Notes created by Gerard Chung, gerard@nus.edu.sg

Quantitative and Qualitative methodology

Debate on what is really real?

- The debate boils down this: What is really real? Does the researcher view the research phenomenon as something that can be observed objectively because there is an external objective reality? Or does the researcher think that reality cannot be objectively observed because everyone brings their subjective "lens" (the researcher also brings her/his own lens) to interpret the phenomenon?
- A quantitative researcher would look for rules or patterns among large groups of people that are generalizable and representative; A qualitative researcher would want to know how each individual interpret the situation because how each individual make sense and meaning and getting different (even unique) points of views is important.

Contemporary positivism: No universal laws; instead we replicate and we revise.

- In Social science, three key paradigms organize how social science researchers organize their observations and make sense of reality: Positivism, interpretivism, and critical social science.
- Positivism paradigm in Quantitative research:
    - Like physical and biological science, we can study society using our five senses.
    - Society can be studied logically and rationally.
    - There are universal laws or rules.
    - However, contemporary positivists recognize the limitations of positivism in social science research.
    - Contemporary positivists recognize that researchers are not as objective because their personal feelings an influence their research and universal laws are elusive in social science research.
    - To reduce these limitations, positivist social scientists ensure their research methods are highly structured (e.g, randomized designs, causality); they do not necessary seek universal laws but seek to replicate their results by doing further testings or revise their findings. Objective truth can be achieved over the long run with further testings, replications of results, and revisions to theories.

Theory and its place in social science scientific methods

- Scientific methods in social science
    - Two types of scientific methods in social science: deductive or inductive
    - The two types differ depending on how they use theory
- Quant methodology in social science research uses the deductive scientific method
- Note that statistics (thus, probability) comes in at the hypothesis stage because in reality, the observations we make (e.g., will higher parenting stress ALWAYS lead to increased maladaptive parenting?) seldom 100% matches our hypothesis (expectations). Probability becomes an important part of the scientific method to help us decide if we have made an important finding or not.

Notes created by Gerard Chung, gerard@nus.edu.sg

Conceptualization and Operationalization

Moving fr*om* the vague to the specifics

- This stage comes only after developing the research question.
- And the key task here is to move from vague ideas to specific definitions and ways to measure what you want to study
    - There are two consecutive steps:
    - Conceptualization involves refining and specifying abstract concepts
    - Operationalization involves developing specific research procedures that allow you to observe the construct of interest.
    - An example: I would like to understanding self-efficacy in the parenting domain. Conceptualization lead me to use the definition of parenting self-efficacy rather than the global self-efficacy. This definition is based/influenced by Bandura's theory of self-efficacy. Operationalization leads to decide on the specific scale to use and how I intend to obtain the data. Given that I intend to study parents with newborn babies. I chose PSE scale that was developed for newborn parents.
- This is mostly a quantitative thing to do because quant aims to be precise and specific and believes that external reality can be observed and measured. Nonetheless, even Quali begin with some initial definitions or anticipated meanings of the construct.
- Concept -> Variable [operationally, this is what the process looks like]
    - Concept is a mental image of an idea or the word we agreed that represents the idea (confidence, self-esteem etc)
    - Concept can have attributes. Gender is a concept with male, female as two possible attributes.
    - Concept is converted into a variable when we develop ways to observe the attributes of the variable.

Being transparent and precise about how we define the concept.

- In everyday life, we use words (e.g., she is so *smart*, they are so *confident* people, I am *stressed*, client is *resistant* to our help) that we generally know what it means, but in precise terms, we can have different agreements on its precise definitions.
    - This is why conceptualization is important to state in precise terms what we mean when we use the term (what does confident means?). We can then debate on the appropriateness of the definition we chose.
- Operationalization translates the variable into observable terms using indicators.
    - Koi is more popular than Liho: "Popular" = "# of sales per day" indicator
    - Depression: CESD scale with 26 indicators
- Obviously, how we operationally define (ie. What indicators we use) will influence the data we collect and consequently our findings.
    - Koi is more popular than Liho:
        - "Popular" = "# of sales per day" indicator?
        - # of Facebook followers? Length of store queue? Demographics of followers?
    - Depression
        - "Do you feel sad of the time?" vs. "Do you have poor sleep?"

