PAI 705

McPeak

Lecture 1

Epistemology – the science of knowing.

Methodology – a subfield of epistemology, the science of finding out.

Social science research is based on using a methodology (and there are different ones to use) to find things out.

How do we know what we know?

How much is direct experience?

How much is passed on?

What does it mean to know things are true?

Agreement Reality. The things we 'know' as part and parcel of the culture we share with those around us.

In what ways does this help?

In what ways does it hinder?

Role of Tradition in establishing agreement reality; don't necessarily have to experience to have knowledge.

Role of Authority in establishing agreement reality; experts stating assessments about what is a fact.

Agreement reality is not the same as proven reality, which is part of what we try to test in social science inquiry.

Problems in Social Science inquiry.

- Inaccurate observation / lack of attention to specific detail. We need to be deliberately paying attention to particular facts and recording them correctly.
    - It is hard to recall a specific item if we were not paying attention to this item.
- Overgeneralization. We see one case and we think it applies to all cases.
- Selective Observation. We see what we expect to see. We dismiss contrary evidence to confirm the version of reality we choose to see rather than allow the contrary evidence to refute our favored interpretation.
- Illogical reasoning. The exception proves the rule. Dismiss evidence that refutes one interpretation as exceptions.

The foundation of social science.

Theory, the logical foundation that provides systematic explanations.

Data collection is the observational aspect.

Social theory has to do with what is, not what should be.

Normative compared to positive approaches.

The approach in research is positivist.

Distinct from what should be, or values, which are normative.

In our program, we do think about normative goals with indicators and objectives.

As an example, we can take the sustainable development goals. Development should look like improvement on these agreed upon indicators.

https://sdgs.un.org/goals

That is a first step: establish what we mean by development and how it will be measured.

Then we have to come up with a means of reporting on the values for the agreed upon indicators.

Overall in social science research we are looking for patterns of regularity; average overall behavior that may have exceptions, but describes the overall tendency.

Overall tendency is where social science research draws on statistics training. We are working to establish evidence about facts based on the 'law of large numbers'.

- In doing analysis and understanding behavior we need to understand the context.
    - What are the formal rules and regulations?
    - Where can you park your car and on which side of the street today?
    - Can you eat food in this classroom?

We also need to think about informal norms and customs.

- Which side of the stairs is for going up and which side is for going down?
- Which side of the sidewalk for east to west crossing of the quad?

A critique of social science is that when we identify a regularity, that regularity is trivial and obvious. We are spending a lot of effort to prove things that are already known. "Female headed households in rural Senegal are poorer than male headed households".

However, what makes things interesting is that which is understood as true might not hold up to investigation. "Actually, they are not statistically poorer due to the role of male labor outmigration and remittances."

In the goal of publishing social science research, we are asked to establish how the findings we are presenting are a contribution to knowledge.

That which seems trivial *ex post* was not understood *ex ante*.

Ex post is after the fact. Ex ante is before the fact.

To stay with the Senegal example, "Fulani female headed households are poorer but Wolof are not due to differing land inheritance customs that vary across ethnic groups."

Figuring out which is the exception and which is the general pattern is not always easy, especially if not treated with some systematic design.

We do test hypotheses by adopting the formal structure of probability, and appealing to the law of large numbers.

What is the general tendency and what is an outlier (and what does the outlier tell us).

Sometimes we appeal to typologies to structure our analysis

Create sub-groups: Data from 330 households in northern Kenya and southern Ethiopia, stratified by herd size and cash income in the first encounter.

Use of medians to create four quadrants. Then using the tools of statistics (t tests) to test for statistically significant differences.

Come up with names to apply to the created sub-groups.

| Group | Herd Size TLU | Total Income per capita per day | Cash Income as % of Total Income | Total Income variability (cv) |
|---|---|---|---|---|
| 1) Left out | 7.3 | $0.20 | 29% | 1.32 |
| 2) Moving From | 7.2 | $0.27 | 46% | 0.90 |
| 3) Staying With | 23.7 | $0.34 | 21% | 0.82 |
| 4) Combining | 26.0 | $0.46 | 35% | 0.63 |
| Significant difference in means by groups, t-statistics | t12 ,<br>t13 ***,<br>t14 ***,<br>t23 ***,<br>t24 ***,<br>t34 , | t12 ,<br>t13 ,<br>t14 **,<br>t23 ,<br>t24 ***,<br>t34 , | t12 ***,<br>t13 **,<br>t14 *,<br>t23 ***,<br>t24 ***,<br>t34 ***, | t12 **,<br>t13 ***,<br>t14 ***,<br>t23 ,<br>t24 ***,<br>t34 ***, |

Note that observation has challenges.

