., Alvarez_Bailey_Nkedi_Zhang_A4_REFLECT

THEORY

1. Define the nature of the problem, issue or need that you want to address (500 words maximum). This is a description of the “what and who” of interest to you. What is the PIN (problem, issue, or need)? What is the population of interest (adolescents, for example)?

2. What is the theoretical basis for your research? Include a diagram or model of the theory you will use for this assignment like the ones I provided during our first class session and the one you provided in Assignment 1. Include a table with a definition of each construct in the theory as you did in Assignment 1. You have to be specific. MORE detail is better. It makes the assignment easier and it is more likely that I will understand your experiment well. No simple big boxes with lots of arrows going both ways, overlapping circles that include just about everything like “proximal, intermediate and distal environment.” Many carefully defined constructs interconnected in one diagram is what you need to create. This model should include the constructs that will serve as the basis for your explanatory research objectives.

FLOW

The more detail you provide in the template, the more easily I can understand what you plan to do. You do not need pages and pages, but if all you put under comparison groups is “treatment and control,” it will be hard for me to assess your understanding of how to set up comparison groups for an experiment. You need to develop the flow chart as a group. Start by agreeing on the topical, explanatory and (if you have any) theoretical objectives (remember that means BUILDING theory, not just using theory). At least one of you must have a good knowledge of the topic. All of you have to know the theory and be able to identify theory-based objectives for understanding or explaining the topic (behavior) of interest. State your explanatory objectives in terms of the constructs of interest in the theory. State theoretical objectives only if you plan to add to (build) theory. Using theory and building theory are NOT the same thing. I would encourage you to think about testing or building theory, but this is not a requirement. You must reach agreement about the objectives. Do so quickly and complete the flow chart quickly. Otherwise, you cannot complete the critical component in this assignment, your critical reflection of the decisions you make in creating the flow chart. The reflection is where you will show me that you have mastered the key concepts of research design and know how to use the design literature. Once you have defined your objectives, spend no more than 2 or 3 hours on completing the flow chart.

Getting your ideas “right” in Box 1 is critical. Here are some pointers to think about as you develop your question and objectives.

Research Questions. Your research question should be a “thick” one. Theory-based research questions deal with explanation. We have discussed this idea in detail. Answering a “good” research question provides the knowledge that we need to solve problems, meet needs,
understand phenomena of interest, or know why a program or intervention succeeds or fails. Theory-based research questions do not focus on “what happens,” but rather contribute to understanding of why and how things happen. **You must pose a “how or why” question that is designed to provide evidence of a direct cause and effect relationship between two or more constructs in the theory you chose applied to the subject (topic) of your research.** The design you create therefore has to provide sound evidence of a direct cause (your intervention) and effect (outcome) relationship. Make sure your question is one that you can answer with an experiment. For example, a question like “What is the relationship between parental monitoring and alcohol use and adolescent substance use,” would **not** be an appropriate question for the experimental group. Your question is **your question**, but make sure you develop it carefully. Later, when you start to answer questions in the Reflection, you may decide that not all of your decisions were “good” ones and see ways you could have improved the question. That is fine as long as you discuss your ideas in the Reflection. It is better than correcting the flow chart because you will show that you do understand how to apply the key concepts of research design to an analysis of your own research. You must establish one outcome variable, two or more comparison groups (at least treatment and control), and at least one (probably several) non-experimental variables that may affect the outcomes of the experiment, excluding demographics.

**Objectives.** This decision about the contributions to the body of knowledge is the starting point for research. It will drive the remainder of the decisions you make. I know that you are probably not experts in the theory. That does not matter because I want to understand what you considered as you thought about contributing to the body of knowledge. Students tend to be able to identify ways to add to the empirical body of knowledge through description, but adding to explanation and theory is harder. Your intervention should focus on improving our understanding of how to achieve behavioral change through one or more interventions. Put another way, the objective of doing an experiment is to see “what works and more importantly what does not work.” Your objectives should put your idea (what you think will “work”) to the toughest possible test. Keep thinking about that – how to make the experiment a truly tough test – all the way through the design process.

You may want to contribute to theory – have a theoretical objective. I encourage you to do so, but it is not a requirement. There are four ways you can contribute to theory. If you decide you do have a theoretical objective, **pick just one of the four** in order to keep the task manageable.

