ORLJ 5040 Final Exam Study Questions

1. Discuss the point of view that states that no worthwhile information about social behavior can be generated in a laboratory.

* This statement mainly concerns about the **external validity** of lab study. Because the lab environments (under the experimental control) can be very different from the reality, which may lead to the **systematical errors** in the observation of social behaviors. For example, experimenters may unintentionally bring errors in eliciting desired responses from subjects. Meanwhile, lab also have **limitations in incorporate complex variables**. For example, many longitudinal studies are not easily to be launched in lab.
* However, lab (experimental) studies do benefit from **random assignment, the ability to manipulate an IV, and offering maximum control**. These benefits improve the **internal validity/causality**. In terms of realism, the argument would be that even though an experiment might not have face validity (mundane realism), it might still have psychological reality (experimental realism).

1. Then discuss the point of view that states that no worthwhile information about social behavior can be generated by correlational research.

* This statement mainly concerns about **the internal validity** of correlational study. Without random assignment and manipulation on IVs, all threat towards internal validity (e.g., selection, maturity, test, and regression) are present. Meanwhile, most correlational research (in particular the cross-sectional data rather than longitudinal data) can only be able to generate the correlation instead of causality.
* However, correlational research has strong generalizability. Researchers can extrapolate results from the sample onto the population. Meanwhile, many variables can be studied at the same time. Thus, more complex study hypotheses can be studied.

1. How would you resolve these two opposing viewpoints?

* There is the “three horned dilemmas” which is the constant tension between **control over behavior, realism of context and generalizability across actors**. Lab experiments (experimental research) has the benefit of control over behavior at the expense of realism and generali­­zability. Correlational research has the benefit of generalizability at the expense of control over behavior and realism if data is gathered via survey.
* There is no perfect research method in reality. Choosing the appropriate research method is based on the research question, resource, and many other factors. However, the limitations in each research methodologies do, to some extent, encourage the usage of mixed method and reparative/longitudinal data.



a. Explain why the one-shot case study, the one group pretest‑posttest design, and the static group comparison design are not experimenting.

* Pre-experimental design (including one-shot case study, the one group pretest‑posttest design, and the static group comparison design) follow basic experimental steps but fail to include a control group. In other words, a single group is often studied but no comparison between an equivalent non-treatment group is made.
* Pre-experimental design has little or no control over the IVs. There are confounded extraneous variables that can jeopardize internal validity and researchers cannot rule out alternative explanations for reducing the systematical errors.
* One-shot Study (O X):
* A single group is studied at a single point in time after some treatment that is presumed to have caused change.
* Problem: no control over the study, no comparison group
* Threats to internal validity: History, maturation, selection, selection by mortality
* One Group Pretest-Posttest Design (O X O)
* A single case is observed at two time points, one before the treatment and one after the treatment.
* Problem: no comparison group and has confounded extraneous variables
* Threats to internal validity: History, Maturation, Testing, instrumentation, maturation by selection
* Static Group Comparison (X O) ( O)
* A group that has experienced some treatment is compared with one that has not.
* Problem: confounded extraneous variables
* Threats to internal validity: Selection, mortality, maturation by selection

b. Explain why the time‑series and non‑equivalent control group quasi-experimental designs are superior to them in terms of establishing causality.

* Quasi-experimental designs are designs that are modeled after experimental design except for the key characteristic of **random assignment** of participants into each condition.
* Through quasi-experimental design some internal validity threats are solved and consequently the causality is improved
* Time series quasi-experimental design (O1 O2 O3 O4 X O5 O6 O7 O8)
* Extension of One-group Pretest-Posttest Design
* Cons
  + Treatment effects can be examined in the context of trend and seasonality
  + solve the problem of maturation, testing, instrumentation, and maturation by selection
* Potential Threats: History
  + There are no automatic techniques to test trend and seasonality
  + It is difficult to detect gradual changes
  + Still not too much confidence can be put in causal statements, but much better than Pre-Post design
* Nonequivalent control group quasi-experimental design (O X O) (O O)
* This is a between-subjects design in which participants have not been randomly assigned to conditions.
* Extension of Static-Group Comparison
* Cons:
  + more control over the extraneous confounding variables: selection and mortality
* Potential Threats: Interaction of Selection & Maturation

c. How do true experiments differ from both the pre‑experimental and quasi‑experimental designs?