The knowledge by people we are observing that they are being observed may lead to changed behavior, making the observed behavior not the 'natural' behavior but the response to being observed.

They might be performing for us so we are recording their construct, not their natural behavior.

The "Hawthorne Effect". Observing workers in a plant in Illinois in the 20s and 30s. They changed working hours, break times, lighting in the plant. Worker productivity went up when these changes were made. Eventually they figured out the productivity went up because people were aware they were being observed, not due to the changes.

Hard to observe behavior if behavior changes in response to observation.

People perform.

People can report strategically.

Note challenge to informed consent. In discussing ethics, we will note there is a requirement that people provide informed consent. There is an inherent conflict in research.

What you see is also filtered against what you are prepared to see.

Our claim as positivist researchers is that we can establish and reveal truth.

“objective truth” can be challenged.

A book by Faulkner called Absalom, Absalom is written about the same set of events, but from different perspectives. It struggles with the idea of what is the truth of what happened when there are multiple perspectives on the same set of events when people have different positions in society.

A passage from the world of fiction:

**Unidentified participant:** Mr. Faulkner—

**William Faulkner:** Yes, sir.

**Unidentified participant:** In *Absalom, Absalom!* is any one of the people who talks about Sutpen have the right view, or is it more or less a case of thirteen ways of looking at a blackbird with none of them [getting it] right?

**William Faulkner:** That's it exactly. I think that—that no one individual can—can look at truth. It—it—it blinds you. You look at it, and—and you—you see one phase of it. Someone else looks at it and sees a slightly awry phase of it, but taken all together, the truth is—is in what they saw, though nobody saw the truth intact. So—so these are—are true as far as—as Miss Rosa and as Quentin saw it. Quentin's father saw what—what he believed was truth. That was all he saw. But the old man was—was himself a little too big for—for people no greater in stature than Quentin and Miss Rosa and Mr. Compson to see all at once. It would've taken, probably, a wiser or more tolerant or more sensitive or more thoughtful person to see him as he was. It was, as you say, thirteen ways looking at a blackbird. But the truth, I would like to think, comes out, that when the reader has read all these thirteen different ways of looking at the blackbird, the reader has his own fourteenth image of that blackbird, which I would like to think is the true one.

http://faulkner.lib.virginia.edu/display/wfaudio29_1

Another take on this idea is that we put interpretations on things that are more rooted in our social science training than they are the lived reality of those experiencing them.

http://www.jstor.org/stable/665280?seq=1#page_scan_tab_contents

Social science is study of the aggregate, rather than the individual.

We are looking for collective regularities.

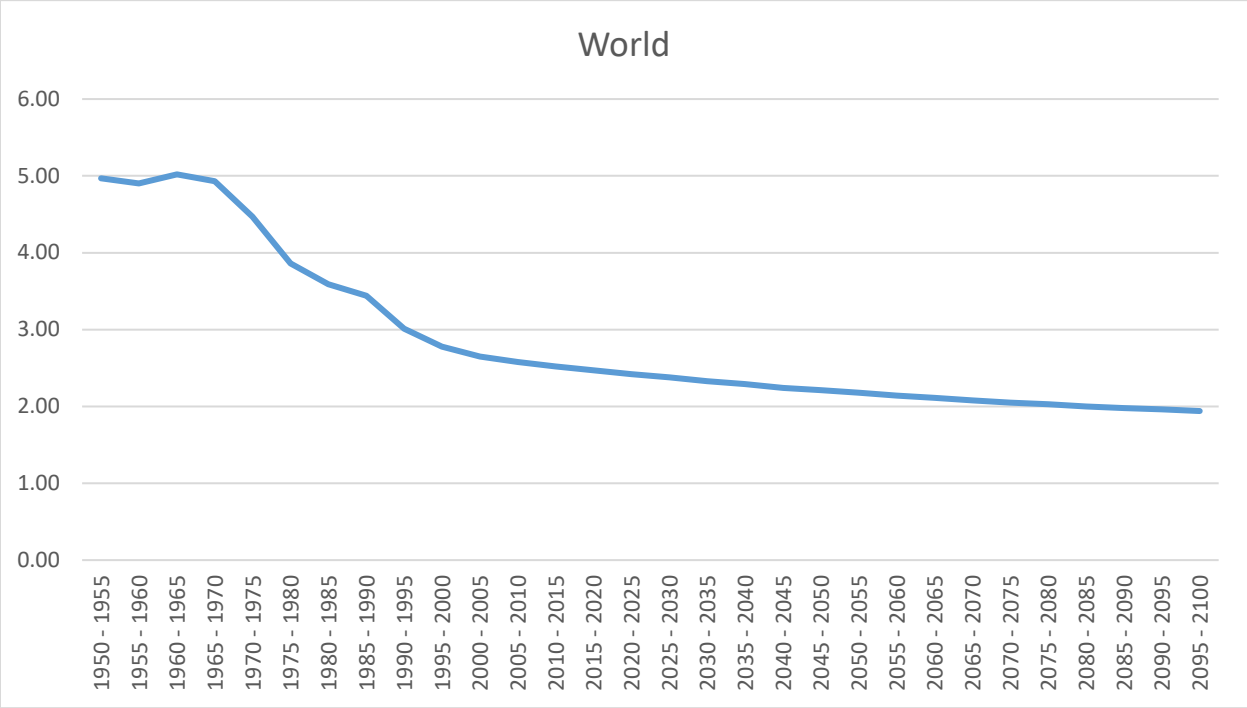
It is fundamentally based on individual decisions, but in aggregate, they create aggregate facts.