**Theory Testing.** Be specific about the contributions in the flow chart – e.g., do not just say “we will test the theory of ...”. Think carefully especially as you explain and justify your decision. Have a reason for selecting one type of theoretical contribution over another. For example, if you are using a theory that is very well established, theory testing probably will not constitute a good contribution to theory. However, even some very well developed theories may have some linkages that are not well established, in which case testing those linkages would be very useful. Theory testing is particularly useful with recently developed theories. Usually, a theory has to prove itself useful to understanding a phenomenon in a number of tests before it is accepted by the scientific community. Testing new theories is therefore very important, but interest in testing generally declines as a theory becomes more widely accepted.

**Comparing Theories.** Comparing the explanatory power of two or more theories is another common approach. We try to decide which of several competing theories offers the best or most complete explanation for some phenomenon. For example, we might have two or three theories that are widely used, with different researchers arguing that one or the other has
better explanatory power. Research that tests the theories through an intervention helps us decide which is more robust, and experiments where both theories are applied is a very good way to decide “which one works best.”

**Building Theory.** Building or extending theory is a very common objective for researchers. One student of mine took a theory commonly used to explain environmental behavior and added constructs from other, related behavioral theories to it. Her objective was to generate a more explanatory theory – one that does a better job than the original of explaining the behavior of interest. She used the environmental education model and added parental and institutional influences on the environmental behaviors or emerging adults (freshmen and juniors in college), drawing on the theory of emerging adulthood.

**Extend Domain of Theory.** Researchers may want to test the degree to which a theory is universally applicable. This involves taking a theory that has shown good explanatory power with some populations, or in some contexts, and applying it with a different population or under a broader range of conditions. Piaget conducted his early work dealing with the developmental stages of children, for example, with children from Europe and North America. Later researchers wanted to know whether the stages of development that he proposed were culturally specific, or whether they occurred for all children. They therefore tested his theory with children from many widely differing cultural backgrounds.

**Specific Design.** Name the design and be as specific as you can. Whatever you do, do not just pick one because it sounds good. In the flow chart, identify the general type of design (true or quasi-experiment) that you will use and the specific research design that you will use, by name (Solomon four group, switching replications, etc.). To do this, look at your notes from class, the material in your text, and outside readings. As you explore the literature about experimental designs, you will see that there are other experimental designs that we have not discussed in class. Consider these options, too, as you make your decision.

**REFLECTION**

The Reflection should take at least 75% of the time you spend on this assignment. To be successful, this must be a group product. This is where you show me that you have a thorough grasp of the concepts we have studied and that you can use the research design literature to improve your own work. Remember that experiments have one purpose – to test for direct cause and effect relationships. Think through what you set out to do. Review the research question(s) and the objectives. Review your answers and think about the strengths in your design and the weaknesses. Do not try to redo the assignment by repeatedly changing the flow chart. The objective is for you to analyze your own work critically. Treat your work as though YOU were the reader, not the researcher. Try to answer three questions. What aspects of your design would convince you of the value of the work – specifically your ability to test direct cause and effect in this case? (2) What weaknesses in the design would you spot? (3) How could you address weaknesses?

There are three critical components you must include in answering Q1-3.

**Component 1: Explain your decisions.** I know what you plan to do. That is in the flow chart. Now I want you to explain why you made the choices. What was your logic? For example, you may have decided to use a Solomon four-group design because you want to minimize the effects of testing bias in the results. You might decide to use a switching replications design because you want all of the participants to receive the treatment for ethical reasons or because you want to know if the effect of the treatment is persistent. You need to provide these kinds of
explanations in the Reflection – not in the flow chart. I will address several specific considerations below – based mostly on the kinds of issues that have caused problems for students in the past.

Component 2: Think critically about your decisions. Identify the strengths and weaknesses in your design with regard to internal validity (establishing direct cause and effect), external validity (nature of the generalization – theoretical or statistical), and explanatory power (contributions to the body of knowledge, focusing on explanation and/or theory building). Identifying weaknesses will not have a negative effect on your grade. I want you to put the flow chart together quickly – not spend days on it. This is your first attempt at building a research design. Being able to examine your decisions before you start to collect data is critical for all research so that you can identify weaknesses and correct them. Do not spend time “redoing” that chart. Explain your decisions (point 1) and what you think may need to be changed based on further consideration (point 2).

Component 3: Use the research design literature – consult it, think about it, and incorporate what you find into your answers. This is what I mean when I say “Use, cite and reference the research design literature.” This is an absolute requirement for success and accounts for 40 points on this assignment. Do not try to answer the questions and then find literature to support what you have already decided to say. Use the literature to decide how to answer the three questions.

