

Research Methodology (in Social Sciences)

↓

design → getting prepared

FUNDAMENTALS

Question?

- needs to be interesting
- has to contribute to literature, to existing body of knowledge

✓ sampling → selecting research subjects

- measurement → doing observations

- design ⇒ { CAUSAL? $A \rightarrow B$
RELATIONAL? $A \leftrightarrow B$
DESCRIPTIVE "A", "B"

✓ conceptual + theoretical
(kuramsal)

↓

usually no data collection or analysis

✓ empirical → data-based

subject - deneke

causal - nedensel

relational - ilişkişel

descriptive - betimleyici

sampling - örnekleme

validity - geçerlilik

reliability - güvenilirlik

Data Collection → final prep stage

- Data collection itself is creating data collection instruments (scale)

e.g.
✓ surveys, questionnaires

- All data collection instruments have to pre-tested & post-tested

- We have to test **VALIDITY & RELIABILITY**

3 types of validity:

① **EXTERNAL** → (about sampling) if the results could be generalized

② **INTERNAL** → whether A causes B (about design)

③ **SCALE VALIDITY** → the scale measures what it is intended to measure

Research has a process that has a beginning and an end:



- research question (should be ^{decisions} "focused" / "well specified")

↓
a field → subfields
(man&org.) (HRM, OB, OT, strategy)

↓
topic (entrepreneurship)

↓
"literature"

↓
research question

- a research report

↓
journal articles
(if you can publish)

American Journal of Sociology

American Sociological Review

Academy of Management Review

Academy of Management Journal

Administrative Science Quarterly

- book

- book chapters

- dissertation

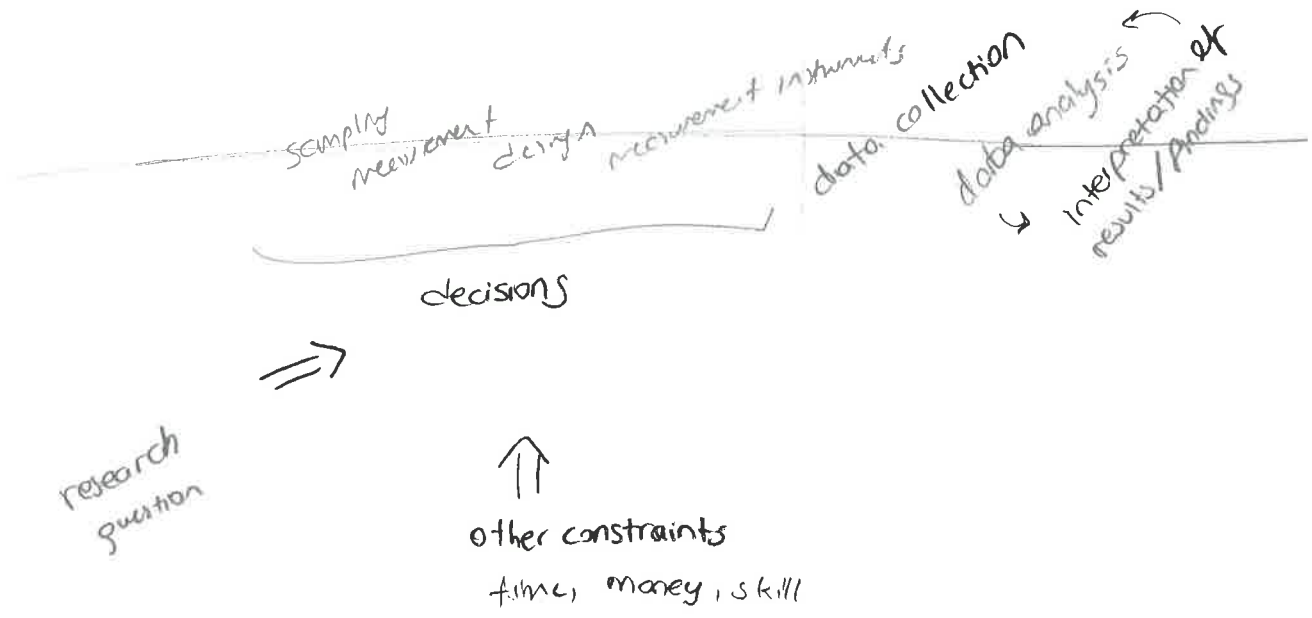
- technical report

- paper

- presentation

- term paper

- oral presentation



Research Methodology in Social Sciences

how

science - social science

- what exactly is the difference?
- implications for methodology?

Science

- nature?

Social science

- we deal with behaviour, relationships, social structure, culture, ideologies, demography

} not a good distinction

what

- social science
- natural science ("science") → older
 - more prestigious
 - emulated by most social scientists

↓ methodology

(positivist) non-positivist
 ↓
 most people
 most journals

⇒ science & social science are

both "intersubjective"

majority of people agree on sth

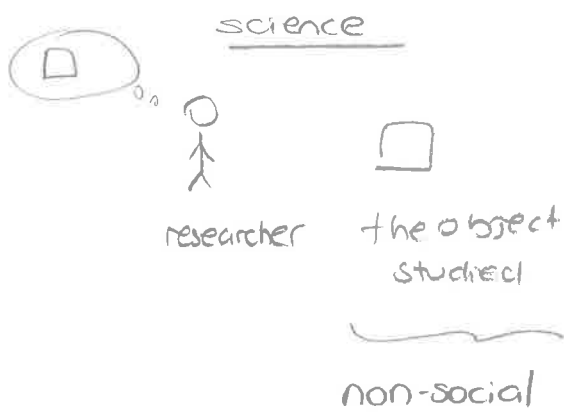
The difference between science and social science:

Science "emulation" social science

- social science is based on an emulation of science

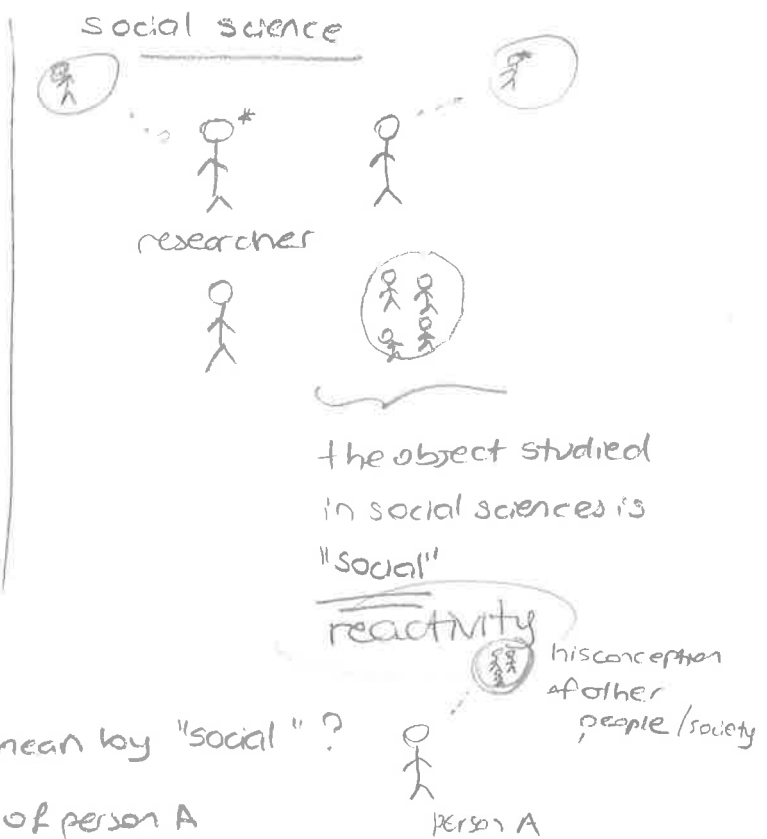
□ The distinction is based upon "METHODS"

↑
objects studied in how
they relate to the researcher
"different"



No reactivity

- The object does not have a conception of the scientist or other objects.



What we mean by "social"?

The identity of person A comes from its conception of society.

"Social" means someone or a group with a conception of society in mind.

- distinction is a matter of reactivity / the relationship between the two

Jargon

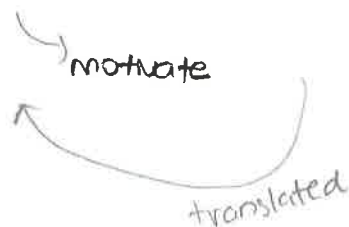
- things to be memorized
- things to be consistently used ✓ writing a paper, proposal, exam
- makes understanding the "topic" easier

Some Concepts

(1) TYPES OF RESEARCH

- The distinction between Applied and Basic Research
 - ✓ Basic research aims to increase our ^{scientific} knowledge on what we are studying.
 - ✓ Applied research addresses practical problems.
 - Practical problems usually motivate basic research

applied - basic



ex:

Title : _____ An application in _____ context

- Authors means "empirical"
(most probably)

- The distinction between empirical & theoretical (conceptual)

✓ In empirical research, there is only kind of data

collection & analysis

= "measurement"

= "observation"

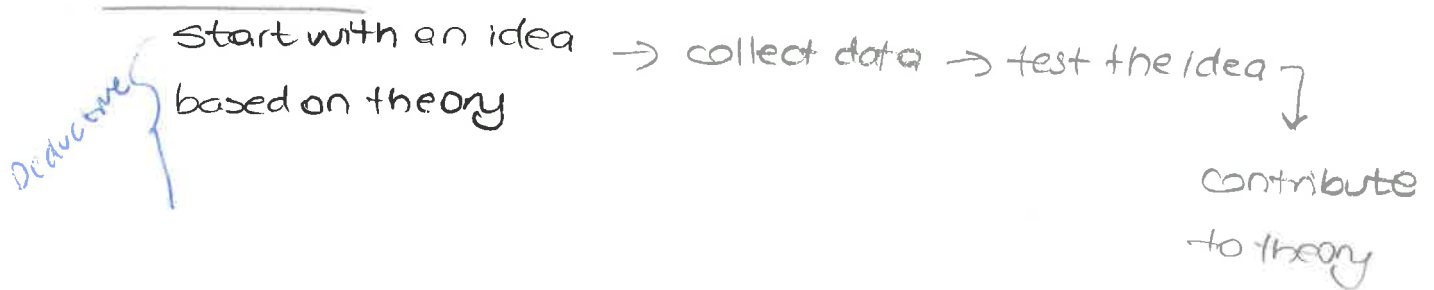
✓ In theoretical research, we have concepts and relationships between them; no data.

- A paper can be empirical and theoretical at the same time.

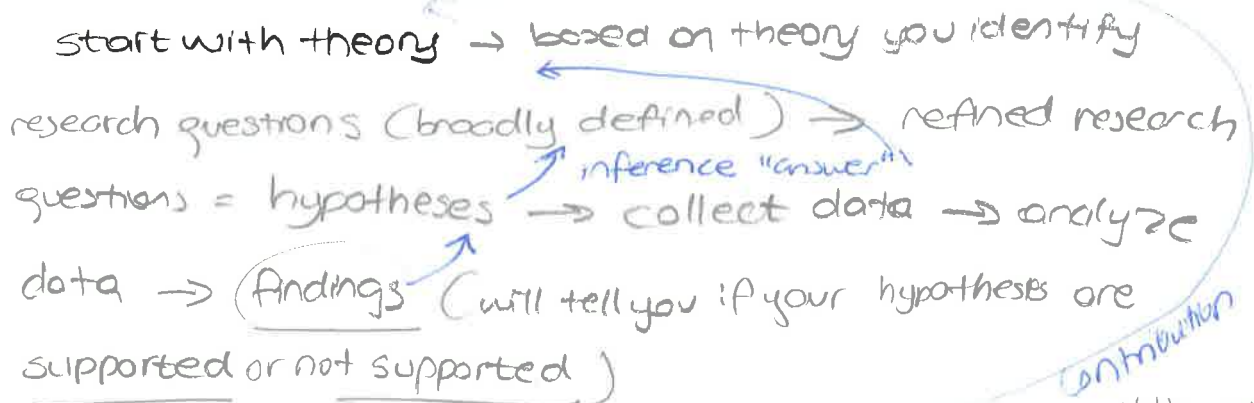
- No paper can actually be purely empirical

- The distinction between deductive & inductive research

Most Papers



Deductive



"discussion conclusions section"

- Inductive research aims at arriving at a conceptual understanding of what we are studying.

- The difference between the two: Deductive starts with a theory. You don't have one in inductive research in the beginning. You hope to have one at the end.

