Research

One definition:

"creative and systematic work undertaken to increase the stock of knowledge, including knowledge of humans, culture and society, and the use of this stock of knowledge to devise new applications"

Do you think this is a good definition? Why? Why not?

QUALITATIVE METHODS

Why even do qualitative research?

https://www.flickr.com/photos/cesarastudillo/2815083396/

Why even do qualitative research?

Every study methodology or study design has particular limitations and capabilities. This is particularly in the context of:

- how generalizable the results are;
- how much you can control for specific variables;
- how much the results reflect the real world.

A metaphor: the streetlamp/spotlight

https://www.flickr.com/photos/darwinbell/2920366009/

A metaphor: the streetlamp/spotlight

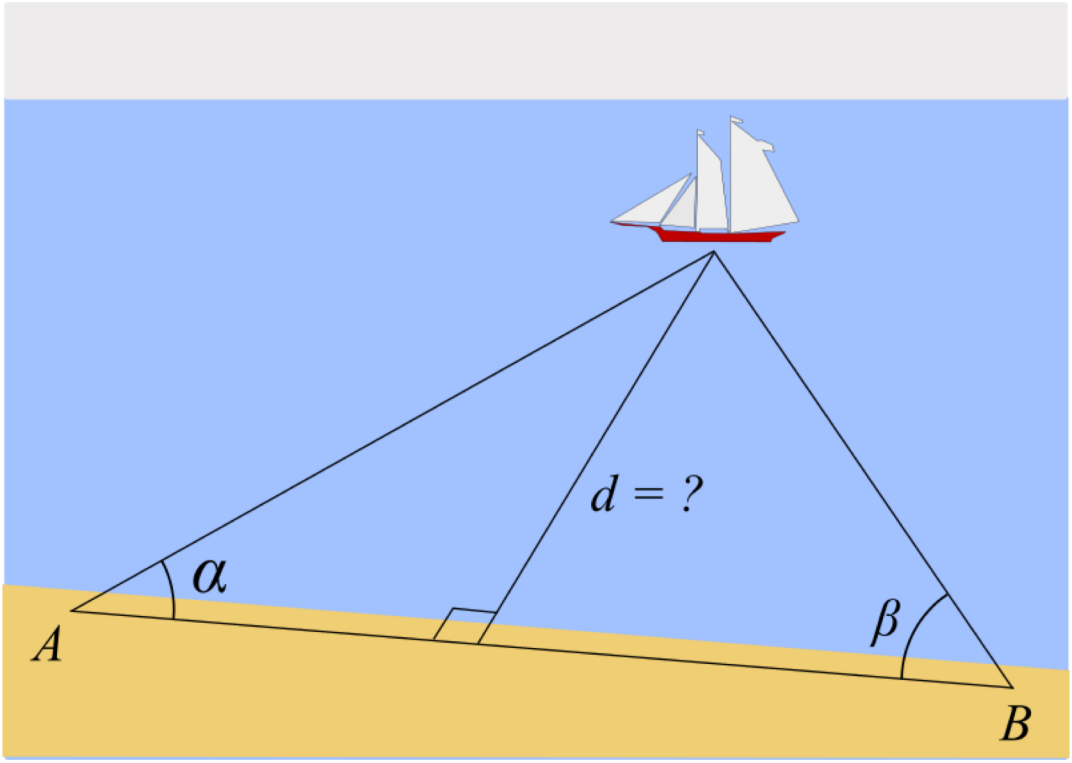
- **Quantitative studies**: measure things that are *easily instrumented* (i.e., located within the light).

- Sometimes the things you want to learn are *outside the spotlight* (i.e., in the dark) and not easily instrumented or well understood. **Qualitative studies** to the rescue!

Mixed Methods!

"Multimethodology or multimethod research includes the use of **more than one method of data collection or research** in a research study or set of related studies. **Mixed methods research** is more specific in that it includes the **mixing of qualitative and quantitative data, methods, methodologies, and/or paradigms** in a research study or set of related studies."

Triangulation!



Triangulation!