* True experiment usually occurs in the laboratory environment, where random assignment of subjects to conditions and we can manipulate the independent variable. True experiment design tries to rule out all obvious alternative explanations.

3.

a. Explain the difference between systematic and random error in experimental designs.

* Random error is caused by extraneous variables whose average influence on the outcome is the same in all conditions. Random errors usually result from events that occur in experimental settings, subject variables, or any variable not under the experimenter’s control.
* Systematic error is caused by extraneous variables whose influence on the outcome is different across conditions of the IV. Systematic error results from extraneous variables that cannot be separated from the IV. These variables can be part of the experimental setting or manipulation, or can be subject variables (differences in subjects across conditions).

b. What are the negative consequences of each?

* Random errors obscure the relationship between the IV and the DV.
* Systematic error distorts the relationship between the IV and the DV.

c. Which is more critical in terms of internal validity? How can they be reduced?

* Systematic error is more critical in terms of internal validity.
* Random error can be reduced by holding extraneous variables constant and incorporating more random subjects in every IV conditions.
* In order to reduce systematic error, we may carefully operationalize variables so that extraneous variables don’t occur with the IV or we may randomly assign subjects to conditions.

4.

a. What is the dilemma between achieving impact versus control in research?

* Control in research is to manipulate the IV and control all other confounding variables to minimize the error, and to interfere with the IV-DV causality relationship.
* Impact in research is to ensure the conceptualization of the IV is close to reality as possible, so that the study have broader scientific implications.
* The dilemma between control and impact implies that as we increase the impact of a manipulation, we are likely to decrease our control over it. The goals of impact and control are in constant tension. In order to deal with the conflict, IV should have enough impact to show up against background noise, but not so much as to have multiple meanings or alternative explanations.

b. Describe a manipulation that has high impact with some loss of control.

* For example, in order to get understand the stereotype threat and older adults’ performance on practice exercise in a library training program. We take a survey of all the older adults who attend the library training program. Ask these subjects whether they fell the stereotype threat and collect the data about their practice performance. Then analysis the relationship between stereotype threat and performance on practice exercise.

c. Describe a manipulation of the same variable with higher control but less impact.

* For the same example, we firstly collect a large group of older adults (e.g., over 300 subjects) who have the interest in attend the library training program. Then, we randomly assigned half of the older adults towards the program, which participants will obviously feel the stereotype threat. For example, stereotype information will be emphasis in the training material over and over again. The other half older adults are assigned to the program, which strictly avoid any stereotype threat. Then we analysis the difference in two groups’ average performance in the practice exercises.

5.

a. Differentiate between internal and external validity

* Internal validity is the extent to which conclusions can be drawn about the causal effects of one variable on another, that is, ensuring that only the independent variable an account for difference in the dependent variable.
* External validity is the extent to which the results of the research can be generalized across the populations and settings.
* Internal validity focus on ruling out the alternative explanations for the statement of case-and-effect relationship. While external validity measure how much the findings are generalizable to other setting, sample, and variables.

b. Discuss the factors that jeopardize each

* Factors that jeopardize internal validity:
  + Maturation: naturally occurring changes in person over time that influence posttest scores,
  + History: external events that take place during the course of research that influence the posttest score, and can be different in different group,
  + Instrumentation: difference in how score or other instrumentations are measured,
  + Regression: extreme score in the pretest tend to regress towards mean in the posttest,
  + Testing: pretest measurement influences posttest measurement,
  + Selection: initial differences between groups of research participants that may affect the dependent variable,
  + Mortality: attrition of participants (loss of subjects) is different in experimental can control group, which cannot be rule out by random assignment,
  + Interaction of Selection and Maturation: The difficulty of reaching causal conclusions because the individuals in the two groups might be growing or maturing at different rates
* Factors that jeopardize external validity:
  + No real subjects (selection biases): the participants differ substantially from the true population
  + No real condition: the participants in the study/testing behavior systematically different from their real behavior.
  + No real setting/environment: the study setting (e.g., lab) is not replicable or representative towards the real world.

c. How would you decide which of the many threats to internal and external validity to worry about in a given research investigation?