1. Consider your design and protocol as a whole, including sampling and planned statistical analyses, with regard to internal validity. Remember that internal validity, for a realist, refers to the confidence that the conclusions that you reach are valid and justified. Internal validity involves both providing evidence that the explanation you propose (my idea that by increasing perceived power I can increase change in behavior) is valid and by your ability to show that other things probably did not cause any change you observe (eliminating alternatives).

What are the greatest strengths of your design and protocol with regard to internal validity? Put another way, what aspects of your design and protocol would reassure a reader that your conclusions are justified? Refer to the specific threats to internal validity that we have discussed in class. Consider each of them carefully as you answer this question. Be specific in your responses. Do not repeat broad generalities about the kinds of strengths your design has in general. Rather focus on your specific design decisions.

For example, in my proposed study I did four things that would give me some confidence in my conclusions. First, I set up a pre- and post-test of behavior for each. Any change that occurred should be the result of the training, and the treatment group should show a greater change in behavior. Second, I tested perceived power pre- and post-test for both treatment and control groups. I am proposing that I can change this condition by training. Any difference between pre- and post-training scores should be due to the training. Further, I should see a greater difference in the treatment than in the control group. Therefore, including this variable increased the internal validity of my study. Third, I also collected data about the intervening variable (behavioral intent). This helped, too. Finally, I plan to run regression analysis where I can confirm (or not) the proposed linkages between all these variables and, more important, I can see if perceived power is a more important (higher beta value) predictor of behavior in the treatment than in the control group. If so, this is one more piece of evidence that my idea is valid and justified. To eliminate other variables, I included a pre- and post-test of knowledge (this is something I really do usually do). Knowledge is not a construct in the theory of planned behavior,
but we do know that knowledge is a pre-requisite for behavior. Therefore, I need to
determine whether prior knowledge affected the outcome. Maybe all my participants
were very knowledgeable before the training – resulting in no affect. Worst yet, perhaps
there was a pre-training difference in knowledge for the two groups. By including
knowledge as a construct and testing pre- and post-training, I should be able to eliminate
or at least evaluate the importance of knowledge in the outcome.

What are the greatest weaknesses of your design with regard to internal validity?
Consider each of them carefully as you answer this question. Be specific in your responses.
Do not just repeat broad generalities about the kinds of weaknesses your design has in
general. Respond in terms of your study. For example, in my study I have a potential
problem because my sample is essentially a volunteer one. I cannot force people into my
training sessions. They sign up because they are interested. For example, people who want
to lose weight sign up in weight loss classes. If my training were “Diet Management for
Weight Loss,” I could very possibly attract people who really want to change their eating
habits. This could mean that almost any training approach will work, that I will not see a
difference due to training method. On the other hand, since I conduct my training in a one-
time, short-duration setting, I probably will not have to worry about mortality (dropouts) and
that is a strength.

2. Consider your design and protocol as a whole, including sampling and planned statistical
analyses, with regard to external validity. Remember that external validity, for a realist,
refers to the degree to which we can extend our findings, and especially our conclusions, to
people, places, conditions, or circumstances other than those under which the study was
conducted. Sometimes we want to generalize statistically and we usually want to
generalize theoretically. Identify clearly your planned generalization.

Identify the key strengths of your study with regard to external validity. Consider both
statistical and theoretical generalization. Be specific. Do not just include generalities
about the design in general. Refer to what you plan to do. For example, unlike some
experiments, my study is not at all contrived, so artificiality should not be a problem. By
including a pre- and post-training test of knowledge and including this in the regression
model, I strengthen theoretical generalization because I can compare the significance of
knowledge (a key construct in other theories) and perceived power as predictors of change
in behavior. If it turns out (which it does when I do this) that knowledge does change, but
that it does not predict change in behavior, I go a long way toward saying that the theory of
planned behavior is a good one when you have to deal with these short-term exposures to
trainees.