- The distinction between idiographic and nomothetic research

→ Idiographic research is about a single case. It could be a person, an organization, a specific group, a country, a business system, a specific transaction, a specific relationship etc.

→ (It does not necessarily generalize to other cases.)

→ Purpose: in-depth understanding of the case

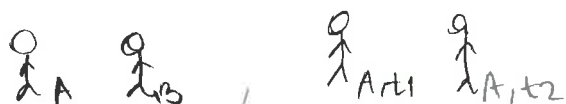
✓ Nomothetic research is about what is ^{valid for} true in most cases.

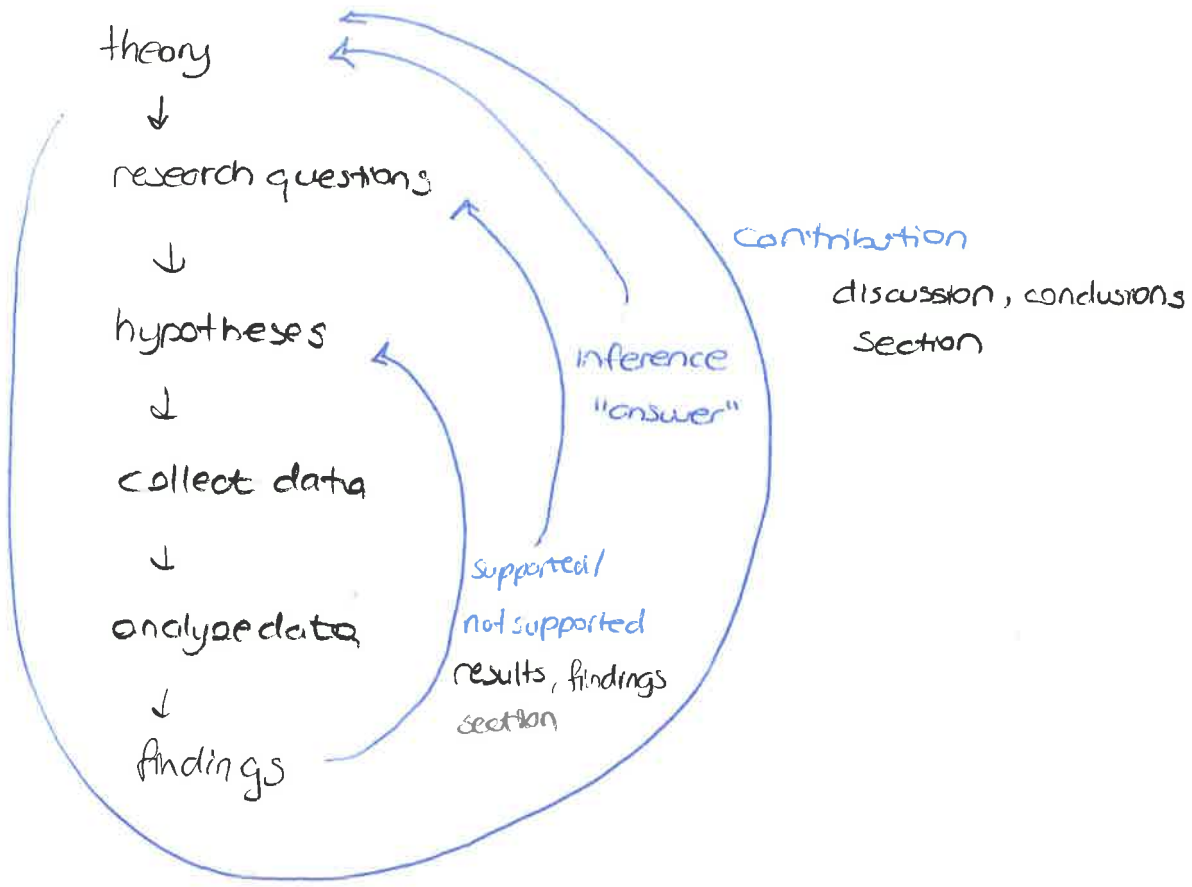
✓ "General" → does not mean true for all cases at all times in all places.

General means generalizable.

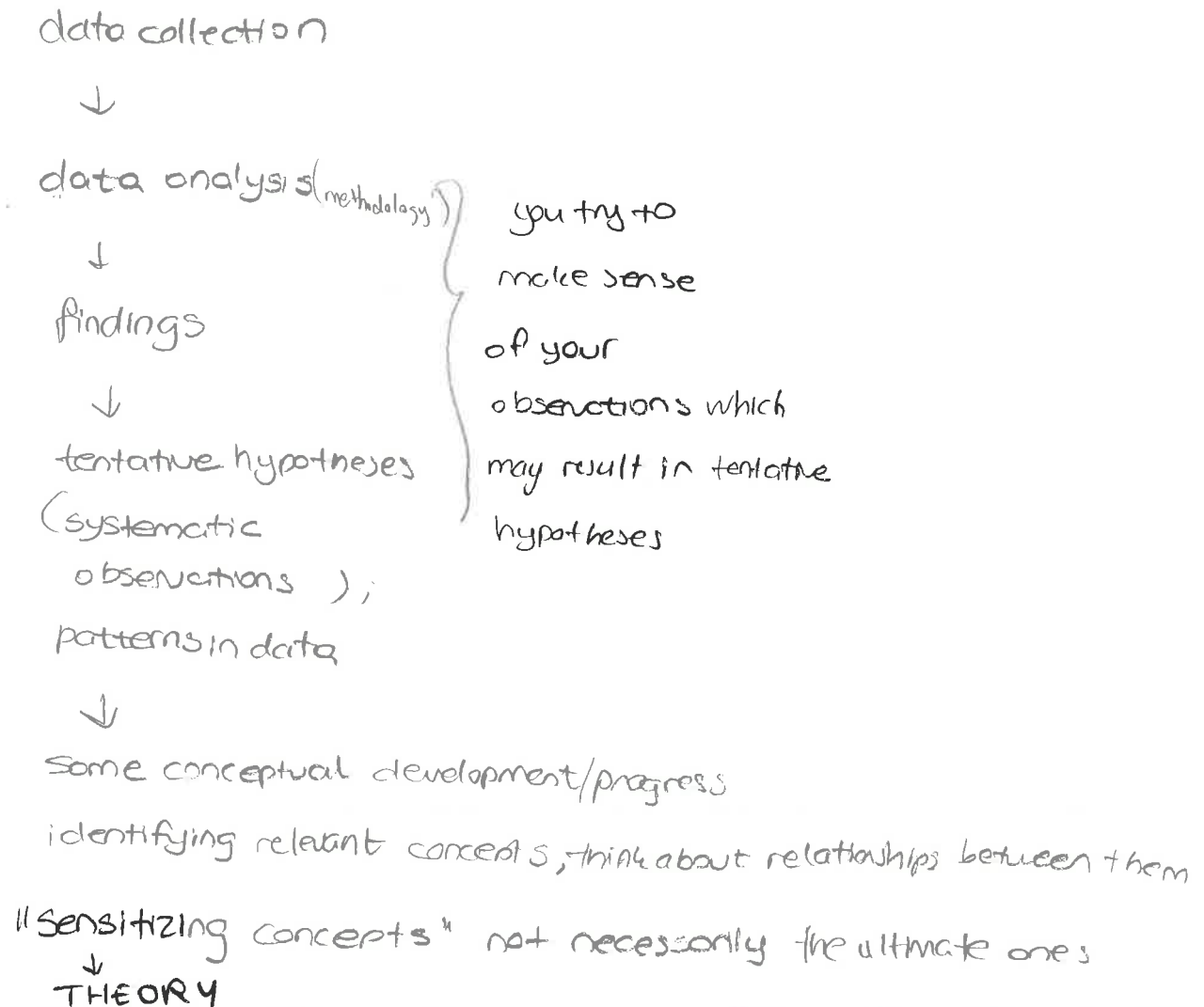
□ Keep in mind: Each social entity is unique

You can generalize but there is always going to be some error.





Inductive Research



Causal



- Does not mean ΔA ^{changes in A} cause ΔB or AA the only cause of ΔB .

✓ Changes in one thing is one of the reasons why some other thing changes...

"science" → change the world
"legitimate" "procedures"

"social science" → change the society

If you can find answers to causal questions

TYPES OF RESEARCH QUESTIONS

(1) descriptive → What are the properties of... ?

what someone is doing
how someone is feeling
where sb/sth is ...
the age, gender, education level...

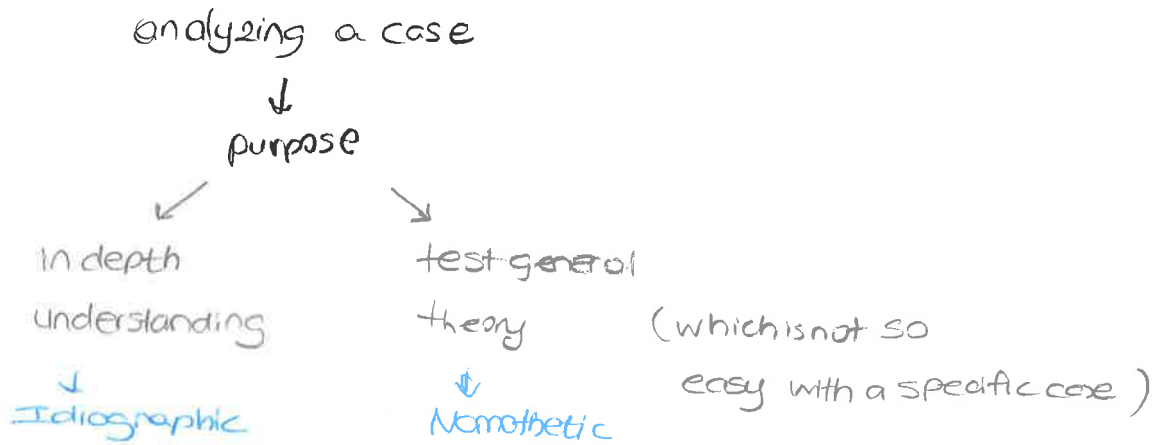
(2) relational → Is there a relationship between A and B?

What is the direction of the relationship?
between ...

(3) causal → Do changes in A bring about changes in B?

Does a change in A cause a change in B?
Can you increase employees' performance by increasing wages?

- Nomothetic research can help idiographic research.
- Mostly, we conduct nomothetic research in management and organizations.



Probabilistic how scientific statements regarding relationships or descriptive characteristics of what is studied.

- even when not explicitly communicated that way

→ Larger organizations have a harder time changing their structures. (= There is a negative relationship between size and likelihood of structural change).



I expect it to be true most of the times, in most places, for most organizations.

✓ This may be wrong; it may not be perfectly generalized.

- cannot be deterministic

} probabilistic

- We believe causal research questions are the most important ones.
- We usually end up doing relational research questions even if causal are our intention.
- Descriptive \rightarrow not very popular, but so fundamental

{ If your research is causal, it is at the same time relational.
 " " relational " " descriptive.

- Causal research is both relational and descriptive.

TYPES OF RELATIONSHIPS

(1) Correlational $\boxed{A} \leftrightarrow \boxed{B}$

When sth changes (a variable), some other thing changes in a systematic way. Then, there is a correlation.

When $A \uparrow$ B mostly \uparrow

or

$A \uparrow$ B mostly \downarrow

When $A \downarrow$ B mostly \uparrow

or

$A \downarrow$ B mostly \downarrow

Same direction = positive correlation

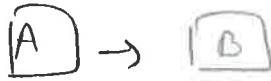
opposite direction = negative correlation

\Rightarrow No correlation; ϕ correlation
zero

$[-1, 1]$

-1 & 1 mean two things are perfectly correlated.

(2) Causal



Cause = ΔA

Effect = ΔB

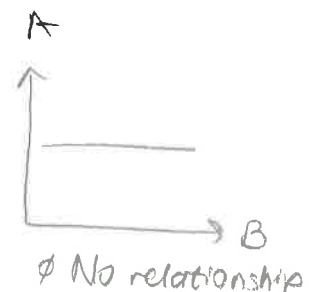
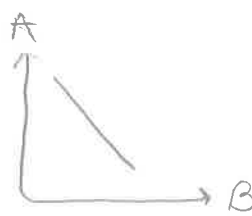
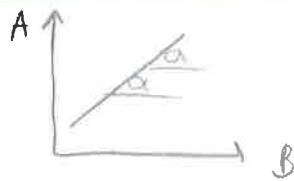
If changes in A bring about changes in B, then there is a causal relationship.