Notes created by Gerard Chung, gerard@nus.edu.sg

Measurement

Minimize the likelihood of measurement errors

- Measurement is the process of thinking about how to translate the abstract concepts so that we can observe them consistently and accurately
- Conceptualization proceeds differently in qualitative research compared to quantitative research. Since qualitative researchers are interested in the understandings and experiences of their participants, it is less important for them to find one fixed definition for a concept before starting to interview or interact with participants. The researcher's job is to accurately and completely represent how their participants understand a concept, not to test their own definition of that concept.
    - If you were conducting qualitative research on masculinity, you would likely consult previous literature. From your literature review, you may come up with a *working definition* for the terms you plan to use in your study, which can change over the course of the investigation. However, the definition that matters is the definition that your participants share during data collection. A working definition is merely a place to start, and researchers should take care not to think it is the only or best definition out there.
- Operationalization helps to understand how we are going to measure and observe our data.
    - Are we going to use standardized scales?
    - Is it going to be self-reported by the respondents? Or direct observations? Or we are going to use administrative data?
- No matter how we are going to do it, we must remember that no measurement process or measure type is going to be perfect.
- We have to be prepared for every possible sources of measurement error.
- We must carefully plan to minimize the likelihood of measurement error by using or designing quality measurements.
- Quality measurements would be high in both reliability and in validity

Quality of measurement: Validity & Reliability

- Reliability refers to the consistency of the measurement
    - A weighing machine that spits out different answers for a person's weight is not reliable
- Validity refers to the accuracy of the measurement
    - A weighing machine that gives the same weight each time I use is reliable. But if it consistently under-estimate my weight, it will not be a valid measurement,
- True score theory tells us that an observed score is made up of two components: the true score and the error
    - i.e., Observed score = True score + Error
        - The true score is the person's true level or ability. But because of error, we do not see the true score but only the observed score.
    - We aim to make the error component as small as possible
    - Actually, the error component is also made up of two components: random error and systematic error
        - Random errors are factors that randomly affect the measurement. For e.g., a person's mood

    - Systematic errors are consistent errors. For e.g., interviewers who use consistently different interviewing style; a noisy environment (e.g., loud traffic)

Reducing errors or keep it to the minimal

- Pilot test your instrument
- Train the persons who will administer the measures
- Double-check the data e.g., errors in coding, extraction etc
- Use statistical measures to account for errors (e.g., structural equation modeling)
- Triangulate your measurements (e.g., use multiple measures of the same construct)

Types of reliability and validity

- The below are some ways to estimate the reliability in a measure.
    - Inter-rater reliability
    - Test-retest reliability
    - Parallel-forms reliability
    - Internal consistency reliability
- Types of validity
    - Construct validity
    - Content validity
    - Criterion-related validity
    - Concurrent validity

Other technical specifics of measurements

- Levels of measurements
    - A variable or the data can be measured in different ways:
        - Gender can have two levels or attributes: Female or Male -> this is nominal
        - Ranking of an outcome: Grades A, B, C… or 1^{st}, 2^{nd}, 3^{rd} -> this is ordinal
        - Data that is continuous: Height, age in years, weight, self-esteem scores on Rosenberg's self esteem scale, number of counseling attended,
- There are four types of measurements:
    - Discrete:
        - Nominal type (e.g., Gender, Race, Marital status)
        - Ordinal type (Rankings).
            - With nominal, you can say that female is more than male
            - With Ordinal, you can say that the attributes are relatively more or less. But you cannot say that 1^{st} is twice more/better than 2^{nd}
    - Continuous:
        - Ratio type has a true zero (age has no negative values)
        - Interval (has no true zero)
- The level of measurement will affect the choice of statistics to be used.

- For e.g., we cannot use mean to describe data that is nominal. The average sex is Male (WRONG). Instead we use percentage or frequency. E.g., The majority of persons are male (80%)
- We use mean for continuous data. E.g., the mean age of respondents is 38.73 years old.