* The internal and external validity is based on the research design.
  + Check the Research methodologies: for example, the randomized experiment usually have higher internal validity. But it may loss the external validity since it requires many manipulations. While, if we use the correlation study, we may have the internal validity issue (e.g., selection)
  + Check the Research subjects: Increasing sample size usually improve both the internal validity and external validity. Because more random subjects will incorporate into the study. Additionally, we can also check whether their exist some significant alternative characteristics of population (in both control group and treatment group).
  + Check the experimenters and research process. Experimenters may unintentional bring the systematical error (e.g., instrumentation issue) towards the study, which threat the internal validity.

6.

1. Differentiate between independent, dependent, mediating, and moderating variables

* An independent variable is the variable that is changed or controlled in an experiment to test the effects on the dependent variable.
* A dependent variable is the variable being tested and measured in an experiment.
* Moderating variable answer the question about “When” or “For Whom” does the independent variable have an influence on the dependent variable. Moderators usually interact with independent variable to have an effect on the dependent variable. It can be subjective variable (e.g., age, race, and socio-economics status) or situational variable (e.g., task conditions or environmental conditions). Moderators often be introduced when a unexpectedly week relationship between independent variable and dependent variable had been found.
* Mediating variable answer the question about “How” or “Why” does the independent variable influence the dependent variable. Mediator transmits the effect of the independent variable on dependent variable and describe the internal psychological mechanism. Mediators often be introduced when there is already a strong relationship between independent variable and dependent variable.

1. Give examples of

(a) how a variable can be conceived of as both an independent and a dependent variable in separate investigations

* Take the variable ‘performance’ as an example. In one study, we investigate what makes an employee perform well. The IV is payment, the DV is performance. In another study, we want to investigate what make an employee more likely to be promoted. In this study the IV is performance, the DV is promotion.

(b) how a mediating variable can be conceived of as both an independent and dependent variable in separate investigations.

* For example, the variable motivation can be the mediating variable between the IV payment and DV performance. In other studies, motivation can also be IV, which could influence the DV engagement. Or the motivation can also be DV, which could be influence by the IV group morale.

(c) Explain how a moderating variable relates to the notion of an interaction, and provide an illustrative example.

* For example, in the study of how payment (IV) effect the performance (DV). The age may be a moderating variable since the young employees and old employees may value the payment with different importance. Young employees face the great finance pressure may care more about the payment, thus the payment effect performance more significantly. However, the old employees may care more about other factors, including teambuilding, schedule flexibility, and health care benefits.

7.

a. Consider the phenomenon of "group morale" as an independent variable.  Specify one way that this variable could be experimentally manipulated in a laboratory using ad hoc groups, one way it could be manipulated in the field using existing, intact groups, and one way it could be measured in the field.

* Ad hoc groups: researchers recruit participants who don’t know each other before and randomly assign them equally into two groups. The task is a multiple player computation game, each team are required to play 5 times. During the task, the treatment group are frequently encouraged and rewarded (the rewards is irrelevant with the computer game) by the system. As a manipulation check, the group members complete a questionnaire designed to test the group members’ overall level of enthusiasm for the group, its tasks, and goals. While participants in the control group won’t receive any feedbacks from the system during the whole process.
* Existing, intact groups: Researchers recruit a team with high group morale and take it as the treatment group and recruit strangers of equal number and assign them as the control group. Whether the team is high group morale is based on the questionnaire survey and observations on the previous activities.
* Measured in the field: participants are employees in a company which has many groups. Participants are asked to complete a questionnaire on group morale. Then researchers can measure the group morale level of each group by their answers.

b. Evaluate the adequacy of each operationalization, specifying explicitly the criteria you are using in making your evaluation.

* Ad hoc groups: it’s manipulated. Not very true to construct. The realism of this experiment is relatively low. The manipulation may cause some systematic error. Manipulation check is needed.
* Existing, intact groups: it’s manipulated. True to construct. Its realism is high. No need for manipulation check.
* Measured in the field: the internal validity is limited since this approach loss the control of study and cannot rule out the alternative explanations. The reliability be not be high because participants may lie in the survey. Field research has strength in its external validity and applicability to real-world settings and phenomena.