— Causality implies correlation, but correlation does not necessarily imply causality.

! If your research is correlational, do not make strong claims regarding causality.

PATTERNS OF RELATIONSHIP

(1) Linear



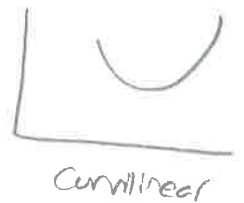
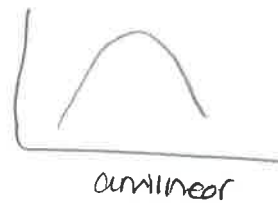
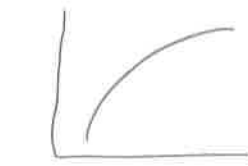
In a linear relationship, one unit change in one of your variables is related to a constant change in the other associated with unit.

→ In social sciences we mostly investigate linear relationships or act as if relationships we investigate are linear. (usually a limitation we acknowledge)

(2) Non-linear

(curvilinear)

(discrete, step function)



curvilinear

curvilinear

a step function

hypothesis

— A hypothesis is a statement of the relationship between two variables.

Examples:

H: Wages are related to performance.

H: Wage changes cause performance changes.

H: An increase in wages increases performance.

H: There is a positive relationship between wages and performance.

} directional

✓ variables

✓ relationship between them

✓ direction (optional)

Hypotheses may be directional or nondirectional.

Null - alternative hypotheses


- In social sciences we use alternative hypothesis. We test null hypothesis.

Null hyp. → covers all possibilities not covered by the alternative

↓
useless

✓ Some people believe refuting the null is how we support the alternative.

Data: a set of variables and attributes

 age = 28 → data set
↓ ↓
variable attribute

Variable: anything that can take on different values

- Each of these values has a name "attribute"

ex: age

Time in Research & Data sets

Time in Research & Data Sets

Time : Cross-sectional research / longitudinal

distinction applies to research as well as data.

longitudinal

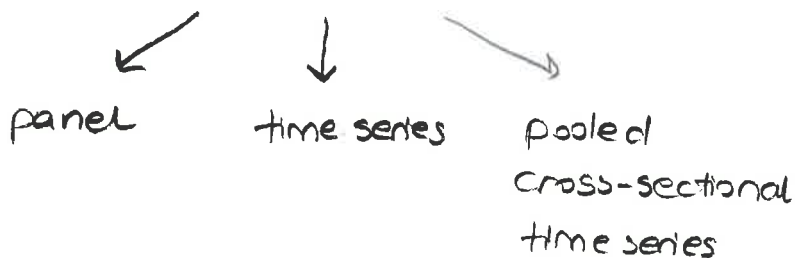
research → process (sth that unfolds over time)

data → observations at multiple points in time (each unit is observed at least twice)

example:

<u>Student</u>	<u>Age_{t1}</u>	<u>Age_{t2}</u>
—	—	—
—	—	—

— Longitudinal data can take different forms:



Cross-sectional

research → description / explanation of existing differences

data → observations at a single point in time (each unit is observed only once)

example: list of the students

<u>Student</u>	<u>Age</u>	<u>Gender</u>
—	—	—
—	—	—
—	—	—

panel

many units, few points in time

time-series

few units (usually one)

observed over multiple points in time

year inflation rate in Turkey

1923 α

1924 α

1925 α

⋮ ⋮

2013 α

panel cross-sectional time-series

multiple units, multiple points in time

<u>year</u>	<u>group IDs</u>	<u>α</u>	<u>α</u>
1	1		
2	1		
3	1		
4	1		
5	1		
1	2		
2	2		
3	2		

When sample size increases, errors will decrease.

— The larger the data set, the better.

unit of observation

entities you collect your data from

- university system, organization

✓ comes from your research question

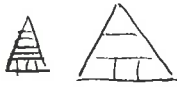
level of analysis

- how you explain the entities you refer to as you explain

unit of observation

→ universities

structural complexity



you are investigating structural complexity of universities.

for ex. why some universities have higher division of labor?

How you measure differences in structural complexity?

measure: # of faculties



- usually in a paper, it is in the "methods" section

level of analysis

What your theory is about?

possible explanations you can make

- environment → uncertainty
- university → size (# of academics)
- subunits, members → entrepreneurial minded rector



- presented in the "theory" section

ontology

- What exists?
- What we can study?
- What can be the object of study?

Two schools:

✓ Realism **

✓ Phenomenology

✓ clear-cut distinction between theory and empirical world

✓ you can take an unbiased observation of your theory.

Qualitative Approach

- general approaches to research related to ontology and epistemology

- Qualitative / Quantitative approaches → not data utilized in research (text / number)

about data

As you transform one form of data to another, you lose information.

epistemology

- How can we know?

Two schools:

✓ Positivist *

✓ Non-positivist

→ - empirical observation

↳ - Theory is not independent of observation. You cannot validate theory as positivists

✓ All observations are theory-driven. "I'm generating observations as I develop theory" They are not saying there is ^{not} an external world out there. Measurement will be driven by definition.

Quantitative Approach

the distinction is about :

Qualitative Approach → non-realist & non-positivist

Quantitative Approach people → realist & positivist

What is science (inc. social science) ?

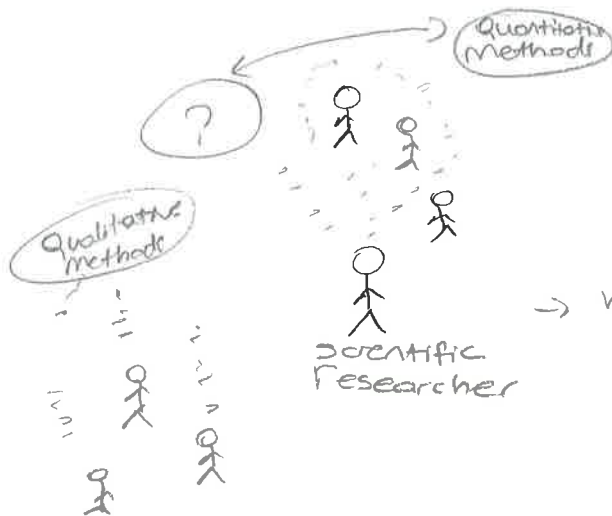
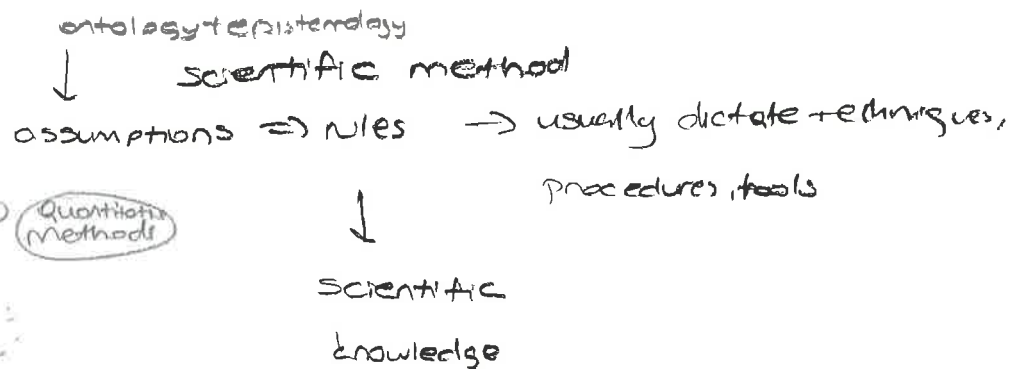
What is the source of "scientific knowledge" ?

Scientific knowledge comes from scientific method
(is a product of)

What is Scientific method ?

ontology + epistemology } shapes assumptions

(assumptions & rules that usually dictate techniques, procedures, tools



→ wants to know what he is doing scientific or not?

he should check if he

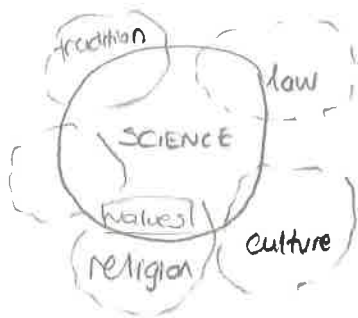
You take your side.
You must be able to acknowledge which rules you are following.

the only way he can make a claim of being scientific is checking the methods he uses are methods of the scientific community

✓ Scientific knowledge } a creation of our minds
methods

- If you go by their methods, people support those methods will say your work is scientific.

ETHICS OF RESEARCH



harm → Physical
emotional

ethics is most basically a definition of obligations towards

(1) other researchers ⇒ you are using other researchers' resources and you are providing them with resources.

i) give credit/acknowledge when you use other people's resources ⇒ cite

(2) subjects of research

not citing → plagiarizing
✓ you are claiming the idea for yourself

(3) supporters of research (society in general or funding bodies)

research → reputation
↓
wages/funds/
→ jobs

cont. →

(1) other researchers

what are your obligations towards other researchers?

ii) do not fabricate data

iii) do not falsify findings \Rightarrow changing results to make research interesting or support hypotheses ...

iv) do not do salami slicing \Rightarrow writing couple or ideas in multiple papers that can actually be presented in a single paper. By doing that, you're inflating your number of publications. It is not ethical.

v) do not engage in complimentary authorship \Rightarrow

vi) do not have other people write for you \Rightarrow

(2) obligations towards research subjects

(i) do not harm

(ii) do not force

(iii) ask for consent

- participation should be voluntary

(iv) inform \rightarrow give as much info. as possible

\downarrow
but also make sure you're not damaging your research

✓ why they were selected

✓ your topic ✓ why is that a relevant topic

✓ why is it important

✓ procedures ✓ risks

} informed consent

(v) promise confidentiality => keeping personal information secret

(vi) if possible, ensure anonymity -> do not collect personal information

(vii) right to service (medicine)

(3) obligations to sponsors of research

(I) generate knowledge that is useful to them or serve their mandate

(II) do not waste

SAMPLING

Empirical research

data

(1) subjects/entities that will be observed (sampling) (sample)

(2) concepts that will be measured - how they will be measured

- development and administration of "scales"

(3) design/modeling

relates to whether your research is descriptive / relational / or causal

Any bad decision regarding these issues results in paper rejection.

good idea + good methodology

r/q : What makes companies successful? → consider it basic research

✓ We will conduct empirical research

sample : select a group of very successful companies

(e.g., the most successful five in each industry)

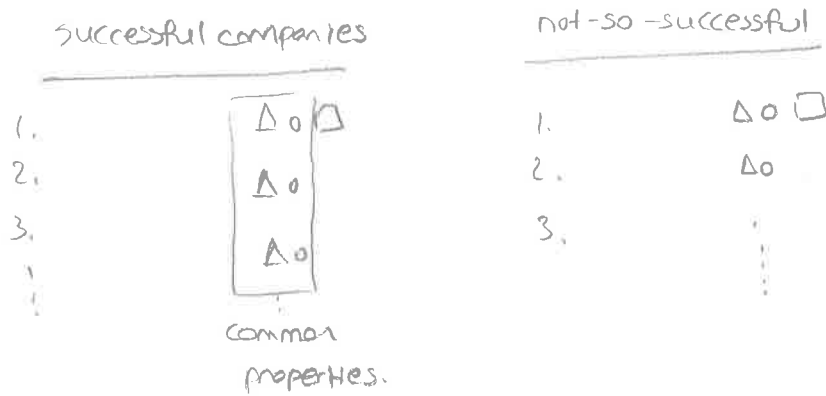
↓
to observe their common characteristics

Then, make a list of factors that make companies successful.

Problem: Sampling on the dependent variable

↳ you try to explain the success, you end up selecting successful organizations. But sampling on the dependent variable results in biased sample.

biased sample → biased estimates → biased interpretation



But you cannot conclude that those characteristics make them successful.

Sampling → selecting "subjects" that will be observed.

- Subjects you will observe depends on your r/q

→ entities to which your r/q refers

↳ population → { = sampling frame }
↕ they may be different sets of entities

Sample → entities you end up observing

- The sample is always a subset of a population (but not any subset of the population)
- But any subset of a population is not sample.

r/q: what makes companies successful?

no temporal restrictions

no spatial restrictions

no restrictions regarding company characteristics

Such as form of incorporation, age, size, industry, ...