Hypothesis and Hypothesis Testing

Hypothesis: a statement about the relationship between variables

- Hypothesis are statements that predict what we expect to find in the relationships between variables. They are actually tentative statements. They can be rejected and revised.
- Specifically, hypothesis states the relationships between an independent variable and a dependent variable
- Independent variable (also called a predictor sometimes) is postulated to explain another variable. Dependent variable (sometimes called outcome variable) is the variable being explained.
- Hypothesis, like research questions, should be clear, specific, and testable.
  - "Koi bubble tea is not popular" is not a good hypothesis. It is value-laden, not testable, not specific.
  - "Koi bubble tea is less popular than Liho as measured by the volume of sales per hour in a day"

Statistical Hypothesis Testing: Putting our hypothesis to the test!

- So a hypothesis is our statement of the expected relationship between two variables. For e.g., Single-parents compared to dual-parent families experience higher parenting stress
- To test if this statement is true, we can use:
  - descriptive statistics to find out the average scores for parenting stress among the two groups of parents. Now, let's say we find that the mean score of stress among single-parent families is higher. However, this difference in scores may just be a chance difference; it is not due to the effects of single-parenthood variable.
    - "By luck one lah" This is like what friends would say to us when we do something amazing. For instance, throwing paper balls into the trashcan from a distance.
    - To prove to our nay-sayers that we really do have the skill – that it was not *by chance* – we repeat the trick. What we are hoping to prove is this: that despite the assumption that we do not have the trick, we successfully did it even though the probability of us being successful is very small.
    - This analogy underlies the concept of statistical hypothesis testing:
      - Assuming that I do not have the skill, I was
      - Assuming that there is no effect (e.g., no differences in stress between the two groups), we found a difference in our sample where the probability of achieving this is small (usually we define small as $p < 0.05$).
  - Statistical hypothesis testing

        - We use statistical hypothesis testing to determine the possibility of obtaining a chance difference (e.g., or an effect) as large as the actual difference found in our study. If the probability is small enough, we decide that the observed difference is not a chance difference. Rather, the observed difference is attributed to the effect of the independent variable
        - This is mind-boggling to understand and will give you a headache and a nightmare to understand.
        - Imagine if someone states that there are pink elephants in the world. We know that this is not possible. Even if possible, this would be a very very very small tiny chance (we cannot be 100% sure because we theoretically have not seen every elephant in the world). But imagine in the moment when you step out of the room, a pink elephant passes by. You would then have to accept that your original hypothesis that there is no pink elephant has to be rejected.
- Note how easily hypothesis can be rejected but how impossible it can be to prove it. This is the idea of falsifiability.
    - For instance, our initial hypothesis that there are no pink elephants in the world can be proven wrong simply because we have seen one pink elephant – based on a single evidence contrary to our hypothesis.
    - Yet, note how the initial hypothesis (elephants cannot be pink) cannot be proven correct. Why? because in order to prove it, we need to search every single elephant in the world from the past to the present.
    - This idea is central to science and is known as Falsifiability:
        - It is easy to reject a statement but hard to prove a statement.
        - A statement can be proven wrong but cannot be proven right
        - "Swans can only be white" -> to prove this, we need to search every single swan. But to prove this is wrong, we only need to find a swan that is not white.
        - "There is no effect between X and Y". To prove this, we need to conduct every possible experiment. But to prove this hypothesis is wrong, we need only find one experiment that finds a relationship between X and Y.