Take a population measure like the Total Fertility Rate (TFR).

TFR is the average number of children a woman will give birth to in her reproductive years conditional upon her birth cohort.

From the United Nations

https://population.un.org/wpp/DataQuery/




Total Fertility Rate for example, that changes over time but is not a collective decision, but a characteristic of individual decisions aggregated to the collective level. Not an outcome of social planning (though may reflect social programs and policies).

The goal is not to understand an individual person, but ‘to understand the systems in which people operate, the systems that explain why people do what they do’ (p. 14).

Variable.  Sets of attributes that can have different values for different units we might observe.

Attributes.  Characteristics of peoples or things.

| Variable | Attributes |
|---|---|
| Age | Number of years since birth |
| Gender | Female, Male, Other |
| Occupation | Professor, Lawyer, Candymaker |
| Social Class | Upper, Middle, Lower, Lumpenproletariat |
| Height | Value in centimeters |

Attributes are agreed upon concepts that sort complicated reality into conceptual ‘bins’.

We are trying to define concepts, and then find ways to measure these concepts.

The ‘bins’ are constructs, choices.

We have to figure out what to do with a transgender lawyer who makes candy on the weekends from a wealthy family who self identifies as working class.

Above, we used constructed groupings:

| **Group** |
|---|
| 1) Left out |
| 2) Moving From |
| 3) Staying With |
| 4) Combining |

So sorting cases into bins is part of the art, or our first step from complicated reality to our artificial constructed version of reality used for analysis.

When looking at two variables, we might see some kind of pattern in correlation. The variables move together in some sense.

Rho=Expected value [((x – mean x)*(y-mean y))]/(standard deviation x*standard deviation y)

That computation is mechanical based on statistical theory.

Do things move together positively or negatively?

Total Fertility rate and wealth?

Height and lifetime expected income?

There is a constant struggle to distinguish between 'correlation' and 'causation'.

Household self-assessment for resilience on a 1-5 scale and responses to how many months out of a year are the able to provide sufficient grains for the family rho=.52.

To put structure on the correlation, we can reach for the idea of causation. That moves beyond statistical manipulation and forces us to reach for a behavioral theory.

Much of what we do in social science research is a search through correlations to try to identify causation.

Why does Total Fertility Rate decrease as expected income increases?

Why do taller people have higher expected lifetime earnings?

We are trying to identify independent variables, the things on the right hand side, the predictors, from the dependent variables, the things on the left hand side, the outcomes.

In classic y=f(x) notation, the y is the outcome, the dependent variable, the x is the independent variable(s).

We spend time arguing about what is the outcome and what is the input. What is the dependent variable? What is jointly determined?

Higher income countries have better governance.

Countries with better governance have higher incomes.

Higher wages go to people who are less likely to be absent for health problems.

People with higher incomes eat more healthy diets.

Q=286-20*p is the demand

Q = 88+40*p is the supply

Also have to worry about the outcome being related to the relationship in question.

Students are highly satisfied with the IR curriculum based on our survey.

Students who have higher satisfaction are the ones that complete the survey.

Results indicate students are not satisfied with the IR curriculum.

Students who are unhappy are more likely to fill out the survey.

Our survey indicates 95% of people who went through our treatment program are employed and no longer homeless.

We did not find 40% of the sample who went through treatment and are still homeless. Because we could not contact them we have 95% of 60% we could locate so .95*.60+0*.40 =0.57 is a worst case possibility.