Population → the set of entities that your r/q refers to

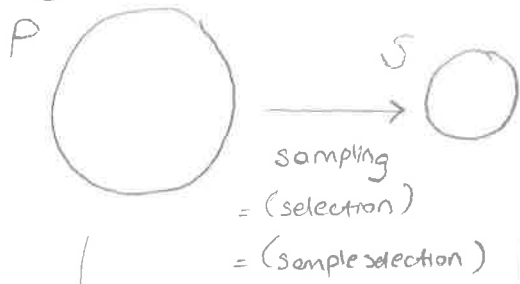
Population frame → a list of all population members

Sample → entities that you ^{actually} observe

Element → any member of the population

Subject → any member of the sample

mostly :



This is what
your questions
about

This is where
your empirical findings come from

generalization ?

↓
Do these findings
generalize to the population?

generalizability = external validity

random \Rightarrow rule-based

randomly \Rightarrow

a certain procedure that
prevents characteristics of
population elements interfering with
selection

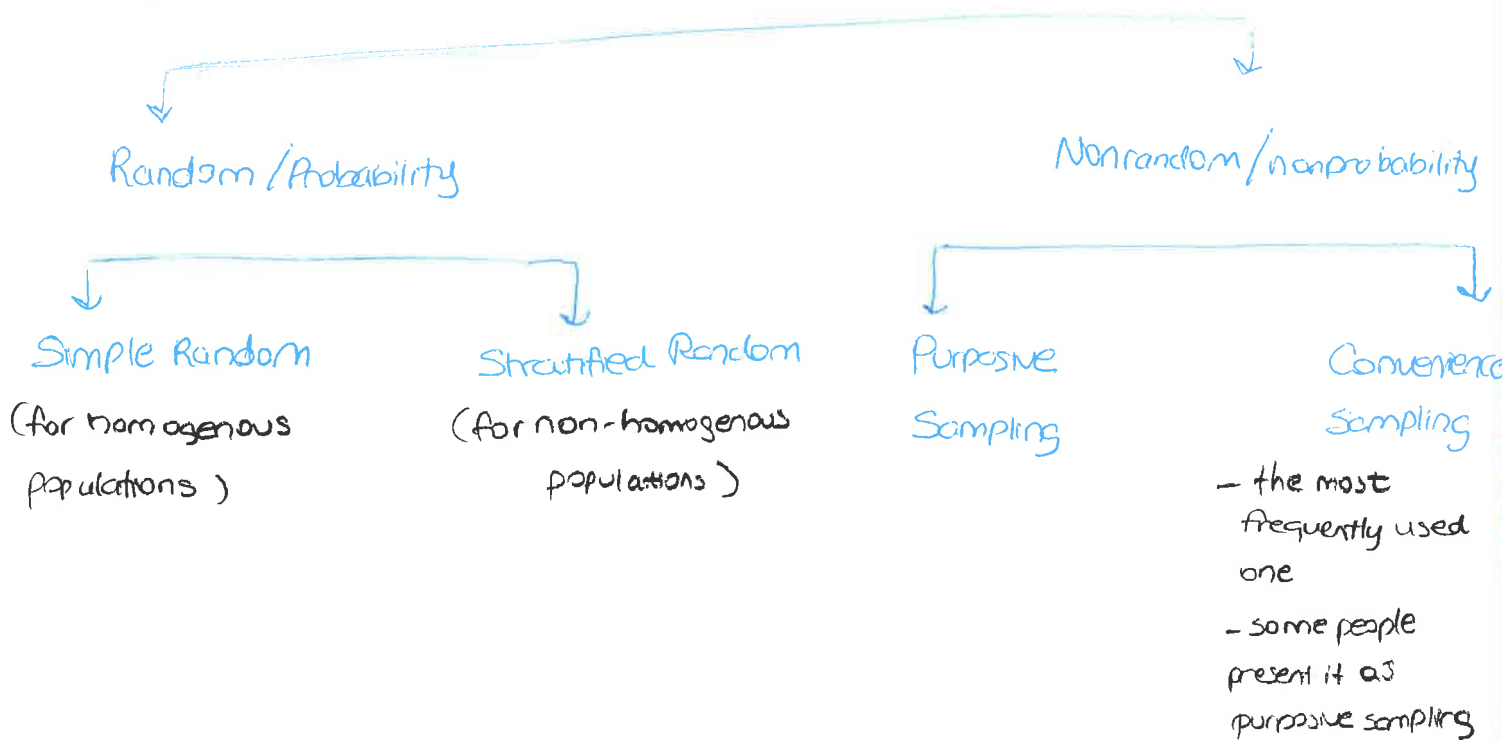
Central Limit Theorem = if you ^{randomly} select a sufficiently large

of elements from a sufficiently large and homogenous
population; then the sample that is constituted by the
selected elements will give unbiased estimates of population
characteristics.

Unbiased \Rightarrow no systematic error BUT nonsystematic error
may exist

We cannot totally eliminate bias. We "minimize" bias
by means of sampling methods. \Rightarrow randomizing

SAMPLING



SIMPLE RANDOM

- develop a randomization procedure
- select a sufficiently large number of elements
- observe all subjects
- findings will be generalizable or will have high/satisfactory external validity
- To administer this method, each member of the population given an equal chance for being included in the sample

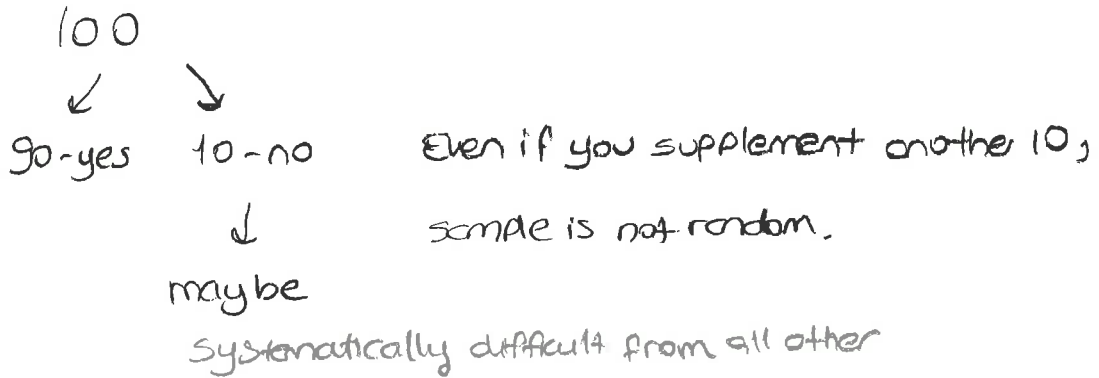
Why this method difficult?

- (1) you need a list of all population members
- (2) some selected elements may refuse to become subjects → one kind of self-selection

If you don't have a list of all population members, then they don't have an equal chance for inclusion. \rightarrow Sample biased

Even if you have a list, some of them refuse to participate.

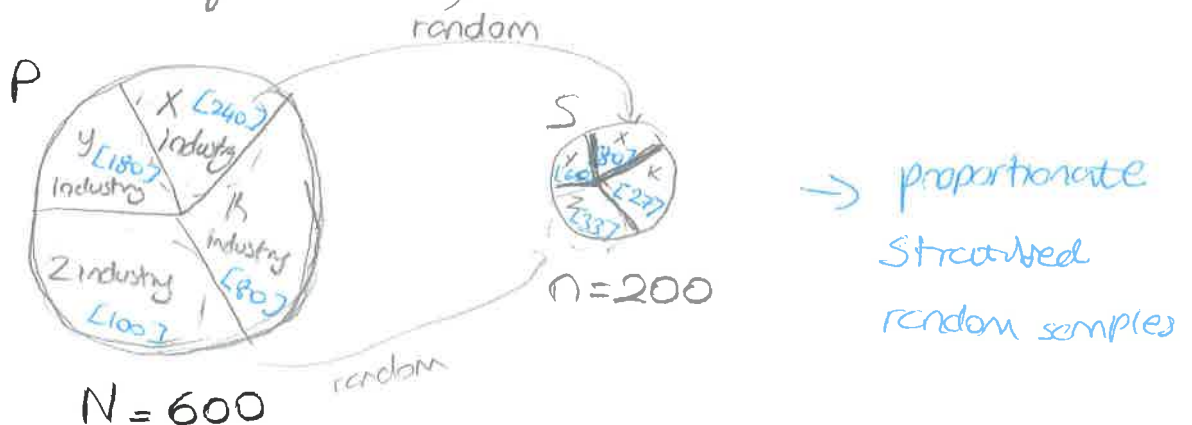
\rightarrow sample biased



STRATIFIED SAMPLING

stratum/strata \rightarrow group(s)

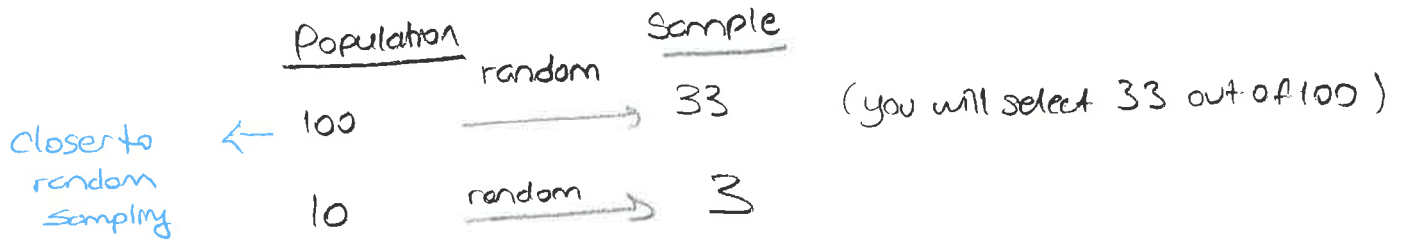
— used for nonhomogenous populations (groups that are systematically different)



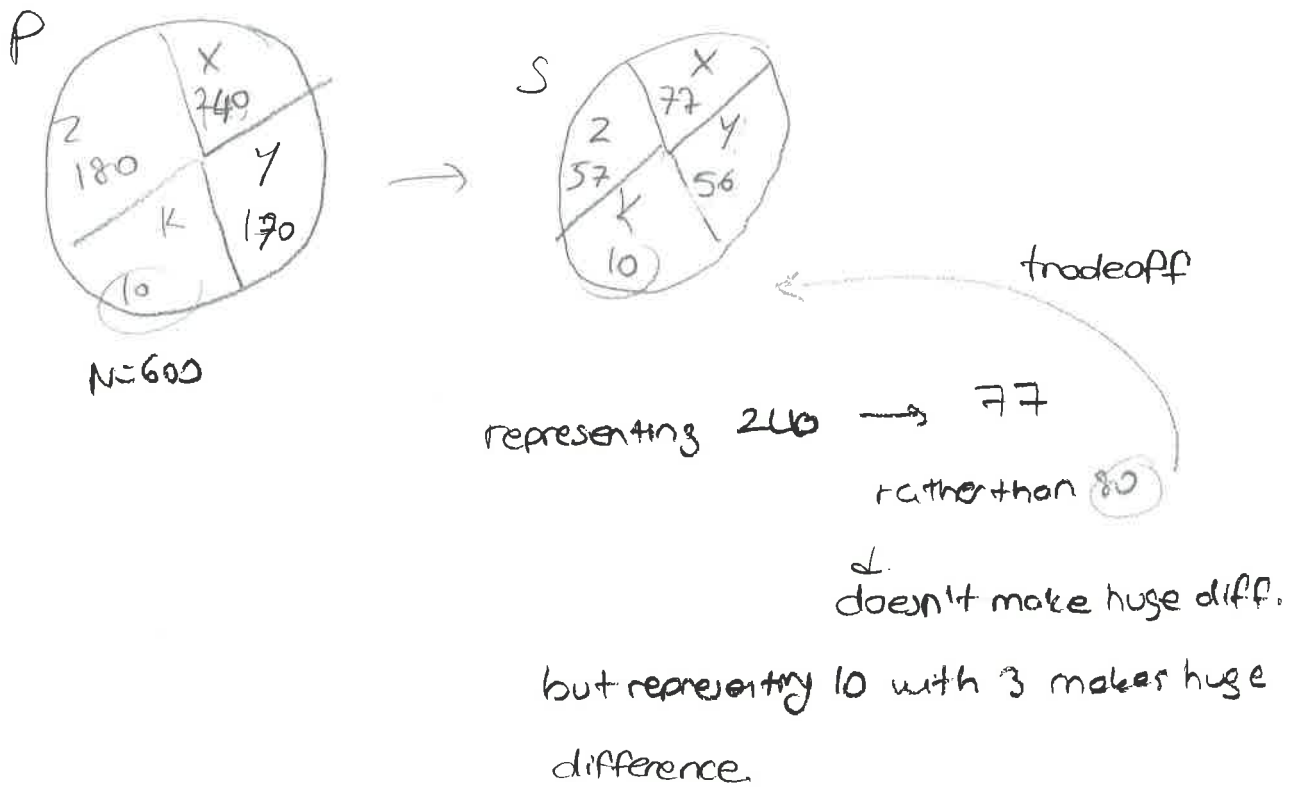
The reason for using stratified sampling \rightarrow when some groups are of small percentage of the population. It is likely that one category will not be represented



- Judging by Central Limit Theorem,



↳ you cannot represent 10 by 3. → include all of them!
Not random sampling!



rules = flexible }
contested }

- You use random sampling when you ^{have to} need generalizability
- Random sampling almost impossible in social sciences.
- Random samples give representative samples
- Representative sample ⇒ your finding will be generalizable

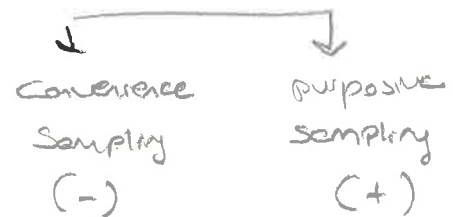
- Sometimes it is good enough to have an informative sample, not necessarily a representative sample.