Identify the key weaknesses of your study with regard to external validity. Consider both
statistical and theoretical generalization. Be specific. Do not just include
generalities about the design in general. Refer to what you plan to do. People volunteer for
my study because they want the training – they presumably have high motivation. This may
pose some problems for statistical generalization. I probably can generalize the general
conclusion that discovery-based learning will, in general, with lots of kinds of people, topics,
places, etc., be superior to the traditional lecture-based learning in creating behavioral
change. However, given the potential for high motivation on the part of my participants, I
probably cannot generalize the degree of change in behavior from my study to other
settings. In situations where motivation is not nearly as high, the amount of change might be
considerably less. I also need to be careful about extending my conclusions to other kinds of
training settings. I have a one-time, short-duration exposure to trainees. We cannot, based
on my findings alone, conclude that the focus on discovery learning would result in similar behavioral changes under very different conditions. This is an opportunity for more research – for example, in classroom settings where there is longer-term, repeated exposure.

3. Discuss the contributions to **explanatory power** of your study – your contributions to the body of knowledge. Remember that explanatory power rests on the degree to which the researcher can explain the phenomenon under study in a complete and robust way. Explanatory power rests, to a large degree, on accumulating multiple kinds of evidence to support (or not) a theoretical explanation. Contributions to explanatory power often come from providing a different **kind of evidence**. For example, if most of the research about a specific topic has relied heavily on a single type of design, the scientific realist would argue that evidence from the other design groups is needed to enhance explanatory power. Contributions may also come from addressing parts of the theory that have been largely ignored in the research base. This may include, for example, verification of some aspects of the theory, extension of the theory to areas where it has not been tested before, or building new components into the theory, to name a few. What are the contributions of your study to explanatory power? What kinds of evidence will we have that we have not had before? What will we know or understand after you complete your study that we did not know or understand before?

**More Guidance**

I received requests from students last year to provide more guidance about responding to the Reflection. Here is my attempt.

**Requirements.** The terms independent variable and predictor variable mean the same thing. I prefer predictor variable. Please remember that **screening criteria are not independent or predictor variables**. Demographics are usually no more than descriptors of the sample, although they are often included as non-experimental predictor variables to control for variance if you think demographics will affect the outcome. **Treatments are not variables.**

Identify **one outcome variable** and at least **one experimental or predictor variable for each of the constructs in the theory** you are testing. The relationship between these experimental predictor variables and the outcome variable lies at the heart of your explanation and your test of direct cause and effect. Experimental predictor variables represent constructs in the theory. For example, I measure behavioral norms when I use the theory of planned behavior, although I do not try to use that as the basis for the intervention.

You must also include **at least one non-experimental predictor variable**. Almost all experiments use non-experimental predictor variables as a way of controlling for factors that could affect the outcome of the experiment – but that you cannot control or manipulate. For example, you probably cannot isolate people from all other learning about a topic (wildlife conservation, eating habits) that you want to address through an intervention. In this case, you will have to include various independent predictor variables to statistically control for the influence of these non-experimental factors in your experiment.

A long list of predictor variables, particularly non-experimental predictor variables, is problematic because you have to think about how you will measure each one and it is preferable to control for these factors through screening to get a homogeneous sample and through the establishment of comparison groups when feasible. Explain your reasons for selecting each non-experimental predictor variable and the steps you took to minimize the need to incorporate non-experimental measurements.
Show that you understand three things: (1) The differences between group, independent or predictor, and outcome or dependent variables; (2) how level of measurement affects statistical analysis; and (3) how to incorporate variables (especially predictor or independent variables and groups beyond the treatment groups) that allow you to make maximum use of your data.

Example that may help, I use the theory of planned behavior as the basis for designing my training programs. My outcome variable is change in behavior – that is my ultimate measure of the success of my programs. However, I also want to know why the program works (or not). Based on the literature, I have decided that the predictor variable that I can probably affect in one-time, short-duration exposures to trainees is perceived behavioral control, and specifically perceived power or self-efficacy. I am spending a lot of time and effort creating training materials and programs that focus on changing this condition because I think it is “doable” in my normal setting and because the literature indicates that perceived power has a very strong influence on behavioral intention and ultimately on behavior. I need to know not only whether people are changing their behavior because of my programs, but also whether perceived power is the key to creating that change. Otherwise, I am wasting a lot of time and effort because discover-based learning, key to perceived power, is a lot harder to do than the simple lecture. Therefore, in a true experiment I could assign people to treatment (discovery-based learning) and control (traditional lectures). I could measure the outcome – behavior – prior to and after the training. However, even if there were a difference between treatment and control group after the training, I would not know why. I at least need to measure perceived power (or self-efficacy) to make sure my focus on discovery-based learning is not a waste of time. I probably should measure behavioral intent, too, because that construct intervenes between perceived behavioral control and behavior in the model. Finally, even though knowledge is not a construct in the theory of planned behavior, we know that knowledge is a pre-requisite (but insufficient) condition for behavior change. Therefore, I also need to test change in knowledge. Therefore, my true experiment would have one dependent or outcome variable (behavior), and three independent or predictor variables (perceived power, behavioral intention, and change in knowledge).