There are three main purposes / phases for conducting Social science research:

1) Exploratory. What do people around here mean by the word 'household'? What is a working definition of this concept that we can use to pose questions. Herd owning, common cooking pot, joint labor in fields, wives and co-wives, labor migration concepts….
2) Descriptive. How many households in this area are female headed? What is the average number of children? What is their experience with formal education? What is the mid upper arm circumference for the kids?
3) Explanatory. Do the kids residing in female headed households have higher or lower educational attainment than kids from non-female headed households? How about nutrition as measured by Mid-Upper Arm Circumference (MUAC)? How can we explain differences in childrens' nutrition with differences in the share of household income controlled by women?

Different methodologies and skills are called for when trying to address each purpose.

In research teams, different people tend to have comparative advantages in the different purposes, but we need to do all kinds at some point.

Kinds of Explanations.

Ideographic explanations. The full listing of all the individual influences that lead to one specific unique outcome. Explain by exhaustive listing the multiple reasons a given unique outcome was realized. Fully explaining the single case / event in question.

- What are all the factors that led to the Brexit decision?
- What factors led to a return to in-person instruction at Maxwell for the Spring 2022 term?

Nomothetic explanations. Identify a set of causal factors that lead to general classes of outcomes when there are multiple outcomes to consider. A general explanation of what leads to a class of conditions and outcomes that is reporting trends on average, and possibly deviations from trends.

- What economic, political, and historical factors have influenced referendum votes on political and economic integration in the post war era?
- What are the most important variables in the characteristics of students in Maxwell that predict they will be enrolled in this class?

In social science theory, we have two main ways of thinking about establishing patterns:  Induction and Deduction.

Induction; moving from the particular to the general.
Take observations of individual cases and make statements about general patterns.
Begins with whether (I wonder whether this factor matters for that outcome) and moves to why (Hmm, it seems to matter, I am going to try to explain why it matters).

> Individual household wealth seems to influence whether children attend secondary school. How to interpret answers on the survey form in terms of overall patterns.

Deduction.  Moving from the general to the specific.  Begins with why and moves to whether.  In other studies, we have found household wealth influences school enrollment rates.  Is this the case in southern Ethiopia?  How to write questions on the survey form to see if patterns seen elsewhere hold up here as well.

Figure 1-3.

