

PAI 705
McPeak
Lecture 5

Conceptualization, Operationalization, and Measurement.

- Much of what we want to study does not exist in physical, tangible, directly measurable reality.
 - For example: Good governance. Development. Participation. Corruption. Diplomacy.
- What we are trying to do is come up with things that can be measured that have some logical relationship with the concept we want to use.
- Another way of putting it is that many of the concepts we use are not real things in the sense that you could trip over them but definitions that we have agreed to use and share to facilitate communication.
- Conceptualization – the mental process that takes fuzzy and imprecise notions – concepts- and makes them more specific and precise.

- Consider concepts like: Poverty. Well-Being. Inequality.
 - What elements are on the 'here is what each looks like' list in my head.
 - You also have a list in your head.
 - Through communication and agreement / disagreement we can refine the sense in which we are using the word.
 - 'Poverty' exists as a concept through a shared agreement what that word means.

- From Kaplan (1964) we can think of three things that we can measure as social scientists.
 - 1) Direct observables. Height, weight, eye color, number of earrings....
 - 2) Indirect observables. Characteristics of a person as they answer questions on a questionnaire, a written record of what happened. Where were you born? Where did you go to elementary school? What was discussed at the Arts and Sciences faculty meeting?
 - 3) Constructs. A 'family of conceptions'. Concepts are constructs, things we create. IQ. Compassion. Prejudice. The mental images, the conceptions, combine to form an agreed upon reality and way to combine direct and indirect observations to create a measure of the construct.

- Constructs are not real, they are created.
- Treating a construct as if it is a real thing is called reification.
- Constructs are not real or observable in and of themselves, but they have relationships to things that are real and observable.
- Conceptualization is the process by which we specify what we mean by a term in our research. Production of a specific agreed upon meaning of a concept for the purposes of our research.
- In making this meaning precise, we identify indicators. These are observable things, the presence or absence of which provide information about the concept we are interested in.
- A part of social research does not take the approach that we should define the indicators that are aligned with a concept, but rather try to tease out what the population we are investigating uses as indicators for a given concept.
 - What are the indicators of wealth and poverty using local concepts? Krishna's stages of progress exercise.
 - What are the indicators of good governance?

- Within our indicators, we may find clusters of related concepts that we can identify as dimension of the concept revealed by groups of indicators.
 - Poverty;
 - Health and Nutrition Dimension, a lack of adequate nutrition as specified by caloric intake measures and anthropometric measures,
 - Financial Dimension, a lack of access to financial capital as measured by past credit history and credit scores
 - Shelter dimensions, inadequate housing as measured by plumbing type, presence of electricity, roofing and flooring materials...

A part of social research is developing a taxonomy of indicators that capture essential elements of the concept.

The book provides the example of starting with the concept of 'genocide' and discovering there are four distinct kinds of genocide which need to be investigated as different phenomena.

- Convenience. Move people out of the way / prevent them from causing future problems.
- Revenge. Avenge past insult or injury.
- Fear. Prevent a growing problem before it becomes too large to confront and jeopardizes your own group's existence.
- Purification. Eliminate from your population elements that are viewed as creating impurities.

Analysis then could proceed on each of the four motivations for launching a genocide that would have more conceptual coherence and provide a better framework for analysis.

The interchangeability of indicators. If several different indicators all represent to some degree the same concept then all of them will behave the same way as the concept would if it was real and could be observed.

- Wealth from me.
 - Herd size.
 - Number of fields.
 - Roofing material.
 - Animals to pull plows.
 - Bank account.
- Wealth from you.
 - Number of shoes owned per child.
 - Number of pots and pans in the kitchen.
 - Value of jewelry owned by household members.
 - Number of children in university.
 - Frequency of hosting social events for community.

These both are capturing dimensions of the same concept but in different dimensions. It would seem reasonable that both approaches capture aspects of the same underlying concept (and we would suspect they co-move).

Specification is the process by which concepts are made more specific.

- We are moving towards nominal definitions of concepts; a consensus about how a term is to be used.