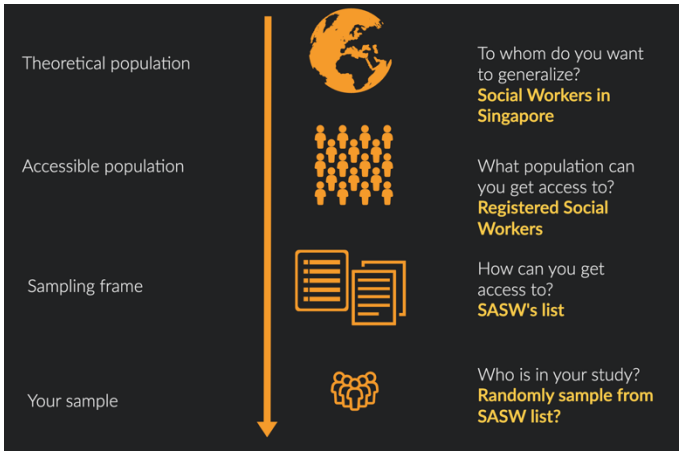
Sampling

Sampling is a multistep process

- The process of selecting units from the population of interest.
- The logic of sampling is that we want characteristics of our sample to represent the population. People are all different. If everyone is the same, we do not need to sample. Just 1 person is enough for our research. But because there is variations in the population, we want our sample to also contain the same variations that exist in the population.
- Sampling is a multistep process and that you can go wrong in many places. From each step to the next, the decision we make or the limitations we face will introduce bias (systematic error) to our research.

Notes created by Gerard Chung, gerard@nus.edu.sg

- In this multistep (see figure below), we first identify the theoretical population and then accessible population. Even if we can correctly identify our theoretical population, limitations in accessing this population will mean that our sampling frame will not be complete or accurate. Furthermore, our respondents in our final sample may not turn up for our study.



- Why does all this matter? Because we want to generalize our results from our sample to the population.
    - This is known as external validity.

Sampling Methods

- Nonprobability and probability methods
- In probability methods, individuals are selected for the sample in such a way that each individual **from the population** has a known or equal probability of being included in the sample i.e., random selection
    - For e.g., the researcher draws numbers from a bag to select the participants for the sample. The choice of the participant is determined by chance – not based on participants' characteristics or researcher's judgement (or bias).
    - Imagine if I choose only participants who are highly motivated to be in my intervention group. If my study shows results that are effective, does it mean that my intervention is effective for all parents? Or only motivated parents?
    - Probability methods include simple random sampling and stratified random sampling
- In non-probability, individuals are selected for the sample based on the researcher judgement (non-random selection)
    - Nonprobability methods include convenience sampling, purposive sampling, and snowball sampling
    - Convenience sampling rely on whoever is available to you. If you want to know people's attitudes on LGBT, you survey those at the mrt station because they are most accessible and you get access to a large group of participants. But would survey findings from participants from Bukit Batok mrt be representative of the wider population?
    - Purposive sampling relies on your judgement or your expertise to sample those you want for your study. For e.g., I need parents and I know that the Centre for Fathering can give me access to this group. Another e.g., would be doing research with homeless people. This is a hard-to-reach population and researchers have to depend on their own

community knowledge to access the locations in which homeless people would be available

- Do we always need to use probability sampling? Nope. Sometimes it is impossible or inappropriate
    - In social services research, do we always have a list of the population of research subjects to randomly select from?
    - For e.g., do I have a list of the population of homeless people in Singapore to randomly select?

Probability theory in Sampling

- Why random sampling is better? What is the logic?
- First, if participants are selected by chance, and not because of biases, it is clear that this is better for forming a sample that is representative of your target population and for generalizing results.
- But for a second reason, random sampling is better because of probability theory.
- Probability theory tells us that with many many infinite random samples drawn, the distribution of means of the variable from all these infinite samples will form a normal distribution! Yes, believe it or not (actually this is known as the Central Limit Theorem)
- Probability theory then gives us a formula to calculate the error of this overall mean from all the infinite sampling mean (this is actually known as the sampling error). In other words, we can now quantify the errors.
- Our demonstration with 50 samplings of 4 poker cards with replacement show that:
    - With only one sample (even if randomly sampled), the sample may not be representative. But with 50 repeated samples, we see that the distribution of number of reds began to resemble the characteristics of an actual deck of poker cards (ie. Closer to poker cards (50% or two red cards)
    - If we had resample even more (up to 100,000 times), the distribution would be even more representative with less variability around 50% (or two reds).
    - If we also had sample 10 cards instead of 4 cars (i.e. bigger sample size), this would further improve our sample's representation of the population
    - Conclusions: 1) Use bigger samples improves representation; 2) Probability theory works by sampling many many times.

