Lecture 3 - Design DHX_MET1 Methodology 1

Stanislav Ježek

Faculty of Social Studies MU

So far we have…

- decided for a research question
- developed a theoretical framework and specified hypotheses

RESEARCH DESIGN?

Research question → Research Project

- Design - a strategy, plan of…
  - How will I find answer to my RQ?
    - How will I test my H?
  - Allows me to assess the VALIDITY of the answer
    - Designs have their strengths and weaknesses

RESEARCH DESIGN - DATA

- What are my variables (phenomena)? – Th. Frmwrk.
- What **data** represent my variables?
  - How do I get the data?
    - Find?
    - How to create data?
  - What will be the limitations of the data?
    - Representativeness
      - people, places, time, phenomena…
    - Validity – certainty about the variables affecting the data (intended, unintended)

BASIC ELEMENTS OF RESEARCH DESIGNS - **CHOICES**

- Gather existing data …. create data?
- Precisely measure a few variables … gather rich, contextual data (many variables)?

- Low interference (lurking) … high interference?
- Field … laboratory?
- High control …. natural occurrence

- Focus on one time …. follow the processes?

So, what is it that we want to know about designs?

1. **Choose** a design based on our RQ and resources
2. **Specify** the design so that it can produce high quality data
3. **Execute** the design
4. … Analyze the data

Types of research questions

- Exploratory – we are not sure about variables/concepts
  - Little established theory, uncertainty about relevant variables, concepts
  - Focus on understanding the phenomenon in its context, on its meaning
  - Proposals of concepts/variables and their values (categorizations) or even theory of their relationships
- Descriptive – we have questions about variables
  - What are variants of phenomena that occur and how frequently they occur
  - Correlational – are there any associations among the occurrences?
- Causal – does one phenomenon give rise to another?

1. The independent and the dependent variable should covary.
2. The independent variable (the presumed causal factor) should precede the dependent variable.
3. No other factor should be a possible cause of the change in the dependent variable.
4. A logical explanation (a theory) is needed and it must explain why the independent variable affects the dependent variable.

Designs - templates for research – by discourse

- Experimental design
  - experimental, quasi-experimental, single-case/small-N experimenting, ex-post-facto
- Survey research
- Observations
- Case studies
- Content/thematic analysis, Grounded theory
- Action research, Evaluation research

Ethnography

- Very general theory - anthropology, sociology, psychology...
- RQs how are things done around here. How are the most basic goals achieved here.
- **Descriptions** with focus on the meaning of behaviors.
- **Participant observation** as the main method of data generation
- Takes a long time, even years
- Interference should be low due to habituation
- Mostly an inspiration for flexible designs with shorter time-frames

Case study

First, look at this case, then we will decide what to do next.

- Exploration, description
- Broadest possible range of methods to gather and create rich data
- A lot of theory is used to make sense of the data
- **Robert Yin** textbooks are the best intro
- Difficult generalizations

- Case: person, group, organization, program, economy…
- N is often >1

Case study – practical features

- Hard to hold on to your RQ
- interactivity, flexibility
- Potential for rich data, triangulation (burden for analysis)
- Need to select case very thoughtfully
- Keeping balance between observing and intervening
- Beware of „success stories“!
- Generalization by theory and replication.

Content/thematic analysis study
Grounded theory study

- Theory generation - proposal
- A number of cases that we compare to see what is common and what is different
- Actually, an analytical approach with so much theory that it becomes a design, still may be used as analytical tool within other designs
- Focuses on texts and interview transcripts – less time in the field
- Produces content units – topics – that may be the basis for the definition of concepts / variables and their values
- GT goes on to formulate a theory of a phenomenon – what gives rize to it, what are the contextual variables that affect it and what are the consequences

Observational designs

How often phenomena occur?

Do various phenomena associate (co-occur)?

- Observing the natural occurrence of variable values (phenomena) and their frequencies
- May be naturalistic/unstructured …. highly structured
- Minimal interference with the observed processes
- You know exactly what you want to observe but it may be difficult to get access and time
- Limited number of variables – good planning necessary
- Descriptive RQs, correlational RQ, longitudinal RQs
- Observation is used as a data generation method in other designs
- Good sampling is key to representative data

Surveys

How often phenomena occur?

Do various phenomena associate (co-occur)?

- Serves the same purposes as observational research but instead of observing we ask for the observations of others
- We may ask about much more than we may observe – the price is we are using untrained observers who are asked about past events – high level of uncertainty
  - Surprisingly high amount of expertise and theory needed to create good data
- Wide range of structuring options
- Huge range – we may ask about anything

Surveys – practical properties

- Survey itself may be an intervention
- Self-report validity
  - people know less than we think
  - the correlation between *saying* and *doing* is small
- Usually a lot of variables must be measured (to achieve meaningfulness)
- Useful to step back and think about existing data
- We can use sophisticated statistical models to assess the fit between hypothesized associations among variables and observed data – econometrics…

Experimenting – for causal RQs

If I do *this*, will happen what I think would happen?

What happens, will it be **only** beause I did *this*?