In the social sciences, **triangulation** is often used to indicate that **two (or more) methods are used in a study** [or set of studies] in order to check the results of one and the same subject…The idea is that one can be **more confident with a result if different methods lead to the same result**.

Triangulation is a powerful technique that facilitates **validation of data through cross verification from two or more sources**…

By **combining multiple observers, theories, methods, and empirical materials**, researchers hope to **overcome the weakness or intrinsic biases and the problems that come from single method, single-observer and single-theory studies.**

Important Research Quality Terms!

- **Reliability**

- **Validity**
  - **Internal Validity**
  - **External Validity**
  - **Ecological Validity**

Reliability

What evidence do you have that, if you did what you did again, you would get the same results?

Reliability

What evidence do you have that, if you did what you did again, you would get the same results?

In this course we talk *some* about how qualitative methods approach this, but we only dip our toes (a lot of people have spent a long time thinking about this!). You might think of this as more a strength of quantitative methods, though.

Validity

How “real” are the results that you’re getting?

Internal Validity

External/Ecological Validity

Internal Validity

Are your claimed results supported by your study, or are there other confounding factors?

Example: Internal Validity

A (quantitative) internal validity example: I am testing whether people are happier about filing paperwork when they are given breaks. I assign all participants randomly to either condition A (no break between the two assigned filing tasks) or condition B (a break between filing tasks). During the break participants can wander around the room, go to the bathroom, or help themselves to cookies and milk (supplied as a courtesy). At the end of the study, I give participants a (previously-validated) survey that measures their level of satisfaction with the filing tasks. I find that participants assigned to condition B express significantly higher levels of satisfaction than those in condition A. I therefore conclude that giving breaks leads to higher satisfaction.

Does this conclusion seem correct?

A (quantitative) internal validity example: I am testing whether people are happier about filing paperwork when they are given breaks. I assign all participants randomly to either condition A (no break between the two assigned filing tasks) or condition B (a break between filing tasks). During the break participants can wander around the room, go to the bathroom, or help themselves to cookies and milk (supplied as a courtesy). At the end of the study, I give participants a (previously-validated) survey that measures their level of satisfaction with the filing tasks. I find that participants assigned to condition B express significantly higher levels of satisfaction than those in condition A. I therefore conclude that giving breaks leads to higher satisfaction.

Example: Internal Validity

In reality, any number of things could be contributing to this finding. For example, maybe people are just happier when you give them cookies and milk. Or maybe most of your participants in condition B chose to walk around, thereby raising their heart rates, and the residual effects made them more satisfied. Maybe you would get the same results if you gave them no break, but force fed them milk-soaked cookies as they filed while walking on a treadmill.

This study has **issues with internal validity**, given that there are many other confounding factors that could have contributed to the measured results.

External/Ecological Validity

How generalizable are your results? They may be valid for your population, but are they also valid for population X?

Do your (e.g., lab) results have anything to do with how people act and how things happen in the real world?

External/Ecological Validity

How generalizable are your results? They may be valid for your population, but are they also valid for population X?

Do your (e.g., lab) results have anything to do with how people act and how things happen in the real world?

Generally speaking a strength of qualitative methods over quantitative methods.

Detour: Another example

Suppose I want to understand the effect of free candy on Utah CS graduate students immediate, self-assessed quality of life.

There are 2 treatment groups: gets-candy and no-candy. Everybody in the gets-candy group will be given a piece of candy and then asked to rate their quality of life on a scale from 1-5. Everybody in the no-candy group is asked to rate their quality of life first, then offered a piece of candy.

Detour: Another example

Suppose I want to understand the effect of free candy on Utah CS graduate students immediate, self-assessed quality of life.

There are 2 treatment groups: gets-candy and no-candy. Everybody in the gets-candy group will be given a piece of candy and then asked to rate their quality of life on a scale from 1-5. Everybody in the no-candy group is asked to rate their quality of life first, then offered a piece of candy.

Assume I have a list of every CS graduate student in the department. If I go through that list and **just pick out the students I know best to participate in the study**, what are some potential problems with my study results?

Detour: Another example

Suppose I want to understand the effect of free candy on Utah CS graduate students immediate, self-assessed quality of life.