- In most cases representative samples will be informative, but not always.

For example, blue-collar workers may not comprehend/answer your questionnaires. Then you might want to go for an "informative" sample.

Non-random Sampling

- No randomization procedure is used



Convenience sampling

- Subjects conveniently available are selected,

convenient: easy to access and easy to obtain measures from

- Not representative but easy and cheap

- Convenient samples reported as purposive samples

Purposive Sampling

population ecology of organizations

↳ observe all members of an organizational population from the time the first organization of the population was established onwards.

Turkish labor unions

1947 → 1980

+ 3500

/ hospitals

/ banks

⇒ any sort of

organization with a common

- We do not present convenient samples as convenient but purposive D

↳ research for a purpose

r/q

data: findings from which can be answer ^{used to} the r/q

You should be presenting your sample as a source of data that will serve your purpose even if it was totally convenient.

purposive sample → a sample that helps you generate data with which you can ^{use to} answer your r/q

r/q: political opportunity, organizational infrastructure, and organi-

zational founding

1947-1980

Sample: a sub-set of Turkish labor unions (those that were established in IST, ANK)

Justification for sampling: (1) selecting TR labor unions

(2) selecting IST and ANK labor unions

"organizational founding"

"political opportunity" → changes in the political environment

"organizational infrastructure" → ^{local} industries

- more than half of the labor unions established in this period established in IST & ANK

- tradeoff between information and bias

How big an issue?

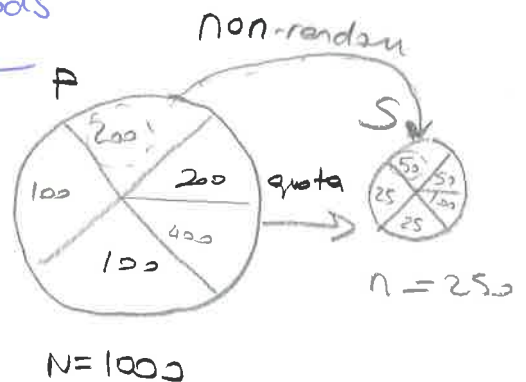
Standards \leftarrow methodology
 \leftarrow existing literature

Types of Purposive Samples

Purposive Sampling methods

Quota Sampling: strata in the population

- The difference from a stratified sample is the non-random selection procedure,



- Research compares conducting "polls" before elections use quota sampling
- It could be a biased sample. There is always a huge error.
- Way to deal with error \rightarrow randomization, but it's not feasible
- another way to decreasing error \rightarrow increasing the sample size

Using no randomization procedure + small sample size would result in tons of error

Judgment Sampling \rightarrow observing the most informative cases or asking questions to subjects who could actually answer questions.

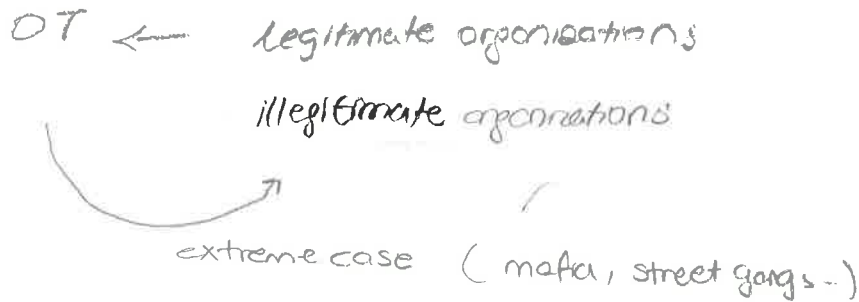
✓ Researcher will not go for a representative sample,

one specific form of judgment sampling is:

"Extreme Case Sampling"

Extreme Case Sampling

- You need information that comes from extreme cases.



Snowball Sampling



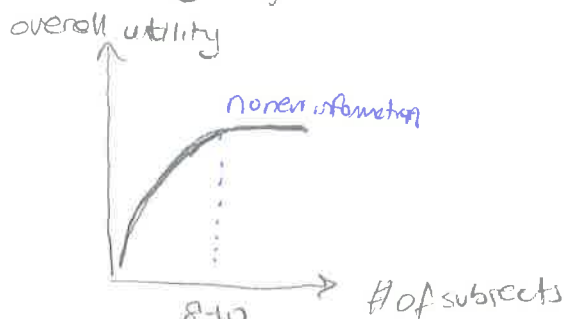
You ask them to name 3 other people ...

Over time, you end up with a huge sample.

- You use it especially when you have no idea of where the information may come from.
- Another reason => You are doing research on relational "networks."

Sequential Sampling

- When marginal utility of an additional subject declines.
- especially in qualitative research



usually for interviews => maybe 30-40

methodology

Sampling

(1) description of the sample justification

Overall 4000 words

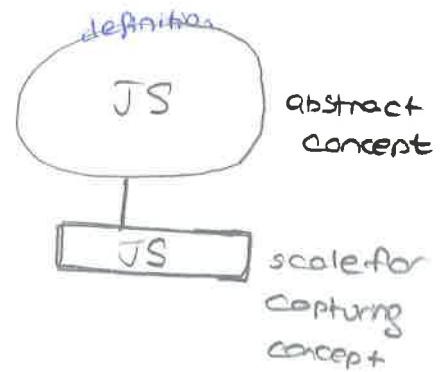
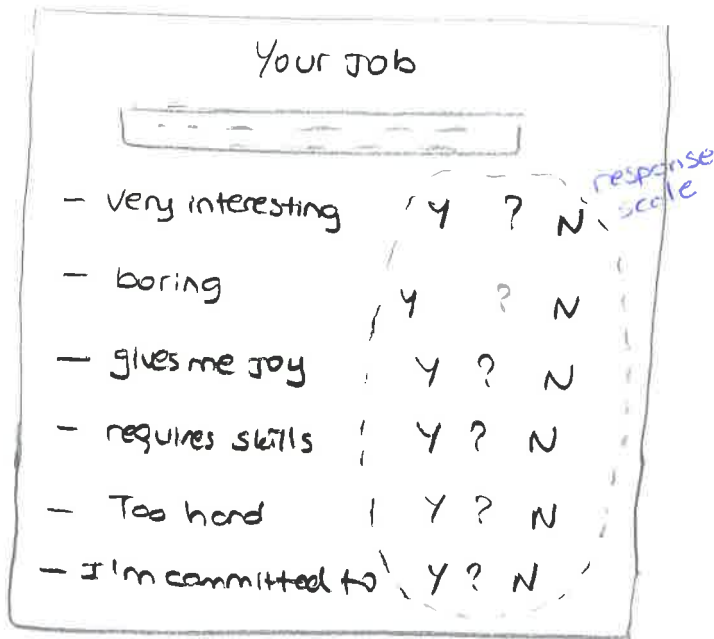
methodology ~ 3000

measurement ~ operationalization

(Rendering abstract concepts measurable)

q = Jobsatisfaction :

a questionnaire will be used.

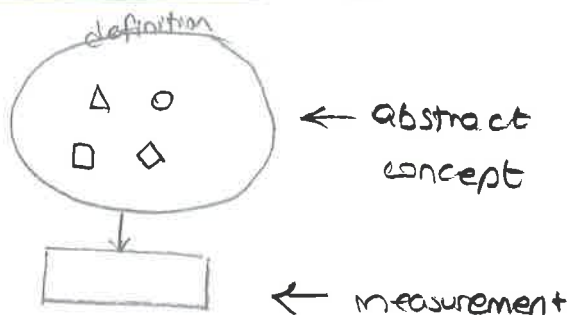


Is this job satisfaction scale valid or not?

#1 Problem : If you do not have a definition, you can not assess if the scale is valid or not.

✓ Face Validity "at face"

Validity of a scale / measurement instrument



scale (a questionnaire, an interview protocol etc)

Validity : the extent to which the scale measures what it is intended to measure

(IV) Transform each subdimension into observable characteristics
(turn them into variables)

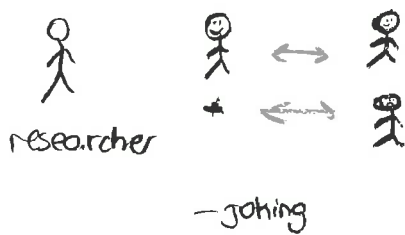
(V) develop the scale

The scale will have items that can capture these observable characteristics

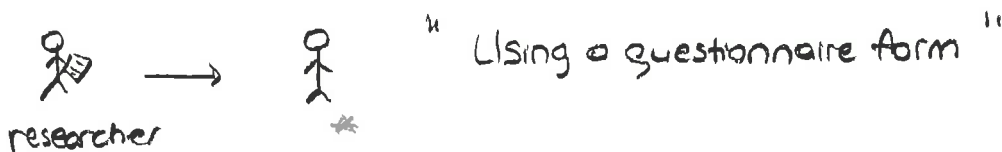
(VI) Test the validity of the scale

job satisfaction = method will be observation through the eye

↳ one specific aspect of JS = satisfaction with coworkers



→ You can have a questionnaire

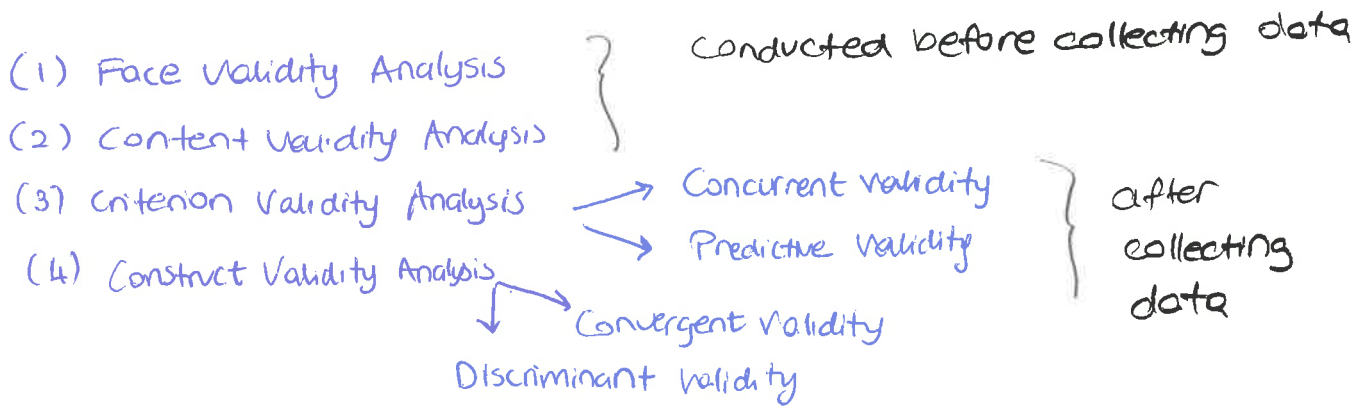


JDI → job descriptive index

↳ have to make sure the items relate to job satisfaction

MEASUREMENT - Validity / Reliability

Validity → the extent to which our scales measure what they are intended to measure



Reliability → the extent to which the scale gives error-free measures

- (1) Internal consistency (≈ reliability)
 - (2) Temporal consistency
- after collecting data

FACE VALIDITY

"experts" (academics, students, actors from the field)
Judgment ↓
quick look

very simple
not costly
may provide very valuable information

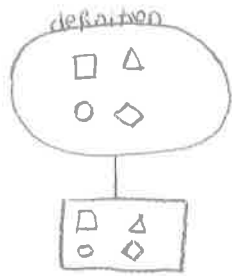
BIAS - even if you know a lot about the topic you are studying, you may be biased.