8.  A statistically significant correlation has been found between alcohol consumption and birth rate in the U.S. for the years 1947-1962.

a. Give as many reasons as you can why we cannot conclude that increased consumption of alcohol was responsible for increased birth rate. In doing so be sure to give alternative explanations of this statistical finding.

* Non-experimental research such as this has three major weaknesses

a) the inability to manipulate independent variables,

b) the lack of power to randomize, and

c) the risk of improper interpretation.

The danger of improper and erroneous interpretations stems from the plausibility of many explanations of complex events. Correlation does not imply causation.

* From 1947 to1962, the average income in the U.S. increased, so people can afford more alcohol. In the meantime, the increased income makes families can afford more children, so the birth rate increased.
* From 1947 to 1962, the average income in the U.S. increased, so people can afford more alcohol. In the meantime, the medical conditions in the U.S. improved, so the birth rate increased.

b. What is the appropriate term for this type of correlation?

Cross-lagged Panel Correlation (CLPC)

**9.**

1. Explain the difference between the reliability and validity of a measure.

* Reliability ask the question about “how stable and/or consistent of a test/assessment is over time?” It refers to how much **random error** is involved in our measure. We estimate reliability by calculating the internal consistency. The more items of a measure are correlated with each other, the higher is the reliability of that measure.
* Validity ask the question about “does a test/assessment measure what it is supposed to measure?” It refers to how much **system error** is involved in our measurement.
* Reliability is a necessary, but insufficient condition for validity.

1. Describe the ways that both reliability and validity can be demonstrated.

* Example: we want to measure the volume of a water-resistant rock which has irregular shape. We can use the displacement method. We find the regular graduated cylinder which at most can contain 1000ml water. We tie the rock with very thing string and then put the rock into the graduated cylinder. Then, we fill in the water till the graduate cylinder is full. We take the rock out of the graduated cylinder with the string. Then we calculate the decrease of volume as the volume of the rock.

1. Give examples of how reliability need not insure validity.

* If we use the ruler to measure the length of a pencil. But the calibration for the ruler we use is not evenly distributed which give us the wrong measure all the time. Since we always use the wrong ruler, the measurement is stable. But it is always wrong.

Reliability:

* Stability reliability:
  + test-retest reliability (e.g., administering the same test at two different times to the same group of participants)
  + parallel/alternative forms reliability (e.g., administering two different forms of the same test to the same group of participants).
* Consistency reliability:
  + split-half reliability (e.g., split the test in half and correlate the scores on one half with scores on the other)
  + internal consistency (e.g., compute the average of intercorrelations among test items that pertain to a certain construct).

Validity:

* Criterion validity: the degree to which the results of a measure corresponds to other valid measures of the same concept, (e.g., does the measure predict an outside criterion, such as the GRE predicting academic success?)
* Content validity: the degree to which the measurement covers all aspects of the concept being measured (e.g., is the content of the measure representative of the content or the universe of content of the property being measured?).
* Construct validity: the degree to which a measure adheres to existing theory and knowledge of the concept being measured (e.g., can the measure be used to legitimately make inferences about theoretical constructs on which your measure is based?)
  + convergence (whether your measure is related to other measures of your construct, as well as to measures of related constructs)
  + discriminability (the degree to which your measure is unrelated to measures of your constructs that are unrelated to your construct)

**10.**

1. What are the major advantages and disadvantages associated with correlational research?

* Advantages:
  + Allows us to study the world without trying to alter it
  + Generalizability
  + realism/impact of variables is high
  + interest in individual differences
  + study of many variables at the same time
  + permits more variation in variables of interest
  + practicality
* Disadvantage:
  + lack of control over IV;
  + No random assignment;
  + all threats to internal vanity are present
  + Correlation does not indicate causation

1. What steps can be taken to increase the likelihood that correlational research will approach causality?

* Establish association:
  + Measure the linear relationship between two variables with wider range of variation
* Establish time priority:
  + Longitudinal studies which IV is measured before the DV
  + Cross-lagged panel correlation data, which offer the autocorrelations, synchronous correlation which will offer autocorrelations (how stable variables over time, and indicates retest-reliability for stable variables) and synchronous correlations (if both synchronous correlations are equal, the relationship will hold over time); although, it won’t wholly solve the time priority problem, it still provides useful insight.
* Eliminate spurious relationships:
  + Determine the partial correlation in the multiple regression, which the effect of IV on DV is condition on all the other variables. Consequently, we can control all the significant but not interesting variables
  + Determine partial explanations can help clarify;
  + Increasing the sample size to make more reliable estimation of the effects.