- The book presents one of the early examples of social science research, Durkheim's study of *anomie*. A sense of unconnectedness, of powerlessness.
 - He used the concept to develop a theory about the incidence of suicide.
 - He hypothesized it might have something to do with whether the person was Roman Catholic or Protestant.
 - Also the weather, contrast northern and southern areas.
 - Also compared different months and seasons.
 - Consider times of social upheaval and unrest.
 - The social upheaval leads to a sense of normlessness.

For Durkheim, *anomie* was a characteristic of society.

Merton in 1938 applies the idea of anomie to the relationship of the individual within society and traces its origin to a disparity between the goals we are told to strive for and the means we have to achieve these goals.

For Jagger it was a good lyric:

When I'm drivin' in my car
and a man comes on the radio
he's tellin' me more and more
about some useless information
supposed to fire my imagination.
I can't get no, oh no no no.
Hey hey hey, that's what I say.

Operationalization choices

- What range of variation am I likely to find?

- Defining 'bins'.

39) Do you **currently have a bank account?** Y N
if yes, is your current balance (circle)

- a. < 7,000 Ksh / 800 Birr (<\$100)
- b. from 7,000 Ksh to 17,499 Ksh / 800 Birr and 1999 Birr (\$100 to \$249.99)
- c. from 17,500 Ksh to 35,000 Ksh / 2000 Birr and 4000 Birr (\$250 to \$500)
- d. > 35,000 Ksh / 4000 Birr (>\$500)

- Covering both positive and negative sides of the response range:

- **Access to services and markets:** How do you assess your ease of accessing the following services and markets, both in the rainy season (R) and the dry season (D).

	Very Difficult	Difficult	Neither difficult nor easy	Easy	Very Easy
		<i>D</i>		<i>R</i>	
Weekly markets					
Administrative services of the Cercle or Département					
Banking services					
Microfinance					

- Don't want to just ask how easy it is, allow for the other direction as well, how hard is it to access.
- In the book the example of attitudes towards nuclear energy; range is not indifferent to strong support, but need to allow opposition.

- Also need to consider how specific you need to be in the information gathering.
 - Would I have been OK with ‘is it easy or hard to get to the market in the rainy season?’ Would I have been better to put it on a 1-10 scale from hardest to easiest to get more variation?
- Variable construction.
 - Defined to be exhaustive of possibilities for the attributes.
 - LGBTQ
 - LGBTQI
 - Democrat / Republican
 - Democrat / Republican /Green /Working Families/ Conservative/ Independent/ Progressive...
 - Every observation can be categorized according to attributes that correspond to that variable.
 - The attributes must be mutually exclusive so you can allocate observations according to the attributes.
 - For the political parties, it might be possible that a person is a working families and a green party candidate.
 - Can resolve by revising each party into a binary or start having other options:

Just Green, Just Working Families, Both Green and Working Families...

Levels of measurement

Nominal measures. Variables who have attributes that are different from each other. Gender. Program of study. Hair color. These measure offer names or labels for characteristics. It breaks the sample into coherent groups, where within the group they are the same by this measure and across groups they are different.

Ordinal measures. If there is some kind of logical ordering, then we have ordinal measures. On the measures of ease of access noted above, those are ordinal. A ranking exercise is ordinal.