- Just as a precaution, have a couple of people to look at your measure.

Judgment: (1) ^{possibilities} seems irrelevant / make a note of why you consider this
(2) seems like relevant ^{response irrelevant and discuss it with the judge!}

↳ revise your scale → discuss with the same person & ask if you
↳ let someone else see it again ^{revision addresses their concerns}

CONTENT VALIDITY



job satisfaction



J/S is emotional
evaluation aspects of work

core

- a definition describes the domain of your construct. It specifies what your construct includes/excludes.
- every definition has to make references to other constructs whose boundaries are drawn by their own definitions.

↓
refer to other constructs

↓
definition

- define your constructs the first time you use them to clarify what you are studying

- If there is a match between your construct and scale, then your scale has content validity

✓ not underrepresenting

✓ not overloaded

- First develop a panel of experts → knowledgeable about your construct and measurement

- These people first work on the definition & then work on the scale. It takes time.

When you will do research on a topic, you can either use ^{- established} scales or ^{- new scales}

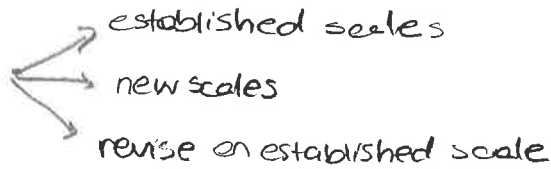
If it fits your purpose, and the definition in your mind, and context, you can use established scales.

↓
sample
cultural and sociopolitical
environment

✓ If there are established scales that you can use, do not develop your own.

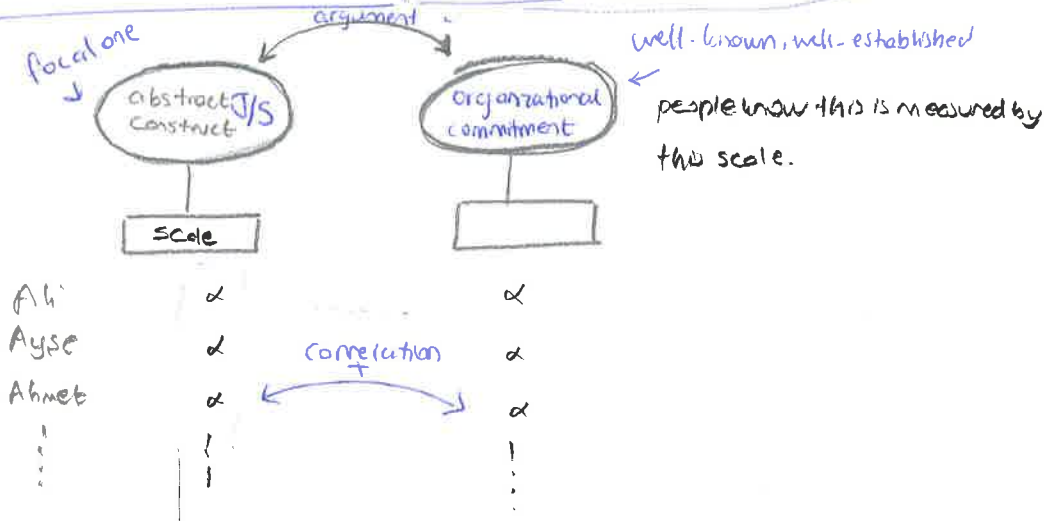
- If you don't have any established scales, you have to develop a new scale.

- Say there is an established scale but you have to revise it, you have to go through the stages of validity again.



✓ an established scale → a scale that has been published in good journals many times. (repeatedly used)

CRITERION VALIDITY Concurrent Validity Analysis



definition = J/S is emotional evaluation of aspects of job (e.g., likes/dislikes)

definition = org. commitment is - wants to remain as a member of the organization, is proud of his organizational membership...

- If conceptually job satisfaction & org. commit. is related, there has to be positive correlation between two.

- If I get a negative correlation (as opposed to my argument), criterion validity is not supported.

+ = C/V supported

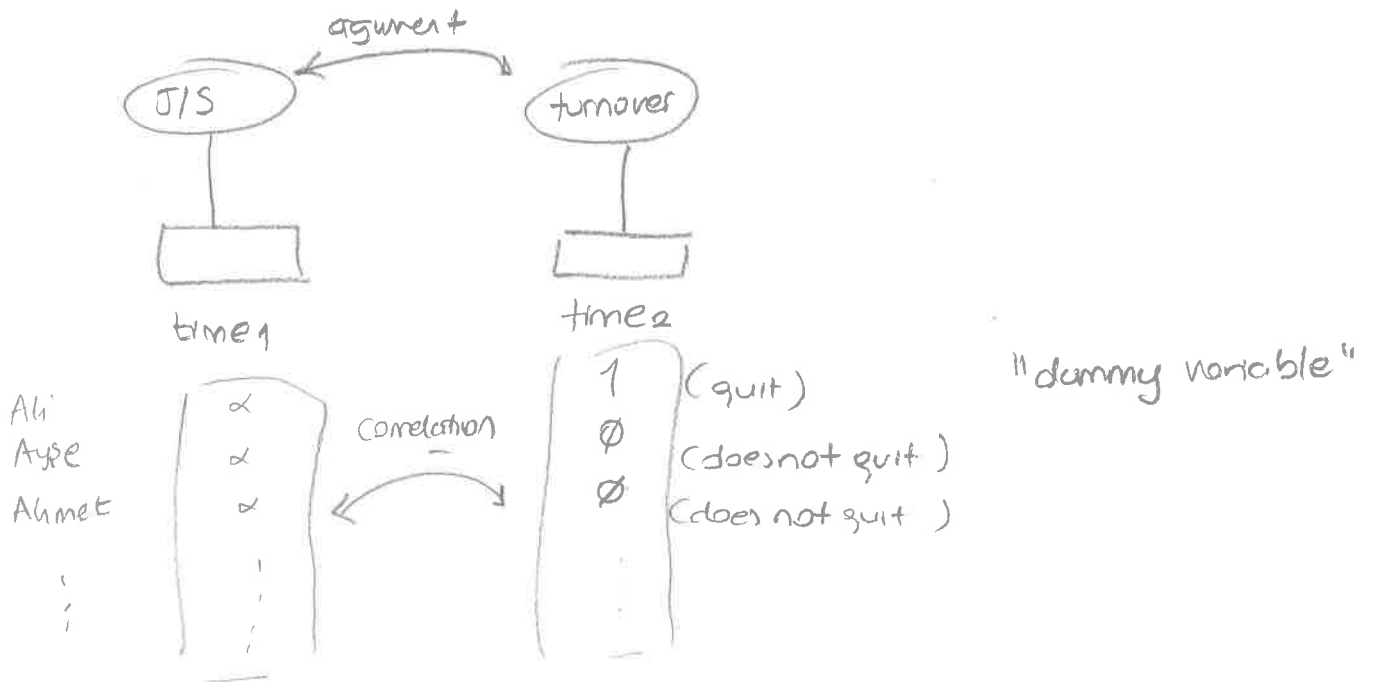
⇒ Scale seems to be valid

- = C/V not supported

∅ = C/V not supported

} which means there is probably
smth wrong with your scale

Predictive Validity



= There has to be a negative correlation between J/S and turnover.

- Conceptually speaking, I should be able to predict turnover based on job satisfaction.

- = p/v of the job satisfaction scale supported.

↓
negative correlation coefficient

✓ for this sort of analysis, you need longitudinal data.

* Why do we say validity is supported based on empirical analysis?

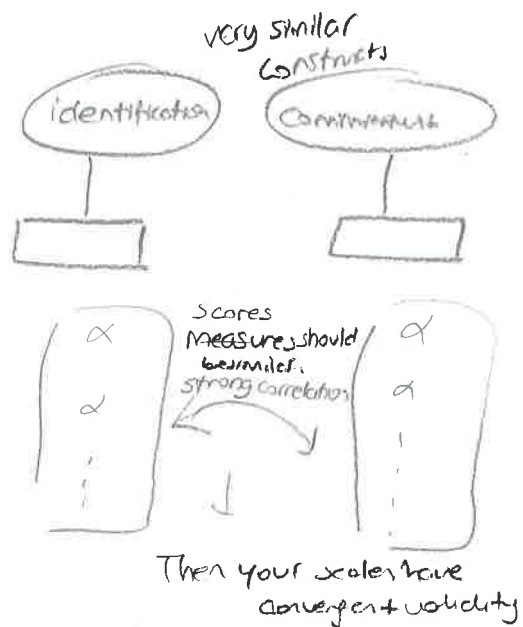
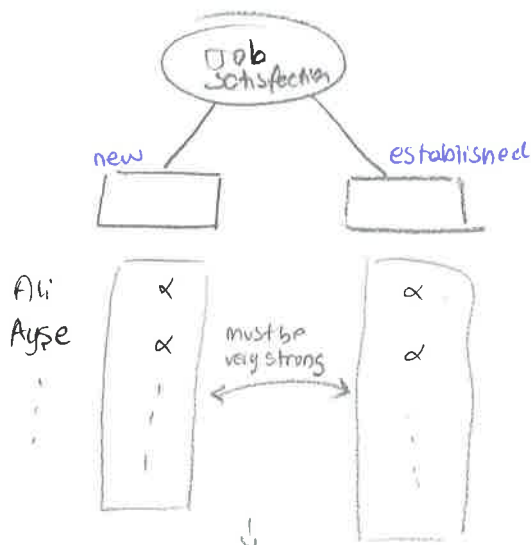
You can't say based on statistical relationship if your scale is valid or not.

You say just supported or not

CONSTRUCT VALIDITY

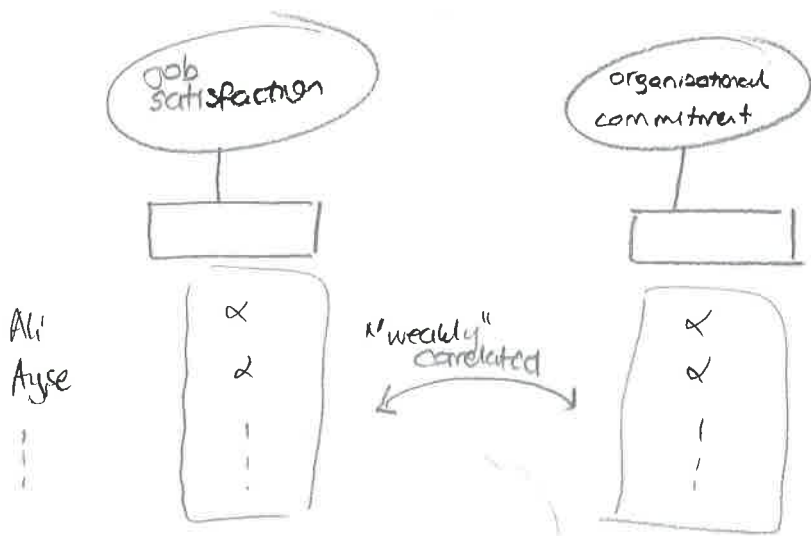
→ convergent, discriminant (konvergenz, diskriminanz)

Convergent validity



If this is the case (the scores you obtain converge / scales are equally successful in capturing the construct) → new scale's convergent validity supported.

Discriminant validity



different / related
Yes Yes

If My scales are capturing different constructs, they should be weakly correlated.

Literature = σ 's $\xrightarrow{.40}$ o/c average estimated correlation coefficient

criterion validity = the correlation between σ 's & o/c must be strong. $\sim .40$ or above

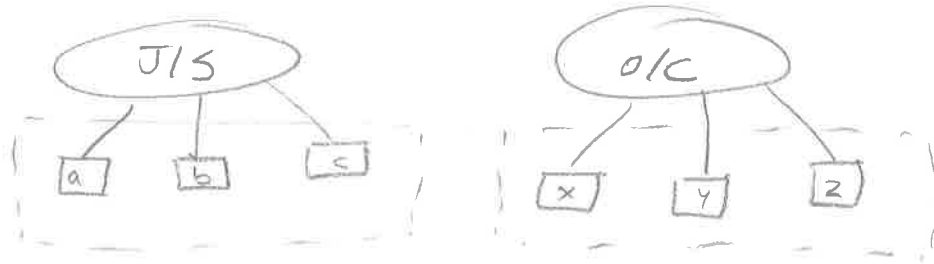
Construct validity = must be weak $\sim .40$ or below

"Weak" and "strong" depend on literature.