Research Designs

The blueprint

- The research design refers to the overall strategy that you choose to integrate the different components of the study in a coherent and logical way, thereby, ensuring you will effectively address the research problem;
- Research design provides the glue that holds various elements in a research project together.
- it constitutes the blueprint for the collection, measurement, and analysis of data.
- Note that the research problem determines the type of design you should use, not the other way around!

The Research Question Guides the Design – not the other way!

Notes created by Gerard Chung, gerard@nus.edu.sg

- The function of a research design is to ensure that the evidence obtained enables you to effectively address the research problem logically and as unambiguously as possible.
- If the research question is to evaluate a program, the suitable research design would be one that helps you estimate treatment effects accurately by comparing your treatment to a comparison group

Types of research designs

- It would be impossible to list and describe every possible research design. I am just going to describe a few
- **Case-study design**
    - **Definition and Purpose**
        - A case study is an in-depth study of a particular research problem rather than a sweeping statistical survey or comprehensive comparative inquiry. It is often used to narrow down a very broad field of research into one or a few easily researchable examples.
        - The case study research design is also useful for testing whether a specific theory and model actually applies to phenomena in the real world.
        - It is a useful design when not much is known about an issue or phenomenon.
    - **What do these studies tell you?**
        - Approach excels at bringing us to an understanding of a complex issue through detailed contextual analysis of a limited number of events or conditions and their relationships.
        - A researcher using a case study design can apply a variety of methodologies and rely on a variety of sources to investigate a research problem.
        - Design can extend experience or add strength to what is already known through previous research.
        - Social scientists, in particular, make wide use of this research design to examine contemporary real-life situations and provide the basis for the application of concepts and theories and the extension of methodologies.
        - The design can provide detailed descriptions of specific and rare cases.
    - **What these studies don't tell you?**
        - A single or small number of cases offers little basis for establishing reliability or to generalize the findings to a wider population of people, places, or things.
        - Intense exposure to the study of a case may bias a researcher's interpretation of the findings.
        - Design does not facilitate assessment of cause and effect relationships.
        - Vital information may be missing, making the case hard to interpret.
        - The case may not be representative or typical of the larger problem being investigated.
        - If the criteria for selecting a case is because it represents a very unusual or unique phenomenon or problem for study, then your interpretation of the findings can only apply to that particular case.
- **Cohort Design**
    - **Definition and Purpose**

        - A cohort study generally refers to a study conducted over a period of time involving members of a population which the subject or representative member comes from, and who are united by some commonality or similarity.
        - Using a quantitative framework, a cohort study makes note of statistical occurrence within a specialized subgroup, united by same or similar characteristics that are relevant to the research problem being investigated, rather than studying statistical occurrence within the general population. Using a qualitative framework, cohort studies generally gather data using methods of observation.
    - **What do these studies tell you?**
        - The use of cohorts is often mandatory because a randomized control study may be unethical. For example, you cannot deliberately assign children to child abuse and no child abuse, you can only study its effects on those who have already been exposed to child abuse. Research that measures risk factors often relies upon cohort designs.
        - Because cohort studies measure potential causes before the outcome has occurred, they can demonstrate that these "causes" preceded the outcome, thereby avoiding the debate as to which is the cause and which is the effect.
        - Cohort analysis is highly flexible and can provide insight into effects over time and related to a variety of different types of changes [e.g., social, cultural, political, economic, etc.].
    - **What these studies don't tell you?**
        - In cases where a comparative analysis of two cohorts is made, a researcher cannot control for all other factors that might differ between the two groups. These factors are known as confounding variables.
        - Cohort studies can end up taking a long time to complete if the researcher must wait for the conditions of interest to develop within the group. This also increases the chance that key variables change during the course of the study, potentially impacting the validity of the findings.
        - Due to the lack of randominization in the cohort design, its external validity is lower than that of study designs where the researcher randomly assigns participants.
- **Cross-Sectional Design**
    - **Definition and Purpose**
        - Cross-sectional research designs have three distinctive features: no time dimension; a reliance on existing differences rather than change following intervention; and, groups are selected based on existing differences rather than random allocation.
        - The cross-sectional design can only measure differences between or from among a variety of people, subjects, or phenomena rather than a process of change. As such, researchers using this design can only employ a relatively passive approach to making causal inferences based on findings.
    - **What do these studies tell you?**