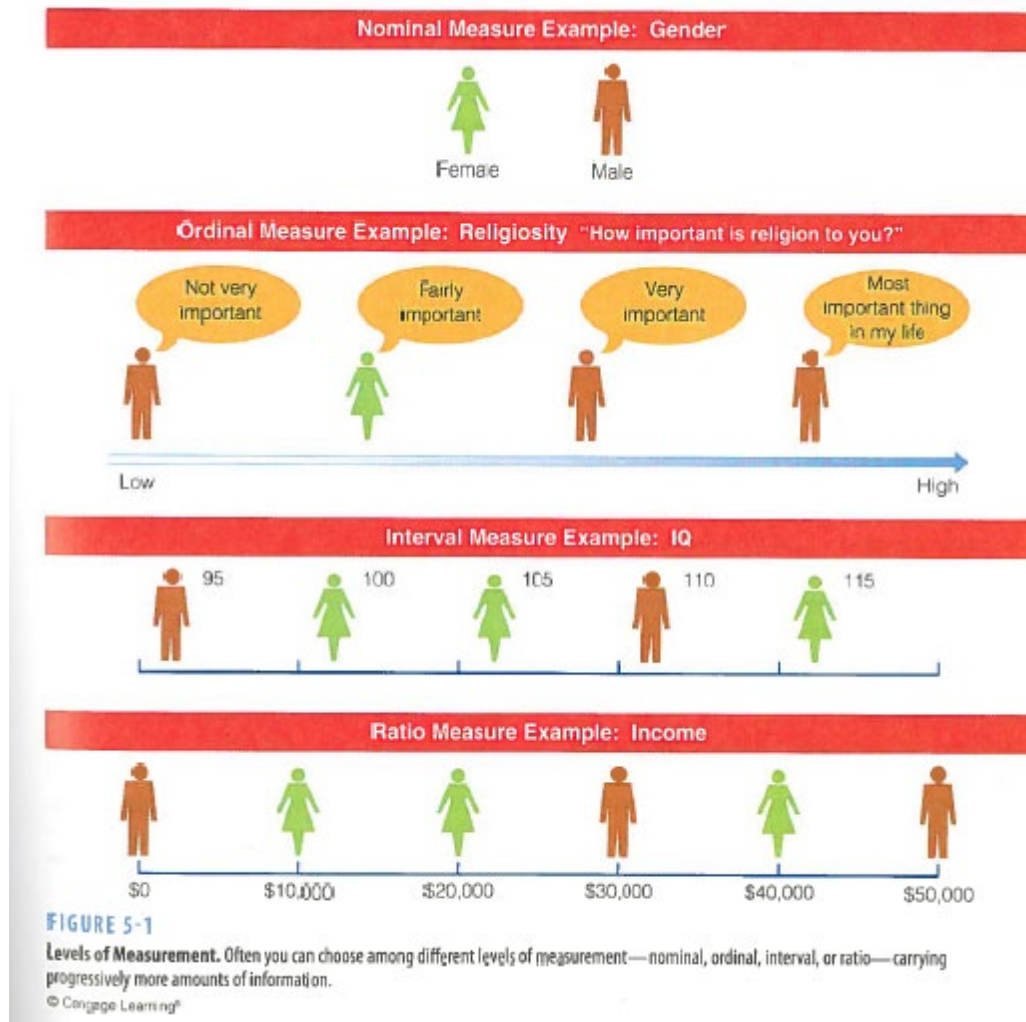
41) We know that households in this area are concerned about problems that could happen to them. We have made a list of concerns people commonly tell us about. I am going to read to you this list of concerns, and I would like you to tell me **which of these you are afraid could affect your household in the coming three months**. If you are concerned about one that is not on the list, please tell us what it is.

(Enumerator, first read through the whole list, then go through one item at a time for the response. If a concern not on the list is mentioned, write it in the space after “other” in row 12)

Type of Concern	Yes or No		Ranking
1. Not enough pasture for animals	Y	N	
2. Not enough water for animals	Y	N	
3. Animal sickness / death	Y	N	
4. Animals loss due to theft / raiding	Y	N	
5. Insecurity / violence / fights	Y	N	
6. Human sickness	Y	N	
7. No buyers for animals you wish to sell	Y	N	
8. Low prices for animals you wish to sell	Y	N	
9. Not enough food for people	Y	N	
10. High prices for things you buy	Y	N	
11. Crops fail	Y	N	
12. Other (describe):	Y	N	

Now, I would like you **rank the concerns you identified**. First, tell me which of the concerns you identified you are **most** afraid of, then the concern you are **next most** afraid of, and so on until they are all listed in order.