There are 2 treatment groups: gets-candy and no-candy. Everybody in the gets-candy group will be given a piece of candy and then asked to rate their quality of life on a scale from 1-5. Everybody in the no-candy group is asked to rate their quality of life first, then offered a piece of candy.

Assume we have a random subset of CS grads selected as participants. Suppose I assign **all the PhD students to the gets-candy group** and **all the masters students to the no-candy group**. What are some potential problems with my study results?

Detour: Sampling

- **Population**: all the people in the world who might be relevant to the research question asked, e.g., all potential touchpad and trackball users.

- **Sample**: a portion of the whole population used in an experiment, e.g., some subset of touchpad and trackball users.

Sampling

You need participants.

You're not going to study everyone in the world.

How do you choose how you recruit and inclusion and exclusion criteria for the study?

Qualitative Sampling

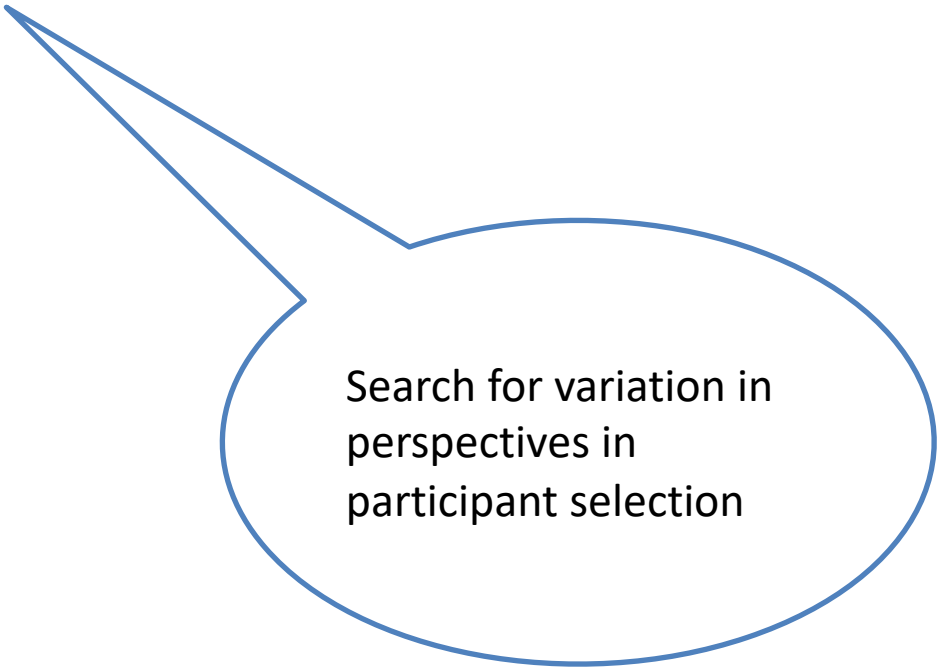
Your goal doesn't have to be representative! e.g.

- Purposeful sampling
  - Maximum Variation/Heterogeneous sampling
  - Homogeneous sampling
  - Typical case sampling
  - Extreme/Deviant sampling
  - Critical Case sampling
  - Total population sampling
  - Expert sampling
- Quota sampling
- Snowball sampling
- Convenience sampling

Qualitative Sampling

Your goal doesn't have to be representative! e.g.

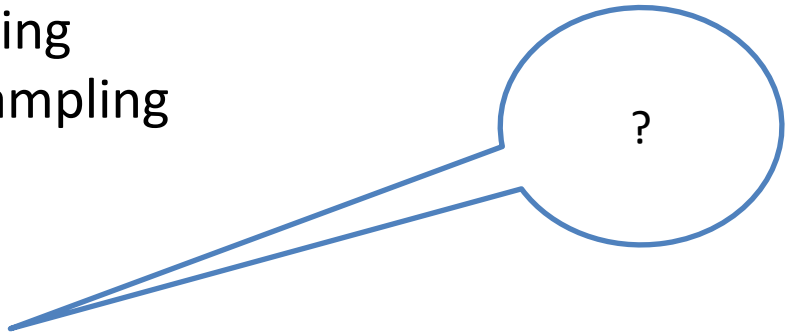
- Purposeful sampling
  - Maximum Variation/Heterogeneous sampling
  - Homogeneous sampling
  - Typical case sampling
  - Extreme/Deviant sampling
  - Critical Case sampling
  - Total population sampling
  - Expert sampling
- Quota sampling
- Snowball sampling
- Convenience sampling



Qualitative Sampling

Your goal doesn't have to be representative! e.g.