        - Cross-sectional studies provide a clear 'snapshot' of the outcome and the characteristics associated with it, at a specific point in time.
        - Unlike an experimental design, where there is an active intervention by the researcher to produce and measure change or to create differences, cross-sectional designs focus on studying and drawing inferences from existing differences between people, subjects, or phenomena.
        - Entails collecting data at and concerning one point in time. While longitudinal studies involve taking multiple measures over an extended period of time, cross-sectional research is focused on finding relationships between variables at one moment in time.
        - Groups identified for study are purposely selected based upon existing differences in the sample rather than seeking random sampling.
        - Because cross-sectional designs generally use survey techniques to gather data, they are relatively inexpensive and take up little time to conduct.
    - **What these studies don't tell you?**
        - Results are static and time bound and, therefore, give no indication of a sequence of events or reveal historical or temporal contexts.
        - Studies cannot be utilized to establish cause and effect relationships.
        - This design only provides a snapshot of analysis so there is always the possibility that a study could have differing results if another time-frame had been chosen.
        - There is no follow up to the findings.
- **Experimental Design**
    - **<u>Definition and Purpose</u>**
        - A blueprint of the procedure that enables the researcher to maintain control over all factors that may affect the result of an experiment. In doing this, the researcher attempts to determine or predict what may occur.
        - Experimental research is often used where there is time priority in a causal relationship (cause precedes effect), there is consistency in a causal relationship (a cause will always lead to the same effect), and the magnitude of the correlation is great.
        - The classic experimental design specifies an experimental group and a control group. The independent variable is administered to the experimental group and not to the control group, and both groups are measured on the same dependent variable. Subsequent experimental designs have used more groups and more measurements over longer periods. True experiments must have control, randomization, and manipulation.
    - **<u>What do these studies tell you</u>?**
        - Experimental research allows the researcher to control the situation. In so doing, it allows researchers to answer the question, “What causes something to occur?”
        - Permits the researcher to identify cause and effect relationships between variables and to distinguish placebo effects from treatment effects.

        - Experimental research designs support the ability to limit alternative explanations and to infer direct causal relationships in the study.
        - Approach provides the highest level of evidence for single studies.
    - **What these studies don't tell you?**
        - The design is artificial, and results may not generalize well to the real world.
        - The artificial settings of experiments may alter the behaviors or responses of participants.
        - Experimental designs can be costly if special equipment or facilities are needed.
        - Some research problems cannot be studied using an experiment because of ethical or technical reasons.
        - Difficult to apply ethnographic and other qualitative methods to experimentally designed studies.
- **Observational Design (non-experimental design)**
    - **Definition and Purpose**
        - This type of research design draws a conclusion by comparing subjects against a control group, in cases where the researcher has no control over the experiment.
        - There are two general types of observational designs. In direct observations, people know that you are watching them. Unobtrusive measures involve any method for studying behavior where individuals do not know they are being observed.
        - An observational study allows a useful insight into a phenomenon and avoids the ethical and practical difficulties of setting up a large and cumbersome research project.
    - **What do these studies tell you?**
        - Observational studies are usually flexible and do not necessarily need to be structured around a hypothesis about what you expect to observe [data is emergent rather than pre-existing].
        - The researcher is able to collect in-depth information about a particular behavior.
        - Can reveal interrelationships among multifaceted dimensions of group interactions.
        - You can generalize your results to real life situations.
        - Observational research is useful for discovering what variables may be important before applying other methods like experiments.
        - Observation research designs account for the complexity of group behaviors.
    - **What these studies don't tell you?**
        - Reliability of data is low because seeing behaviors occur over and over again may be a time consuming task and are difficult to replicate.
        - In observational research, findings may only reflect a unique sample population and, thus, cannot be generalized to other groups.
        - There can be problems with bias as the researcher may only "see what they want to see."

        - There is no possibility to determine "cause and effect" relationships since nothing is manipulated.
    - Sources or subjects may not all be equally credible.
    - Any group that is knowingly studied is altered to some degree by the presence of the researcher, therefore, potentially skewing any data collected.