General or Research or Working Hypotheses. This is a key decision because it plays a crucial role in setting up your intervention or treatment, deciding who comprises the theoretical population, and other decisions. For example, I use the theory of planned behavior as the basis for my Extension training. In that case, my theoretical hypotheses are that (1) behavioral change will be greater using a discovery learning approach than a traditional lecture and (2) that perceived power will be the most important construct that I can affect in the kind of one-time, short-duration training sessions I conduct in Extension. Once I decide that, my treatment is defined. Make sure you are very clear about the constructs represented by the treatment and which constructs will serve as the basis for the independent and dependent (or predictor and outcome) variables. Most research questions focus on just a few of the key concepts and proposed linkages between them that are present in a theory – especially experiments. Do not try to test every relationship in the theory or include every construct. Focus on a specific, “doable” test of some component in the theory.

Theory-Based Intervention or Treatment. Describe the theory-based intervention you will use as the basis for your experiment. Provide enough detail for me to understand how the intervention draws upon the theory and for me to see how you will implement the intervention. That is, explain what you will “do” to the participants in the treatment group(s) in the flow chart. You may have more than one treatment. Remember that simply testing someone or asking them to fill out a questionnaire is NOT a treatment. A treatment is something that you (or some other member of your team, or perhaps even some other group altogether) do to people.
The treatment was meant to **change** people. The example of the health program in Oregon is a good one. The researchers did not implement the health program. They just convinced the Oregon Dept. of Human Resources to divide participants into two groups. And, of course, they collected the data. If the design is factorial, describe the different levels of each factor included in the experiment. If you are using additional comparison groups, such as male and female groups, also describe them and provide a justification for using them.

**Hypotheses.** Remember that one-tailed statistical hypotheses are superior to two-tailed hypotheses. For example, my main statistical hypothesis is that participants in the treatment (discovery-learning) group will demonstrate greater behavioral change than participants in the comparison (lecture) group (statistically significant at the 0.05 p value). Another hypothesis is that perceived power will be a stronger predictor of behavioral change than knowledge or behavioral intent, indicated by beta scores.

**Sampling.** You probably need to address the following factors in sampling: why the specific theoretical population, accessible population, sampling frame (if you use one, which you well may not) and the screening or criteria for inclusion in the sample. Do not repeat what you put in the flow chart. I want your reasoning – for example, why you think the accessible population adequately reflects the theoretical population? You can include specifics you thought about here. For example, it is very highly likely that you will need to use screening criteria. I will probably be disappointed if I see that you are screening based only on demographic traits. Show that you know how to set up reasonable and logical screening criteria that will help you make sure that your sample is homogeneous with regard to characteristics that may affect the outcome of the experiment. This is critical to establishing direct cause and effect. One key point here is to make sure that the sampling logic, specific type of sample and replacement procedures **are appropriate for maintaining experimental control and are “doable.”** Explain how you will assign the participants to the treatment and control groups. The only case in which you should fail to assign participants to groups is that in which your design is a quasi-experiment using some sort of matched comparison (control) group. You may have other comparison groups as well, such as males and females, or under 21 and over 21 in factorial designs in particular. Justify your decisions.

**Data Collection Procedures.** Your comments should focus on aspects of data collection that might affect the outcome(s) of the study. This is not a methods class and this is of relatively little concern here.

**Statistical data analysis.** I encourage you to use statistical analyses for this experiment, but this is not a requirement. Experiments do use qualitative analysis as well. You will have the tests for each hypothesis in the flow chart. Think of this as your data analysis plan or strategy. Explain how the steps in your analysis fit together to answer your research question. Focus on the logic of the process, not the details of the analyses. Consider procedures you can use to make your analyses more reliable and/or precise. Your data analysis strategy should be appropriate for the design. Here is an example.