- Purposeful sampling
  - Maximum Variation/Heterogeneous sampling
  - Homogeneous sampling
  - Typical case sampling
  - Extreme/Deviant sampling
  - Critical Case sampling
  - Total population sampling
  - Expert sampling
- Quota sampling
- Snowball sampling
- Convenience sampling

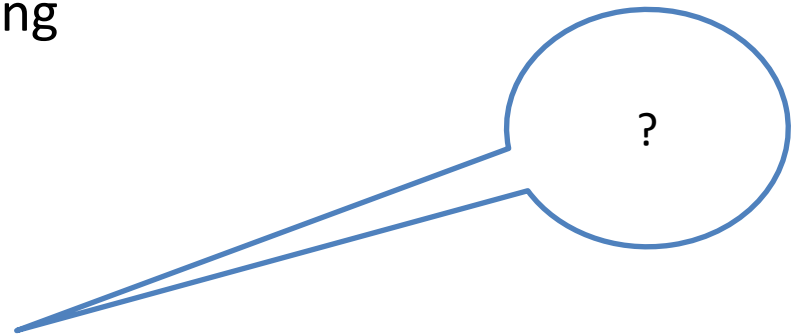


Qualitative Sampling

Qualitative Sampling

Your goal doesn't have to be representative! e.g.

- Purposeful sampling
  - Maximum Variation/Heterogeneous sampling
  - Homogeneous sampling
  - Typical case sampling
  - Extreme/Deviant sampling
  - Critical Case sampling
  - Total population sampling
  - Expert sampling
- Quota sampling
- Snowball sampling
- Convenience sampling



Related term: Saturation

Related term: Saturation

- “Diminishing returns”

- More data will not lead to more information related to the research questions

Related term: Saturation

- "Diminishing returns"

- More data will not lead to more information related to the research questions

Now we're done with the detour. We had just finished talking about reliability and validity.

https://pixabay.com/illustrations/important-colored-pencil-pen-red-2794684/

Different methods are good at different things.

Different methods are bad at different things.


Fast
Cheap
Good
Pick any two.

Different methods are good at different things.

Different methods are bad at different things.

(like what?)

Remember these?

- **Generalizability**
- **Precision/Control**
- **Realism**

Joseph E. McGrath. Methodology Matters: Doing Research in the Behavioral and Social Sciences. *Readings in Human-Computer Interaction: Toward the Year 2000*, 2nd edition, ed: Baecker et al., 1995.

“Although you always want to maximize [generalizability, precision/control and realism] simultaneously, you cannot do so.”

Why?

e.g., a carefully-controlled laboratory experiment increases precision, but decreases realism (and also generalizability)

Why?

e.g., a carefully-controlled laboratory experiment increases precision, but decreases realism (and also generalizability)

e.g., conducting a field study can increase realism, but decreases precision (and also generalizability)

Why?

e.g., a carefully-controlled laboratory experiment increases precision, but decreases realism (and also generalizability)

e.g., conducting a field study can increase realism, but decreases precision (and also generalizability)

e.g., a well-designed, large-scale questionnaire increases generalizability, but decreases realism and precision

“First, each strategy has certain inherent weaknesses, although each also has certain potential strengths…The first strategy you are encouraged to address…is: Does the material, as presented, properly reckon with the strengths and weaknesses of the research strategies it encompasses?”

Like methods, different forms of data have different strengths and weaknesses.