- Interval measures.
 - The actual distance between the attributes of the variable has some meaning. Beyond just the order, the distance between categories now matters. Income. GPA. GRE test results. IQ test results.
- Ratio measures.
 - These are a category of interval measures that have the following characteristic: the attributes composing a variable have a true zero point. Age. Number of semesters enrolled in Syracuse University. Years teaching on the faculty. Income.



- You can go from an interval measure and turn it into a nominal measure; say you define mean income, then $\text{rich} = 1$ if above the mean, $= 0$ if below the mean for observation i .
- Nominal does not allow you to go to interval in the other direction though.

- Though you can go from a set of nominal to construct an interval. Yes=1, no=0, add up answers.

▪ 42) Please tell me if anyone in your household has used the following services since this time last year.

Type	Government		Private		Mission		NGO		Other (describe)
Health Services	Y	N	Y	N	Y	N	Y	N	
Veterinary Services	Y	N	Y	N	Y	N	Y	N	
Education Services	Y	N	Y	N	Y	N	Y	N	
Security Services	Y	N	XXXXX		XXXXX		XXXXXX		XXXXXXXXXX

<i>Level of Measurement</i>	<i>Arithmetic Operations</i>	<i>How to Express the Fact That Jan Earns \$80,000 a Year and Andy Earns \$40,000</i>
Nominal	$= \neq$	Jan and Andy earn <i>different</i> amounts.
Ordinal	$> <$	Jan earns <i>more</i> than Andy.
Interval	$+ -$	Jan earns <i>\$40,000 more</i> than Andy.
Ratio	$\div \times$	Jan earns <i>twice</i> as much as Andy.

Criteria of Measurement Quality

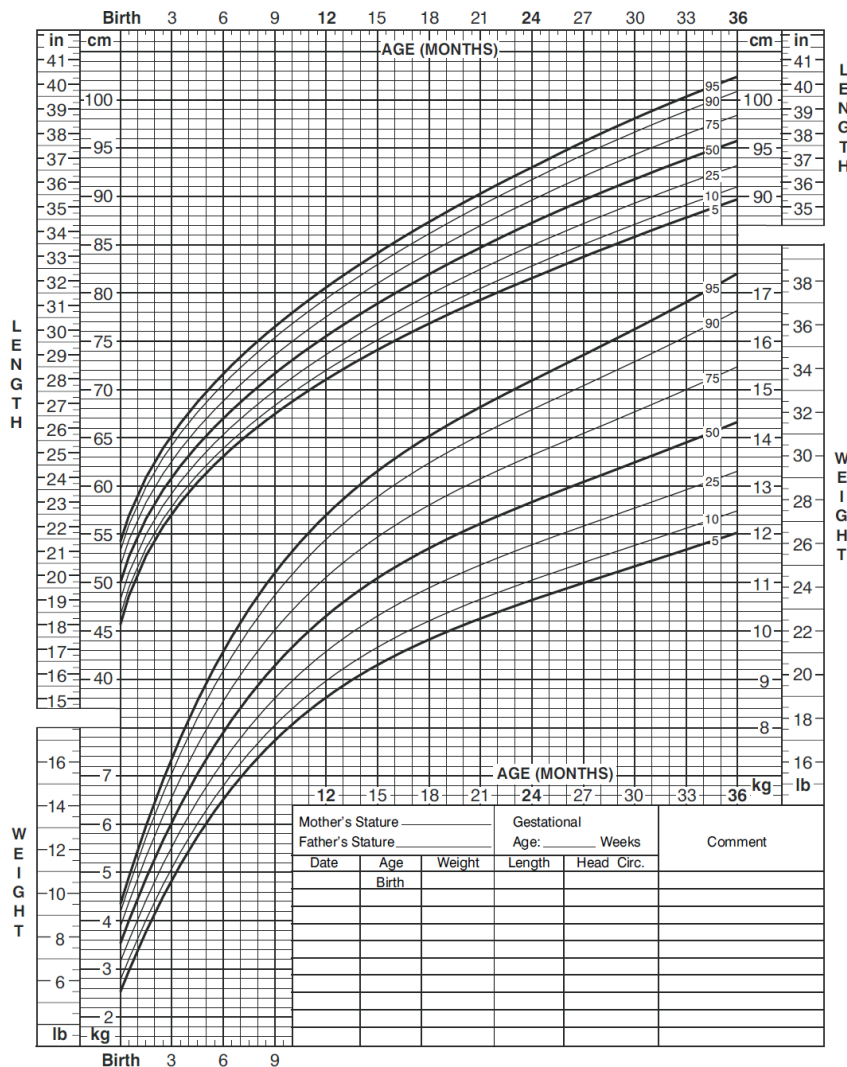
- Precision. At what level of specificity do you need the measurement?
 - NDVI, 8 km by 8 km, 2 km by 2 km, 100 m by 100 m?
 - Age in days, weeks, months, year?