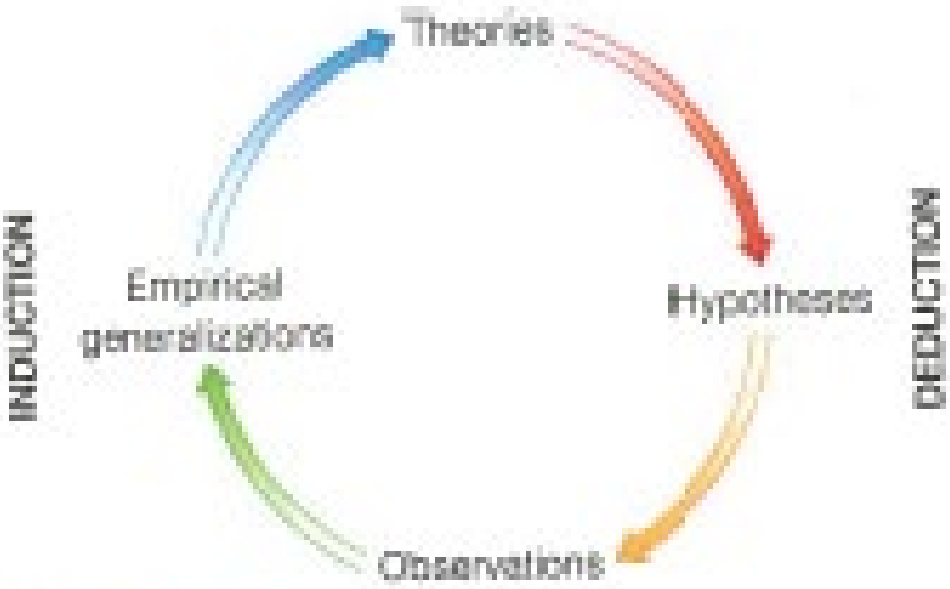


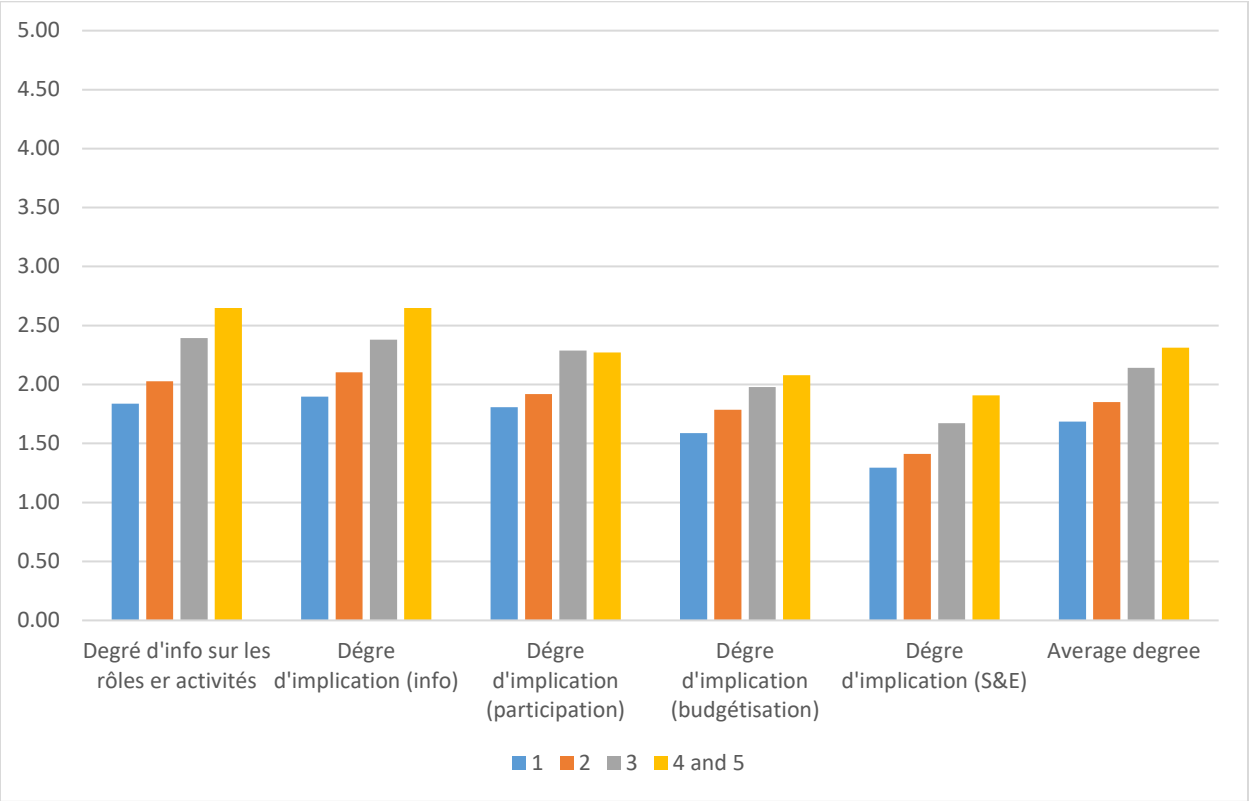

**FIGURE 1-3**

**The Wheel of Science.** The theory and research cycle for the social sciences can be compared to a relay race: although all participants do not necessarily start or stop at the same point, they share a common goal—to examine all levels of social life.

Source: Adapted from Walter Wallace, The Logic of Science in Sociology (New York: Aldine deGruyter, 1971). Copyright © 1971 by Walter L. Wallace. Used by permission.

## An inductive result

**Degree of involvement in community development activities (y axis) grouped by self-scored resilience on 1-5 scale (x axis)**




Observing this pattern can lead to a theory that could be tested in a deductive manner.

Determinism versus agency; $y=f(x)$ or $y=f(x)+u$.

By setting up structure of average behavior we have not described individual decision making or motivation, but a description of the aggregate.

Part of the challenge we face is to make decisions about quantification. Putting numbers on things. What year were you born? But why years? Months? Days?

For Resilience and degree of implication in monitoring and evaluation we created a Likert scale: 1,2, 3, 4, 5

The interpretation is relative to the constructed scale.

> Qualitative data. Are you young or old? What is the threshold that divides? Again, there are decisions we make when we analyze data to report out findings.
>
> We also have quantification of qualitative concepts if we agree to measures and constructs.

What does social science research look like?

Overview of Research Proposals (but also of research projects):
Abstract.
Introduction.
Review of the Literature.
Specify the Problem / Question / Topic noting explicitly why this study is needed and what is being added to knowledge.
Research Design.
    Data gathering methods
    Sampling frame and population to whom this research is applicable
    Research Ethics
Plan for analysis of data, noting methods and standards of evidence
Discussion and conclusion; what will we know if you do this research that we don't know now (reiterate now that it is all detailed as described above).

*Tell them what you are going to tell them, tell them, then tell them what you told them. Aristotle.*