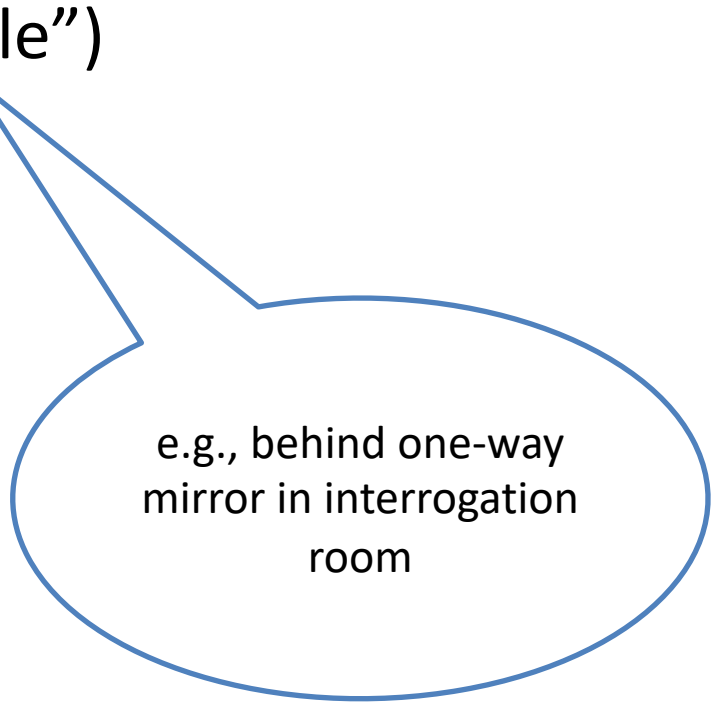
Simplified Classification of Measures in Social Psychology

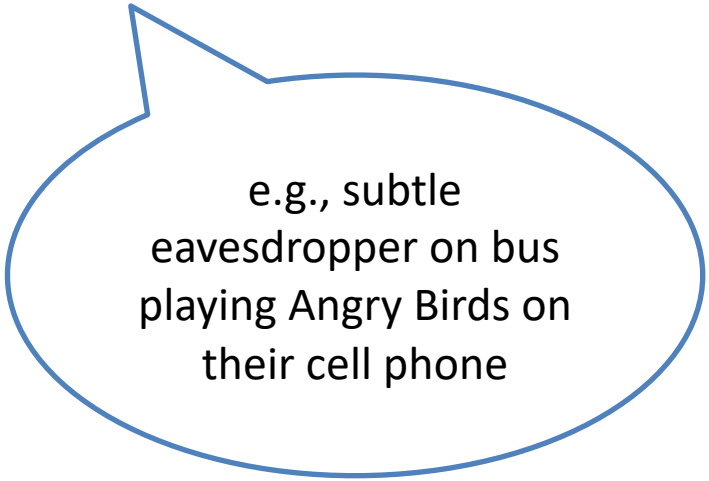
Simplified Classification of Measures in Social Psychology

- Self-reports (e.g., questionnaire)
-
-
-

Simplified Classification of Measures in Social Psychology

- Self-reports (e.g., questionnaire)
- Observations (observers may be "hidden" or "visible")
-
-





Simplified Classification of Measures in Social Psychology

- Self-reports (e.g., questionnaire)
- Observations (observers may be “hidden” or “visible”)
- Archival records (e.g., public Twitter feeds, private diaries)
-

Simplified Classification of Measures in Social Psychology

- Self-reports (e.g., questionnaire)
- Observations (observers may be “hidden” or “visible”)
- Archival records (e.g., public Twitter feeds, private diaries)
- Trace Measures (e.g., wear on museum floors, web logs)

Self-Report

Self-Report

- Versatile – ask anything!
- Low in cost (setup and per-participant)

- Potential Influence/Bias

Observations

Observations

- Versatile
- Hidden observers might reduce influence/bias

- Potential Influence/Bias with visible observers
- Observer errors
- Only overt behaviors
- Time-intensive
- Ethical concerns with hidden observers

Archival Records

Archival Records

- *Potentially* low on bias

- Potentially subject to bias
- May not be available / may only loosely relate to phenomenon of interest

Trace Measures

Trace Measures

- Unobtrusive
- Less bias

- May not be available / may only loosely relate to phenomenon of interest