Birth to 36 months: Boys

Length-for-age and Weight-for-age percentiles

NAME _____

RECORD # _____



Published May 30, 2000 (modified 4/20/01).
 SOURCE: Developed by the National Center for Health Statistics in collaboration with
 the National Center for Chronic Disease Prevention and Health Promotion (2000).
<http://www.cdc.gov/growthcharts>



Reliability.

See also <http://www.socialresearchmethods.net/kb/reotypes.php>

- Whether a particular technique applied repeatedly to the same object yields the same result each time.
- Reliability does not ensure accuracy.
 - Your scale is not set to zero as a baseline.
- Subjectivity is a problem that can work against reliability.

- Grading the same problem one after the other to ensure the same points are taken off for the same problem in an answer.
 - Put away the memos since I am grumpy.
- Your respondents may also not be reliable.
 - Answering strategically
 - Answering based on what just happened and extrapolating
 - Sometimes it is your fault for asking something so precise there is no way a person will be able to answer. Exactly how much did you spend on groceries in your first shopping trip of 2018?
 - Or asking something so obscure that they make up an answer on the fly that might not be the same answer they would give if you asked again. Who had the most passing yards for the total season for the past five SU starting quarterbacks?
 - Ask people things they are likely to know the answers to.
 - Ask things that are relevant to them.
 - Ask things to a reasonable degree of precision.



FIGURE 5-2

An Analogy to Validity and Reliability. A good measurement technique should be both valid (measuring what it is intended to measure) and reliable (yielding a given measurement dependably).

© Cengage Learning®

- Test – retest method. Health Hazard Appraisal.
 - Ask people a set of baseline questions about characteristics and behavior.
 - Ask same questions again about 3 months later.
 - Only 15% reported the same information in each round.
 - One of the hazards with panel data is that you catch these kinds of things as well, notably people getting younger.
- Split half method.
 - I have ten questions that operationalize a concept.
 - Randomly assign half of the questions to one group and the other half to another.
 - Randomly assign who gets the different questionnaires.
 - Do they predict the same level of the concept in the population or do they predict different levels?

- Also used in writing test questions; Write a 100 point exam to test knowledge on a topic, assign questions worth 50 points to instrument A and 50 to instrument B, test half the group on A and the other on B, then switch and the ones who took A take B, the B takers take A, and see if the scores for an individual on A and B are acceptably correlated.
- Use measures that have some history of use elsewhere.
 - Risk ranking in East Africa led to risk ranking in West Africa
 - GRE
 - Five point scales
 - IQ tests
 - May sometimes have to adapt.

Reliability of Research workers.

- Verify they actually asked the questions
- Have different ones code the same items to see if they code them the same and use and discrepancies as training material.
- Verify translations
- Verify orthography

- Build in drop down options when possible. If you don't you are setting the stage for the following:

County	Community	Village
Dehas	Rero	Debelaraka
Dehas	Raro	Dembela reka
Daese	Rero	Dembel reke
Daese	Rero	Dambel reka
Das	Rero	Demibela raka
Das	Rero	Dembela reka
Das	Raro	Dembele raka
Dehas	Raro	Dembela raka
Dahas	Raro	Dambala raka
Dahas	Raro	Dembela
Dahas	Raro	Dembela raka
Dehas	Raro	Denbala raka
Dehas	Raro	Danbala raka
Dehas	Raro	Dembelaraka

These are all the same place.