**Example of analysis plan.** My experiment would have three steps. First, I would run a paired t-test comparing treatment and control groups for the outcome variable (behavior change). This is a paired t-test because I am testing **change** in behavior, not just behavior. Therefore, I have to use a test that looks at both the pre- and post-training score for behavior. My statistical hypothesis was that “Trainees who receive a discovery-based learning experience will show significantly (alpha 0.05) greater change in behavior than those who receive a traditional lecture-based learning experience.” This test will reveal whether my hypothesis was confirmed or not – whether the discovery-based
learning was superior with regard to behavior change. Second, I would run a paired t-test for each of three predictor variables, knowledge, perceived power and behavioral intent. These tests would reveal whether discovery learning affected the predictor variables. My hypothesis was that all three variables would be significantly greater for the treatment than the control group. I need to use a paired test because I am using change in knowledge, perceived power and behavioral intent. For all of these tests the data have to be interval. The scores have to be normally distributed. The variance must be equal for the two comparison groups, treatment and control. Finally, I will use linear regression to examine the relationships between the predictor and outcome variables.

Qualitative Data Analysis

Do not repeat the steps you described in the flow chart. Focus on what each step in the analysis will reveal – why you are doing it. There is an example below, but do NOT just repeat what I say here as your response. Your analysis process should answer your research question. There is no cookie cutter answer to this question. Provide a robust explanation that demonstrates that you understand the connections between design choices and analysis process.

Example for Qualitative Data Analysis. One of my students compared chronic (long-term, multiple occasions) and temporary (short term, one occasion) homeless people. She did not use an experiment, but her data analysis can give you an idea of what you need to explain. She used interview data – narrative information. Her first step in data analysis was to use her interview notes to create a profile of each interview. She identified the responses of each participant that corresponded to each of four constructs of interest, group identity, group cohesion, leadership and goal definition. She also identified any responses that did not “fit” with any of these constructs. This step allowed her to capture the key components of each individual set of responses to the interview question. Her second step, conducted every time she accumulated three or four new interviews, was to identify theoretical and emergent themes. Themes are responses that have a shared comment – that are similar in meaning. Theoretical themes are those that “fit” one of the four constructs of interest. Emergent themes are those that do not “fit” within any of our pre-identified constructs. This procedure allowed the student to identify the commonalities and differences in responses among respondents. Ultimately, she compared the chronic and temporary homeless groups in terms of their potential for self-organization, based on the themes that emerged for each group.

More Guidance – An Example

I got a request to complete an example. Look at the Haiti Extension Experiment as a complete example. However, know that one is complex and I am providing another example here. After careful thought, I decided to limit this to one aspect of the assignment – sampling considerations. I am afraid of creating a “cookie cutter” kind of approach on your part if I do an entire assignment. I take the case of a quasi-experiment designed to test which of three theoretical approaches to teaching adolescents to control bullying in their social environments is most effective. The following is a bullet list and that’s FINE for this assignment. This is my Step 1 – explain what I decided to do. In a real submission I would have a reference for almost every point in this list.

A. Comparison Groups
   1. We included a control group (no treatment) in the design in addition to the three treatment groups based on the three theories because there is a lot of focus on teaching
adolescents to control bullying – in school, in the media, by parents. The use of this control allows us to estimate the effect, if any, of this general learning on the experimental effects of the treatments and improves ability to establish difference in cause and effect between the three approaches.

B. Assignment to comparison groups
1. There is no possibility of assigning individual youth to the different comparison groups, which is why we chose a quasi-experiment.
2. We are asking local BGCs (Boys & Girls Clubs) to implement the training programs. It is not reasonable to expect the local club leaders to have no preference in which of the three programs to implement. We do not have funds to hire people to do it for us. Therefore, we will also be unable to assign clubs to treatments randomly. We will assign specific clubs to one of the three treatment groups based on the club leaders’ preferences. This is a major weakness in the design because it introduces potential bias into the implementation of the intervention.
3. The adolescents are more likely to feel at ease with training on this sensitive topic if the educator is a trusted adult. This is a major justification for the BGC as the implementing units. We would weaken the ability to show direct cause and effect by introducing a new group of strangers into a known environment. Using trusted adults as the trainers reduces the potential for fear and anxiety reducing response to the training.
4. The quasi-true control group will consist of students in schools outside the communities served by the collaborating BGCs.