$\rho = -.40$

strong enough to support criterion validity

weak enough to support construct validity



Multiple items trying to capture the same construct

Ali

Ayşe

	α	α	α	α	α	α
Ali	α	α	α	α	α	α
Ayşe	α	α	α	α	α	α
⋮	⋮	⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮	⋮

Factor Analysis

P	a	b	c	x	y	z
a	1	S	S	W	W	W
b	S	1	S	W	W	W
c	S	S	1	W	W	W
x	W	W	W	1	S	S
y	W	W	W	S	1	S
z	W	W	W	S	S	1

Discriminant validity supported!

If two items capturing the same constructs, they should be strongly correlated, but strong correlation may unlikely to be perfect corre.

RELIABILITY

→ Is your scale error-free?

Internal Consistency

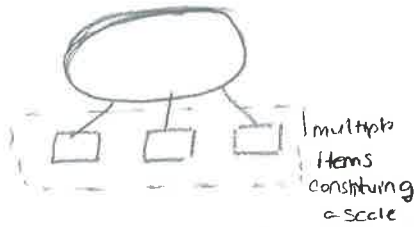
Temporal Consistency

Inter-item Consistency

Split-half Reliability

Test-retest Reliability

Parallel-Form Reliability



Ali	α	α	α
Ayşe	α	α	α
...

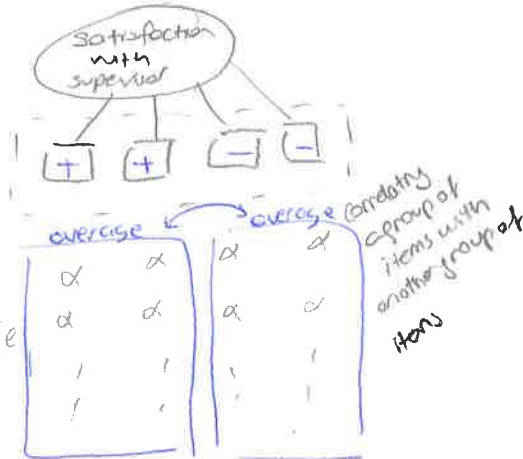
They should be positively correlated.

✓ Is the overall inter-item correlation strong enough?

Cronbach's alpha

(Alpha) $> .70$

(standard in literature)



Ali	α	α	α	α
Ayşe	α	α	α	α
...

Your Supervisor

- Nice person Y? N?
- Helps me when I need Y? N?
- Nasty Y? N?
- Waste of time Y? N?

item = positively worded

negatively worded

Some people answer positively worded items in a systematic way and negatively worded items in another systematic way.

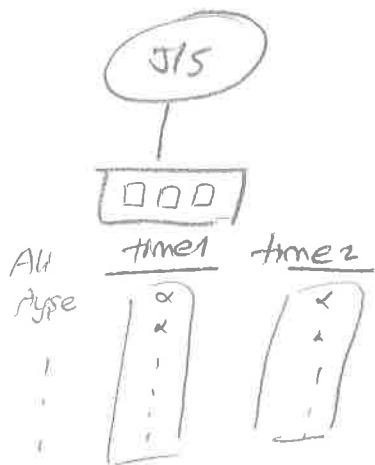
"response set" → particular ways of responding

(if there was a response set, alpha would be low. And I get a weak correlation between averages, because of response set! You change the wording. (you may get rid of negatively worded items))

Temporal Consistency

Test-retest Reliability

- you need longitudinal data



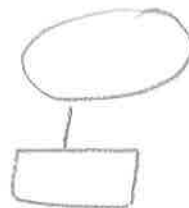
$p > .70$ → make sure you check the literature for the standard

If your scales error free
the way people respond to you at time 1
should be highly correlated the way they respond to you at time 2.

time 1 = } inter-item reliability
time 2 = }

then test-retest

Parallel-form Reliability

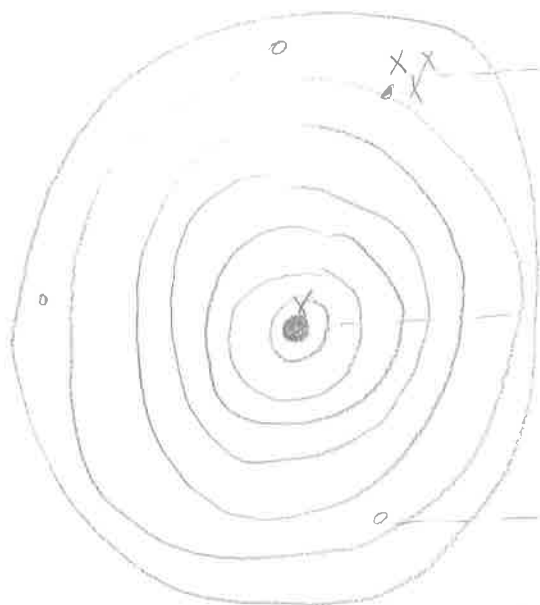


- you are using a questionnaire: scale appears on page 3 as slightly different version of the same scale (items reordered) appears on page 16.

✓ to make sure subjects are actually reading items

✓ to test reliability if they get it differently?

If your scale is reliable, scores you get from page 3 should be similar to scores from page 16.



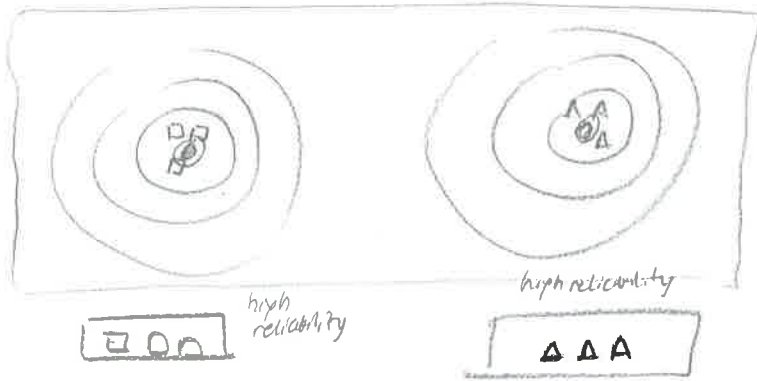
reliable/not valid

valid and reliable

not valid/not reliable

cannot be non-reliable but valid.

For smt to be valid, it should be reliable.



high reliability
□ □ □

high reliability
△ △ △

discriminant validity

strong support

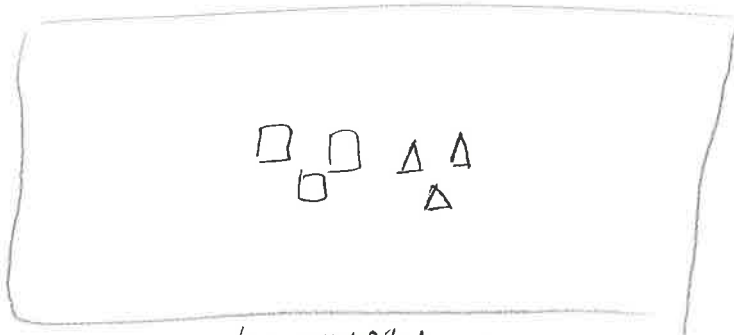


low reliability

probably has low reliability

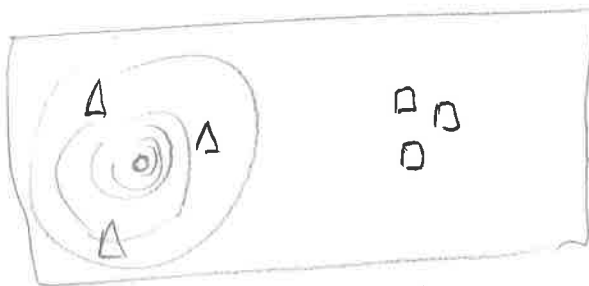
no discriminant validity

weaker support



high reliability but discriminant validity not supported.

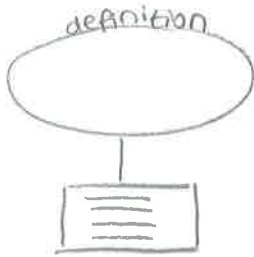
weakest of all



discriminant validity supported
not reliable

discriminant reliability supported
reliable

Measurement



- face
 - content
 - criterion
 - construct
- more subjective
- less subjective
- ↳ because we rely on statistical criteria

When validity analyses are most important?

- new construct
- new scale
- importing a scale
- new sample (different from the ones that have already been studied) (new occupational groups)

OB, Psychology, HRM

Validity analyses equal "psychometric property of the scale"

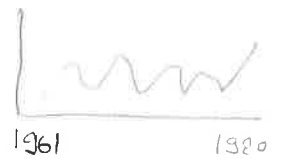
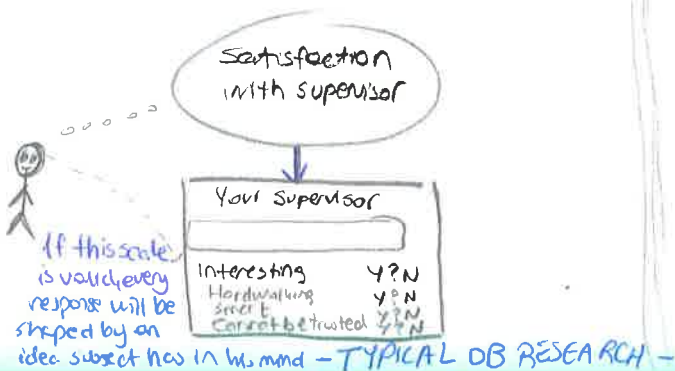
- typically multi-item scales in psychological research
- quantitative analysis mostly address multi-item scales
- questionnaires are typical data collection instruments

OT, strategy

- scales are typically single item
- secondary data → archival data

No matter what your data collection instrument is, you should be concerned with validity. ^{still}

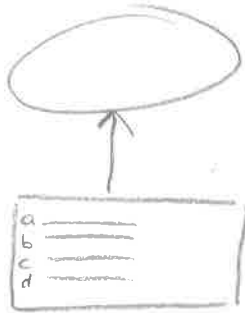
example :



- # of times "turmoil" was cited in newspapers
- # of people in prisons for political reasons ⇒ as a proxy, but does not directly capture "political turmoils"
- some historians' judgments

- In typical OB research, items are reflective which means each & every item reflect a common underlying concept.

- In non-reflective scales -



Sum of the items make up the construct.

political turmoil: the extent to which the political system is under the risk of disintegration.

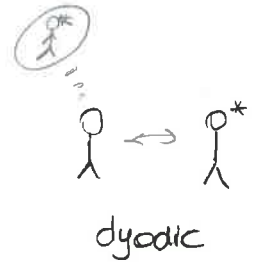
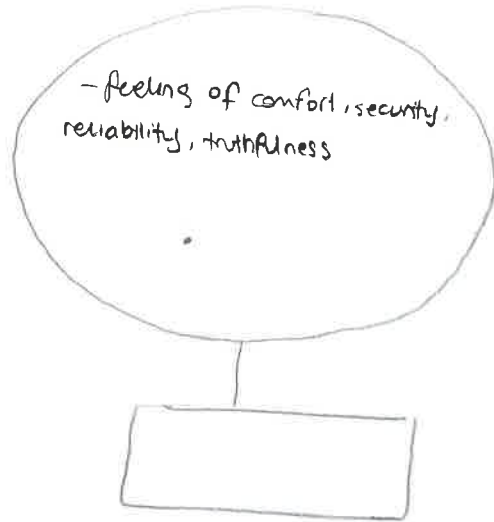
- quantitatively oriented researchers "count"

Cross-Cultural Measurement

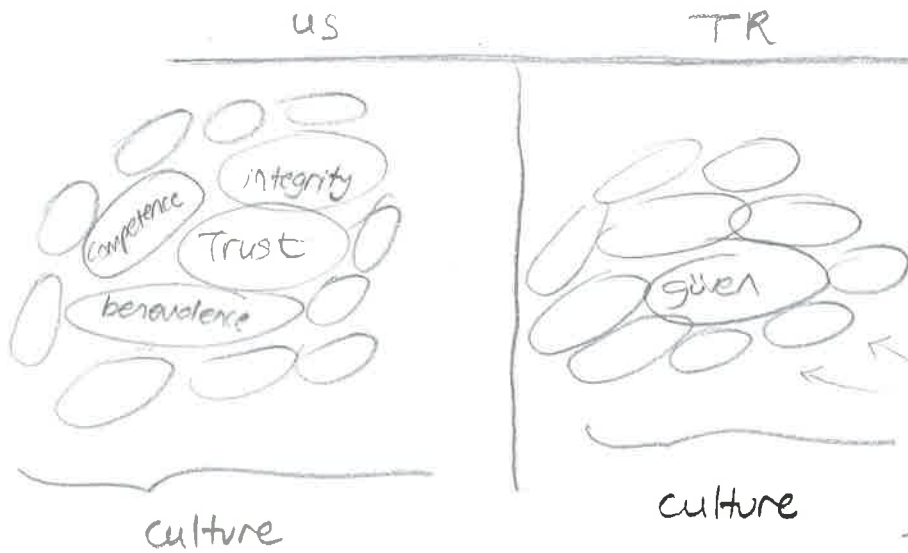
translation = false sense of security =>

Huge source of "BIAS"

To an American, trust is about "control, exchange"

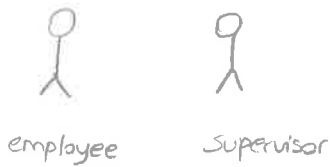


-> Doing cross-cultural research requires redefining the concept.