**11.**

1. What do you see as the major ethical issues concerning researchers of social behavior in psychology?

* Limitation of free choice
* Deception
  + Withhold information
  + Providing incorrect information
* Harming
  + Diminished self-respect
  + Exposing people to stress
  + Withhold the benefits from the control group
* Mistreating participants
  + False promise
  + Breaches of confidentiality

1. Provide specific examples and rules of thumb that you would use in discriminating between ethical and unethical approaches to acquiring knowledge about human behavior.

* Rules of thumb:
  + Respect for participates
  + Beneficence: the magnitude of potential harm outweighs the benefits
  + Justice: used fair procedures and outcomes in the selection of research subjects
* For example, Cambridge Analytica Ltd (CA), as a political consulting firm, acquired and used personal data about Facebook users from an external researcher who had told Facebook he was collecting it for academic purposes. The personal data of up to 87 million Facebook users were acquired via the 270,000 Facebook users who used a Facebook app called "This Is Your Digital Life."

**12.**

1. What is the difference between a within‑subjects design and a between‑subjects design?

* Between-subjects study design: different people test each condition, so that each person is only exposed to a single condition. Each subject is exposed to a single condition, and the means of different group (condition) are compared
* Within-subjects (or repeated-measures) study design: the same person tests all the conditions, and the individual differences are compared. Each subject is repeatedly treated and tested, and the variations cased by different treatments appear with the same person rather than between different group of people.

1. What are the major advantages and disadvantages of each design?

* The major advantage of within-subject design:
  + shorter session and less time per participant;
  + less change the participants will experience fatigue;
  + avoid the carry-over effects
  + simper to design and operate
* The major disadvantage of within-subject design:
  + Require a lot of participants in each group to gain the power;
  + Differences between subjects within a given conditions introduces errors (variance within the group)
  + Not sensitive to smaller difference since we take the average
* The major advantage of between-subject design:
  + Require fewer subjects to achieve power;
  + Higher sensitivity and precision;
  + More control over extraneous variables since each subject serves as its own control (baseline);
  + More statistical powerful since it goes beyond the average to the change in each individual difference;
  + Less random error
* The major disadvantage of between-subject design:
  + Contamination/order/carry-over effects;
  + More fatigue from taking part in too many different conditions;
  + Practice effect or sensitization
  + More complicated to design and need extra effort to prevent contamination or order effects;
  + Only experimental/manipulated variable can be used instead of individual difference variable (e.g., age and gender).

1. Are there times when a within‑subjects design is preferable? Give examples.

* Anytime effects or differences may be subtle and difficult to detect when averaged across groups. Because it would sensitive enough to pick up smaller differences in individual. For example, when solving puzzles, differences between individual solving times might only be around one minute within a certain group. However, this difference might not be detectable if it is averaged across the group using a between-subjects design.
* Track an effect over time: If the research question is how a skill develops after a certain number of practice sessions, a within-subjects design would be necessary to capture the changes in participants’ skill at multiple time points
* When the sample size is small but there are a lot of factors (variables) to include.

13.

1. How do social scientists conceive of causality?

* There are 3 criteria generally considered as requirements for identifying a causal effect: empirical association (i.e. independent and dependent variables must vary together), the temporal priority of independent variable (i.e., the variation in the independent variable comes before variation in the dependent variable), and most importantly non-spuriousness (i.e., all other possible influences/explanation on dependent variable should be eliminated).
* Evidence that meets the other two criteria — identifying a causal mechanism, and specifying the context in which the effect occurs — can considerably strengthen causal explanations.

1. What circumstances do they look for, or what do they try to establish through their observations when trying to demonstrate causality?