C. Theoretical population
1. Adolescents in metropolitan areas in Florida who attend public school – Florida has a culturally and socially diverse population. At this stage, it is better not to try to extend results to other less diverse populations because prior research has established differences in bullying intensity and frequency based on cultural differences among populations. Bullying tends to be both more intense and more frequent in socially diverse populations. This is NOT demographics, but rather refers to social and cultural definitions of what is treated as acceptable behavior toward other children.
2. We chose metropolitan areas because it would be impossible to maintain separation of the test group (members of BGCs) and the quasi-true control (general school population) in small communities.
3. Metropolitan areas in Florida account for a high percentage of all adolescents in the state, which increases the ability to extend conclusions statewide.
4. We require students who attend public school for two reasons. (1) Bullying is very common in school settings which may not be true among home school groups or youth in individual home school environments (we found no research in such environments). (2) Public schools are the primary educational setting experienced by most adolescents.

D. Accessible population [BGCs in metropolitan areas in Florida]

E. Sampling frame
1. BGC of Florida provided a list of BGCs in the state.

F. Screening criteria
1. No reported incidents of serious bodily or psychological harm due to bullying at BGC events. This would indicate that the problems of bullying are beyond the typical and would introduce greater variance into the research setting without adding to explanatory power.
2. Members have not received training about bullying within the past one year. We need to establish the direct effect of the three training programs. Prior “contamination” through training would greatly limit our capacity to establish direct cause and effect.
3. Group leader has not been highly proactive in personal approaches to training members about bullying (e.g., has followed what BGC considers a typical pattern of discussions of this topic with members). This criterion eliminates another potential non-experimental source of variance in the outcome.

G. Sample size considerations
1. Variance within and between BGCs is a concern.
2. We will conduct extensive pre-experimental research to identify these sources of variance and those data will drive many considerations of sample size both with regard to number of BGCs and number of individual youth in the BGCs. Variance between BGCs could result from several factors, including:
   i. the importance placed on anti-bullying behaviors and education by the group leader,
   ii. the extent to which the local school system provides education about bullying and anti-bully programs,
   iii. the social setting of the metropolitan area and the specific community served by the BGC,
   iv. and the degree to which parents have tried to help their children avoid bullying (as the bully or the victim).
3. We will use a confidence interval of 0.90 or 0.95. This decision will rest on the variance and changes in sample size needed to achieve the 0.95 interval. If variance is great, we may need to reduce the confidence interval to 0.90 or 0.85.
4. Precision is +/- 5 on standardized post-test scores, typically sufficient when the instruments were used to establish differences between pre- and post-test.

H. Procedures
1. BGC of Florida has agreed to provide a letter to all group leaders in the selected metro areas asking them to volunteer to participate in this study and expressing the state and national leadership’s interest in this research.
2. They have also agreed to follow-up with a second invitation if initial response rate is low. The endorsement of the statewide council will be important in securing volunteer clubs for this effort. Outsiders suggesting the training would most likely not be acceptable to BGC and would not engender trust in the collaborating groups.

I. Response rate
1. The time and effort by club leaders is not excessive, but each intervention does consist of 5 lessons, each of which requires approximately one hour.
2. Some parents may object to the training, in which case we anticipate that local leaders, even if interested, would withdraw or never volunteer to participate. We will raise the issue in the informed consent procedure.
3. We have a grant that will provide each participating club with $1,500 for participating in the study. We believe this will encourage participation, but is insufficient to provoke leaders who do not want to participate into doing so. We will verify this with the statewide council prior to firmly establishing the payment.
4. In a worst case scenario, we will reduce the treatments to two programs (the most commonly used program in schools today and a new program that has recently emerged) to reduce needed sample size

J. Replacement procedures
1. We have rejected replacement because this study is sensitive to external events (e.g., some child hurt by bullying, stronger emphasis on anti-bullying during experimental implementation)
2. We therefore plan to oversample BGCs to account for simple loss of statistical power due to drop out (mortality)
3. We will follow-up with BGCs that do drop out to try to implement our post-experimental tests of knowledge and behavior with their members in order to understand whether there was any effect of the educational program.

4. We will interview BGC leaders to understand why they dropped the program and incorporate their insights into the interpretation of our results, particularly our interpretation of the degree to which we can generalize our conclusions. If the reasons are varied and do not seem to be related to the specific interventions, we will not treat this factor as a threat to generalization. If there is some pattern related to the training itself or the topic of the training, we will need to carefully limit generalization and may need to avoid generalization beyond the sample.