- **manipulation** with an independent variable
- **measurement** of dependent variable, outcome
- **control** of intervening variables

Experimenting– features

- **Causal inference**
  - causality generalises better than association (coincidence)
- Ex is **interactive**, w/ strong emphasis on **context**
  - potential for further exploration, case study, qualitative work.
  - Ex demanding in terms of **control, interference**

- Ex ca be small, flexible
- Ex achieves **representativeness** more flexibly than eg. survey
  - theoretical generalization and replication
- Ex is demanding in terms of **current knowledge**

Specific applied designs

Applied, limited in generalizability, high usability

- Evaluation
- Action research

Ethics
The need for control can lead us astray

- Putting pressure on individuals to participate in experiments through coercion, or applying social pressure.
- Giving menial tasks and asking demeaning questions that diminish participants' self-respect.
- Deceiving subjects by deliberately misleading them as to the true purpose of the research.
- Exposing participants to physical or mental stress.
- Not allowing subjects to withdraw from the research when they want to.
- Using the research results to disadvantage the participants, or for purposes not to their liking.
- Not explaining the procedures to be followed in the experiment.
- Exposing respondents to hazardous and unsafe environments.
- Not debriefing participants fully and accurately after the experiment is over.
- Not preserving the privacy and confidentiality of the information given by the participants.
- Withholding benefits from control groups.

EXPERIMENTATION IN DETAIL

We need enough theory and focus that we can identify:

Dependent variable (DV)

- one or very few
- measured as precisely as possible

Independent variables (IV)

- one of few
- **manipulated** so that it can create as a large effect so that we can detect it, or estimate it with sufficient precision
- manipulation check – when it is not obvious we were successful in manipulating the IV

Intervening, extraneous, confounding, nuisance variables

- variables associated with both DV and IV
- observed(measured)
- controlled – by design, statistically

If the hypothesis is true and everything is executed properly…

- the IV will differ among perticipannt solely due to our manipulation,
- IV will correlate with DV
- the effects of all other variables on DV are controlled for

Thus, if there is a reasonable theory of IV having a causal effect on DV, we can consider the correlation between IV and DV as support fora causal effect.

What is *control*?

- Making sure the intervening variable does not bias our estimate of the effect of IV on the DV
- We are trying to prevent the intervening variables to correlate with IVs, DV or both
- Fixing the intervening v. – make it a constant so that it cannot correlate
- Randomize the IV so that it cannot correlate with the intervening variables
  - The intervening will still correlate with the DV but their effect will not affect the effect we want to estimate
  - Pairing, matching, balancing is a non random way of achieving this when IV is categorical
- Measure and control statisticaly – partial/part correlation

Internal validity – the concept of success

- A correlation between IV and DV is considered an internally valid evidence for the causal effect of IV on DV when we can argue tha all known and unknown intervening variables have been controlled.
- Then the experiment is said to be internally valid.
- Can we say that the differences in DV attributed to IV are really solely due to the differences in IV?
- Difficult, therefore it has the form of an argument open to discussion.
- Low to high (not „is“ „is not“).

Examples of experimental designs

1. Pretest – posttest single(experimental) group
   - DV is measured before and after experimental manipulation is done
   - IV has as many values (levels) as many there are different manipulations
   - Within-subject design
2. Two or more group posttest only design
   - DV is measured after experimental manipulation has been done in each group diffrently
   - Experimental and control groups terminology
   - Each level of IV is assigned to different participants – between-subject design

- Two-group pretest-posttest design
  - 1 + 2 – add a pretest in each group of 2 – mixed design
- Four-group Solomon
  - 2 + 3 – Two groups without pretest, two with pretest

Generic threats to internal validity

- History – anything that happened between pretest and posttest could have made/biased the effect
- Maturation - …even the naturally running processes in our bodies, getting older, hungry, tired
- Testing - …the act of pretest measurement itself could have affected the posttest or even reaction to the experiental manipulation
- Selection – unbalanced groups – any known/unknown differences between the groups could have made/biased teh effect
- Mortality – what if the reasons for subjects‘ leaving the study correlate with the DV?
- Regression to the mean - when the groups are made according to the level of DV (or related variable). Extremes have higher probability of change towards mean.

Not all experiments are **true** experiments

- True experiments
  - We have full control (down to each individual) over the manipulation of the IV
- Quasi-experiments
  - Manipulation is slightly limited by the fact that experimental groups have been formed prior to experiment. We still decide what level of IV will be assigned to each group
  - Informally – any experiments with obvious design weaknesses (1, 2)
- Ex post facto (post hoc) studies, natural experiments
  - Technically not experiments but correlational studies
  - The IV has occurred naturally, by itself, by someone else‘s choice – we just observe it
  - We use the terminology and statistics of experiments (and ideály dreamof an upgrade)

Field-experimenting

Issues

- Randomisation
- Manipulation
- Ethics
- Limited control

Perks

- **Ecological validity, generalizability**
- Less reactivity
- Availability of people

External validity - generalizability

- To different participants, populations
- To different settings
  - In teh context of lab vs. field – ecological validity.
- To different times

SUMMARY

- Designs are templates for research
- Most developed designs are for studies requiring most control

- For exploration with little theory – qualitative designs
- For descriptive and correlational purposes – observations, surveys, ex-post-facto
- For causality inference we need experiments