* The sort of control that is necessary for a randomized experiment is most easily achieved in laboratory settings, where extraneous variables can be controlled so only the variable of interest changes across conditions. Randomized experiments can also be conducted in real-world settings. However, real-world or field experiments, although often stronger in external validity, tend to be weaker in internal validity because it is usually more difficult to control the effects of extraneous factors.
* They are trying to show that the differences in the independent variable lead to differences in the dependent variable while eliminating all other possible influences on the dependent variable.

1. How do experiments provide the kind of data necessary for establishing causality?

Experimental research provides the most powerful design for testing causal hypotheses because it allows us to confidently establish the first three criteria for causality — association, time order, and non-spuriousness. There are at least three elements of experiments:

1. Two comparison groups (e.g., control group and treatment group) to establish the empirical association. Randomized experiments allow researchers to decide which participants are assigned to which levels of the variable that is believed to be the causal variable.
2. Variable in the independent variable before assessment of change in the dependent variable, to establish time order.
3. Random assignment to the two (or more) comparison groups, to establish nonspuriousness. Randomized experiments, in which individuals are randomly assigned to levels of the independent variable, also allow the researcher to minimize the effects of extraneous variables that might otherwise be confounded with the causal variable of interest.

Frequently, this control, although maximizing internal validity, compromises construct and external validity

14.

1. Name of research design: Quasi-experimental design (Static-Group Comparison Design)
2. Threats to internal validity and ability of the researcher to make causal statements

* Selection: there may have a preexisting difference between the 100 girls who have used the program and 100 girls who have not.
* Selection by maturation: We might suspect that the people who became interested in program energy would naturally spend more time in arithmetic. They might have naturally higher competence in arithmetic
* We cannot guarantee that all the girls participate in the program indeed enjoy the activity. Some girls who attend the program may dislike the video game in the program.  At most, this quasi-experiment can show the relationship between this specific program and competence in arithmetic. But we cannot tell whether girls view arithmetic as an enjoyable activity that could improve their competence in arithmetic.
* Data from any static-group comparison design can be used to establish that two variables’ correlational relationship. Ambiguity in the interpretation of the nature of the association between variables, however, is a reason for researchers to exercise caution in making causal inferences from this sort of research design. If our purposes are limited to assessing the degree of covariation among variables, a static-group comparison design is adequate for the task.

1. How to redesign the study

* Pretest-posttest nonequivalent control group design. All girls in this study are given both pretest and posttest in their competence in arithmetic. There are two advantages of this design over the static-group comparison design. First, this design helps us argue that the independent variable is responsible for variation in the dependent variable, rather than the other way around. Second, we measure (rather than randomize away) preexisting differences between the groups on the dependent variable, which can help us argue against the selection threat to internal validity.

15.

1. Name of research design: Quasi-experimental design (Pretest-posttest nonequivalent control group design)
2. Threats to internal validity and ability of the researcher to make causal statements

* Selection: the student in this study is from different universities: a leading engineering school and a state teachers’ college. For example, the teachers from engineering school may have a stronger background and teaching experience than the teacher in the state teacher’s college for the same course. This may explain students’ different performances in the posttest.
* Selection by maturation: though their pretest is at a similar level, their interest, ability, and engagement in learning statistics may be very different. Students from leading engineering schools may feel that statistics is very important and directly useful for their engineering-related study or career. Consequently, they spend more time studying statistics and pay more attention during the class. However, students from teacher’s college may feel that learning statistics is hard to directly improve their practice in their future career (as a teacher or educational researcher). Consequently, they spend less time. This may explain their difference in the posttest.
* History: it is not sure whether there are some external influences happened differently in these two school. For example, the engineering school in this year plans to give the best students in this class some research assistant positions. This may lead students more interested in this class.

1. How to redesign the study

* I will advise the researcher to ask the students from the same school (e.g., the leading engineering school) to participate in this study. The same class from the same teacher will be opened in two sections (both in the morning).  One class section will use Cookbook Method, the other class Theory Method. The only difference between these two sections is the textbook suggested. All the other materials in the class (slides, exams, and homework) are the same. The slide is not designed to be highly linked with the books. Students are recommended to read the textbook in every class.