ASSESSMENT CRITERIA

<table>
<thead>
<tr>
<th>Assessment Criteria</th>
<th>Possible Points</th>
<th>Your Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basics</td>
<td></td>
<td>10</td>
</tr>
<tr>
<td>Followed instructions, including full APA citations and references</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basic requirements: uses one of the three theories indicated, includes only one</td>
<td></td>
<td></td>
</tr>
<tr>
<td>outcome variable (clearly identified), describes a treatment that is based on</td>
<td></td>
<td></td>
</tr>
<tr>
<td>the theoretical constructs of interests, states one or more hypotheses or</td>
<td></td>
<td></td>
</tr>
<tr>
<td>propositions to test, includes at least one non-experimental variable (something</td>
<td></td>
<td></td>
</tr>
<tr>
<td>that can affect the outcome, but is not part of the researcher’s hypotheses – not</td>
<td></td>
<td></td>
</tr>
<tr>
<td>demographics)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flow Chart</td>
<td></td>
<td>20</td>
</tr>
<tr>
<td>Question is explanatory, not descriptive, is based on theory, and is appropriate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>for an experimental design</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The contributions to the body of knowledge (research objectives) draw on the</td>
<td></td>
<td></td>
</tr>
<tr>
<td>strengths of experimental designs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clearly identifies and distinguishes between the theoretical population,</td>
<td></td>
<td></td>
</tr>
<tr>
<td>accessible population, and sampling frame</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sampling approach is based on replication sampling logic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Whether quantitative or qualitative, analytic procedures are appropriate for</td>
<td></td>
<td></td>
</tr>
<tr>
<td>analyzing the data that result from the experiment based on the statistical</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hypotheses or propositions, the sampling procedures, and the level of measurement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>of the variables</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incorporates data analysis techniques to test the effect of the predictor variables</td>
<td></td>
<td></td>
</tr>
<tr>
<td>on the outcome variable(s)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Establishing Direct Cause &amp; Effect</td>
<td></td>
<td>40</td>
</tr>
<tr>
<td>Question moves beyond a single pathway of cause and effect to incorporate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>complicating factors, such as additional experimental factors other than the</td>
<td></td>
<td></td>
</tr>
<tr>
<td>treatment or intervention</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The research hypotheses are not simple and obvious but rather incorporate more</td>
<td></td>
<td></td>
</tr>
<tr>
<td>complex approaches to understanding causality and using controls to establish</td>
<td></td>
<td></td>
</tr>
<tr>
<td>evidence of cause and effect</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Demonstrated an understanding of the importance of the concept of control in</td>
<td></td>
<td></td>
</tr>
<tr>
<td>experiments in the decisions made regarding design choice, sample selection and</td>
<td></td>
<td></td>
</tr>
<tr>
<td>use of statistical control of non-experimental factors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Justified the selection of the specific design proposed from the perspective of</td>
<td></td>
<td></td>
</tr>
<tr>
<td>establishing evidence of direct cause and effect and explicitly explains how the</td>
<td></td>
<td></td>
</tr>
<tr>
<td>design reduces non-experimental sources of variance in the outcome</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Incorporated sampling procedures that maximize the researcher’s ability to detect direct cause and effect relationships.

Decisions regarding the definition of the theoretical and selection of an accessible population demonstrate an understanding of the importance of replication sampling logic in experimental designs.

The analytic procedures maximize the researcher’s ability to demonstrate direct cause and effect.

**Adding to the Body of Knowledge**

Distinguishes between the specific results that the study will generate and the broader contributions to the body of knowledge that the researcher will be able to make (the conclusions).

Identified the major strengths of the study with regard to the researcher’s ability to draw conclusions about direct cause and effect.

Explained whether generalization will be theoretical or statistical identified the specific strengths of the study with regard to statistical and/or theoretical generalization.

Was able to identify specific threats to internal and external validity and explanatory power that would be of concern in the design – did not just discuss general threats common to all experiments.

Identified any steps taken to reduce the potential impact of these threats on the researcher’s ability to draw conclusions about direct cause and effect relationships and extend those conclusions to the theoretical population.

Was able to explain why the specific threats identified are of concern in this design, despite steps taken to eliminate them, and explained why the threats could not be eliminated either through alternative designs, sample selection, or statistical techniques.

Justified the sampling procedures in terms of the researcher’s ability to generalize the conclusions that s/he draws based on the research.

---

| Consulted, used, cited and referenced extensive required and additional materials about research design, and particularly true and quasi-experiments | 40 |
| Consulted, cited and referenced the literature about sampling to justify the approach selected | 40 |
| Consulted, cited and referenced the research design literature about the nature of the research question, development of hypotheses, and the use of theory in research | 40 |
| Used and referenced the literature about internal and external validity | 150 |

**Total** 150