Household #	Is the IED Afrique/DFC project active in your community?	If no, are aware of the (<i>describe project</i>) in the treatment village (<i>name</i>) implemented by the IED Afrique DFC project?	If yes, have you been involved in discussion about this project over the past two years?	Has there been an investment proposed for your community	Reason 1	Reason 2
103	NON	OUI Reboisement et construction foyer de femmes	NON	JE NE SAIS PAS	0	JE NE SAIS PAS
104	NON	OUI Reboisement et construction foyer de femmes	NON	JE NE SAIS PAS	0	JE NE SAIS PAS
105	NON	OUI Reboisement et construction foyer de femmes	NON	JE NE SAIS PAS	0	JE NE SAIS PAS
106	NON	OUI Reboisement et construction foyer de femmes	NON	JE NE SAIS PAS	0	JE NE SAIS PAS
107	NON	OUI Reboisement et construction foyer de femmes	NON	JE NE SAIS PAS	0	JE NE SAIS PAS
108	NON	OUI Reboisement et construction foyer de femmes	NON	JE NE SAIS PAS	0	JE NE SAIS PAS
109	NON	OUI Reboisement et construction foyer de femmes	NON	JE NE SAIS PAS	0	JE NE SAIS PAS
110	NON	OUI Reboisement et construction foyer de femmes	NON	JE NE SAIS PAS	0	JE NE SAIS PAS
111	NON	OUI Reboisement et construction foyer de femmes	NON	JE NE SAIS PAS	0	JE NE SAIS PAS
112	NON	OUI Reboisement et construction foyer de femmes	NON	JE NE SAIS PAS	0	JE NE SAIS PAS
113	NON	OUI Reboisement et construction foyer de femmes	NON	JE NE SAIS PAS	0	JE NE SAIS PAS
114	NON	OUI Reboisement et construction foyer de femmes	NON	JE NE SAIS PAS	0	JE NE SAIS PAS
115	NON	OUI Reboisement et construction foyer de femmes	NON	JE NE SAIS PAS	0	JE NE SAIS PAS
116	NON	OUI Reboisement et construction foyer de femmes	NON	JE NE SAIS PAS	0	JE NE SAIS PAS
117	NON	OUI Reboisement et construction foyer de femmes	NON	JE NE SAIS PAS	0	JE NE SAIS PAS
118	NON	OUI Reboisement et construction foyer de femmes	NON	JE NE SAIS PAS	0	JE NE SAIS PAS
119	NON	OUI Reboisement et construction foyer de femmes	NON	JE NE SAIS PAS	0	JE NE SAIS PAS

Validity.

The extent to which an empirical measure adequately reflects the real meaning of the concept you are investigating.

We are actually measuring what we say we are measuring.

“real”? “Actual”? I thought we were all post-positivist and such.

Face validity. It is plausible that the measure has something to do with the concept we are trying to measure. Height and winning the presidential election? Winning Iowa and winning the election? Winner with more letters in last name?

Criterion related validity / predictive validity. GRE predicting success in a master’s program. Undergraduate GPA predicting graduate GPA. CV at the time of hire for faculty predicting future research productivity.

Construct validity; the degree to which a measure relates to other variables as expected within a system of theoretical relationships. As a measure of the degree to which people commute to the university area, I record each day if cars have moved on streets from Westcott to Ostrom. Alternate side parking suggests my measure does not have construct validity.

Content validity. This is concerned with how much the chosen measure covers the range of meanings included within a concept. My method of recording vehicle parking patterns is

flawed on content validity as well; should I not also look at parking garages and parking lots on campus? I capture one dimension, street side parking in the area, but not others, garage and lot use. Might want to look at bus use as well.

Tensions between reliability and validity (in statistics you may have seen consistency and bias as the related conceptual pair).

From one way of looking at it, the quantitative, closed form questionnaire is more likely to be reliable. The qualitative, open ended, capture all the different directions that could explain something is more likely to be valid.