"Güven" needs to be defined in relation to

- You ask people what does trust mean to them.
- You can ask a panel of experts.



American
"control"
Turkey
"obligation"

"obligation" → likely to characterize our trust
invitation to ingroup
"enemy"

Guidelines for Cross-Cultural Research

(1) Is the imported concept relevant?

✓ corporate strategy ↔ state

Studying state as a force in
shaping corporate strategy relevant
in Turkey, not that relevant in US

boards of directors ↔ composition → in relation to corporate
size performance

↳ relevant in US not relevant in Turkey

(2) Does it have the same meaning across different contexts?

↳ trust → US
→ Turkey

↓
Yes, you can do comparative
research

(3) Can it be operationalized the same way across different contexts?

↓
no, no
quantitative
comparisons

developing a scale
for capturing the construct

↳ if yes; translate
↳ if not; develop a new scale

best methods translation - back translation
OR committee translation

✓ pilot testing is important \Rightarrow to a smaller sample

Based on your findings from your pilot test, you can revise your scale

DESIGN / descriptive / relational / causal research

↓

organizing several elements of research (observations, time, treatment groups)

- Hawthorne Studies / experiments (mid 1920s and early 1930s)

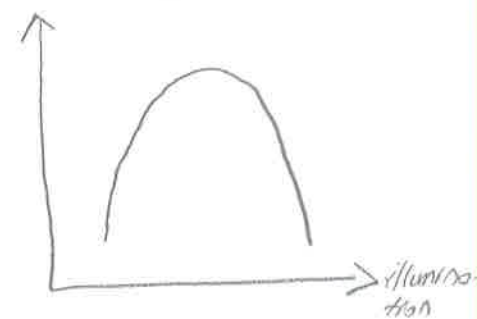
one r/q = how physical work conditions influence job performance

// (Elton Mayo interpreted the findings) \rightarrow Human Relations Approach

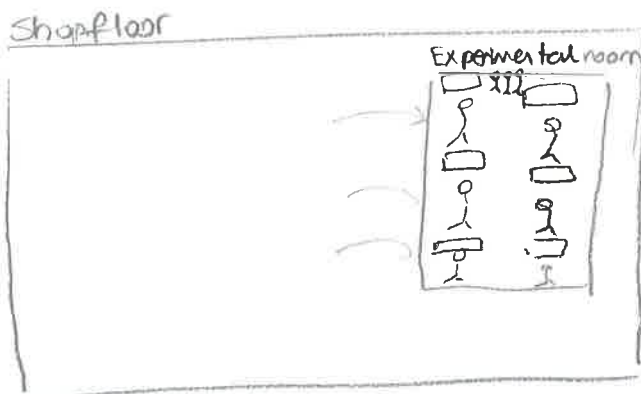
↳ prior to Mayo \rightarrow Taylor (Scientific Management)



productivity



There must be an optimal level of illumination, where productivity is highest



- female subjects

- exp. room allows researchers to manipulate illumination & measure changes in productivity

illumination	productivity
1. 100	100
2. 150	150
3. 50	200
4. 100	250

No relationship.

There is a relationship between illumination & productivity

↓
inference regarding cause-effect → we can.

↓
would that inference be valid or not???

Any conclusion based on these 2 observations would be questionable. (L) ^{may be}

initial inference: illumination and productivity are not related.

incidental

↓
Valid or not? ⇒ questionable

✓ If proper design is used, no one will question your inferences.

Using this design we cannot draw a valid conclusion. The design has to be changed.

Elements of research / design

(1) observation. $\rightarrow O_1, O_2, \dots$

- any kind of measurement
 - simple / complex observations
- ↓ ↓
- counting measuring IQ
of items

(2) Treatment $\rightarrow X$

- usually one
- probable cause
- simple / complex
- In social sciences treatments are mostly complex.
pay ↑, training ---

(3) Time \rightarrow moves from left to right

O_1 O_2

O_3 O_4

(4) Groups = treatment (experimental) group \rightarrow must
control group \rightarrow optional

(5) Assignment to groups

Assignment to treatment or control groups can be
random or non-random

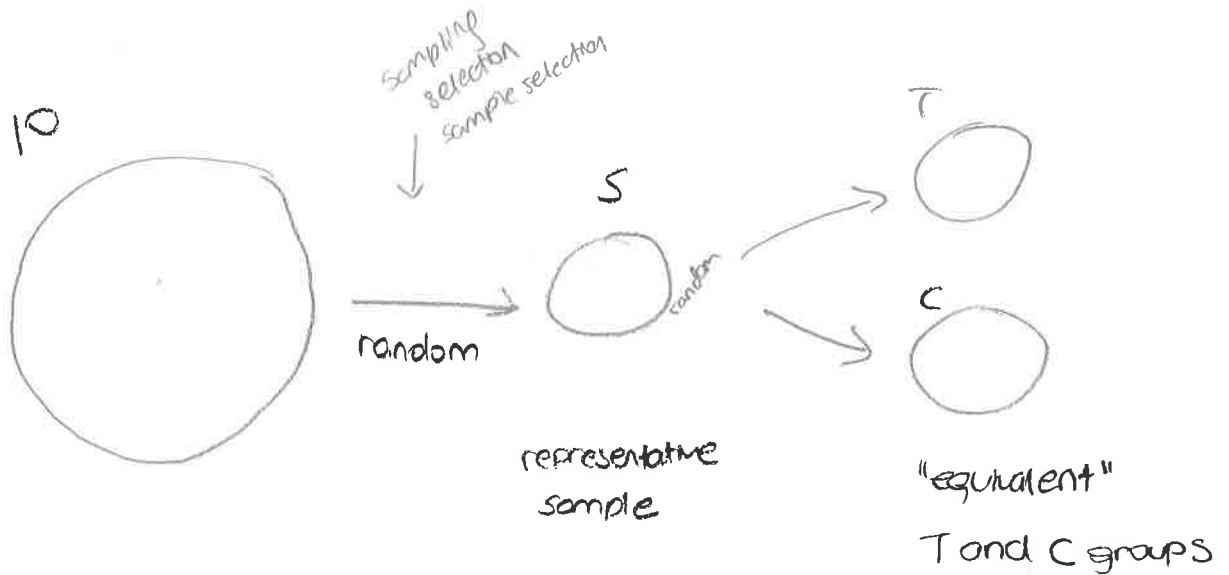
assignment to treatment & control groups

random

- randomization procedure
- equivalent T and C groups

non-random

- non-equivalent T and C groups



DESIGN

Non-experimental

- Control group is missing.

Quasi-experimental

- T&C groups
- non-random assignment

Experimental (True experiment)

- T&C groups
- Random assignment
- Treatment

higher internal validity

Internal validity = validity of claims regarding causal relationships

Non-experimental

O1 X O2

↓
A in
illumination

X O1

popular
management
books

Quasi-experimental

~~pretest posttest design~~

N O1 X O2

N O3 O4

N X O1

N O2

Experimental

R O1 X O2

R O3 O4

- different rows represent
different groups

- pretest posttest experimental
design

post test only design

R X O1

R O2

HOW DESIGN HELPS

✓ By eliminating threats to internal validity
(sources of bias)

- nonexper. design suffers from many threats to internal validity
- experimental design has highest internal validity

Threats to internal validity

(1) History



$$O2 - O1 = a$$

Inference = a was caused by X.

- internal validity of this claim
is weak for a number of reasons,
(1) history.

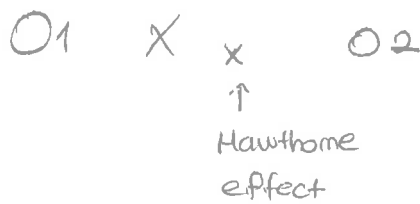
- History is anything that is happening during the experimental process not under researchers' control but may have an influence on subjects ("latent treatment")

Hawthorne experiments: performing increasing no matter how much light there was in the environment

Q = can performance increase over an extended period of time without a cause?

→ - Hawthorne effects

workers were trying to "reciprocate" by working harder



performance training performance

R O1 X O2

$O2 - O1 = X + \text{history}$

R O3 O4

$O4 - O3 = \text{history}$

$[O2 - O1] - [O4 - O3] = X + \text{history} - \text{history}$

magnitude of the effect

N O1 X O2 $(O2 - O1) = X + \text{history}$
 N O3 O4 $(O3 - O4) = \text{HISTORY}$

$X + (\text{history} - \text{history})$

These histories can cancel out because

assignment is random. Control groups and randomization eliminate bias of history effect.

Quasi-experimental designs do not eliminate history effect because of non-randomization - T & C

groups are not equivalent. What if group experiences as history cannot be equal to what C group experiences as history

Experimental Design

R O1 X O2

$$O2 - O1 = X + \text{history} + \text{maturation}$$

R O3 O4

$$O4 - O3 = \text{history} + \text{maturation}$$

$$[O2 - O1] - [O4 - O3] = X$$

(3) Testing Effect → "being observed" shaping future observations



$$O2 - O1 = X + \text{history} + \text{maturation} + \text{testing}$$



exam → testing effect

Experimental Design

R O1 X O2

$$O2 - O1 = X + \text{history} + \text{maturation} + \text{testing}$$

R O3 O4

$$O4 - O3 = \text{history} + \text{maturation} + \text{testing}$$

$$[O2 - O1] - [O4 - O3] = X$$

(4) mortality → subjects dropping out

perf tran perf
O1 X O2
n=350 n=310

~~O2 - O1~~

scores are not comparable.

This is the problem you cannot solve by design. easiest solution, no pretest, just make sure those receiving treatment present in the posttest

R O1 X O2
n=100 n=91

R O3 O4
n=100 n=91

~~O2 - O1~~

~~O4 - O3~~

X O2
n=310

Solution?



$O_2 - O_1 = \checkmark$
 meaningless!
 does not give
 valid ideas on
 your treatment

40

- The results may be different.

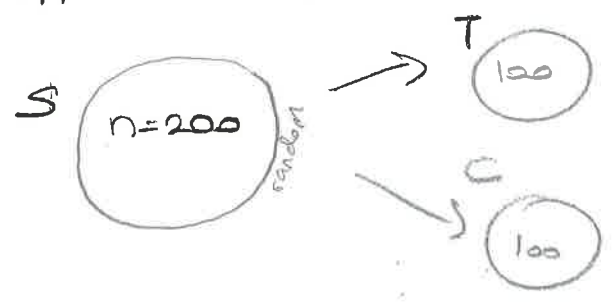
Mortality # of
 - Not a big deal if people dropping out not big.

PARTIAL SOLUTION:

✓ Are dropouts systematically different from others based on observed characteristics (age, gender, occupation, education, growth need ...)

If you can show they are not systematically different, not a big deal...

Suppose an experiment:

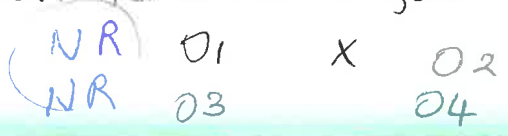


R $n=100$ X $n=31$ ₂ ~~$O_2 - O_1$~~ not feasible

R $n=100$ ₃ $n=94$ ₄ ~~$O_4 - O_3$~~ not feasible

~~$[O_2 - O_1]$~~ - ~~$[O_4 - O_3]$~~

If they are not systematically different, still your experiment turns into a quasi-experimental design.



$n=100$ each

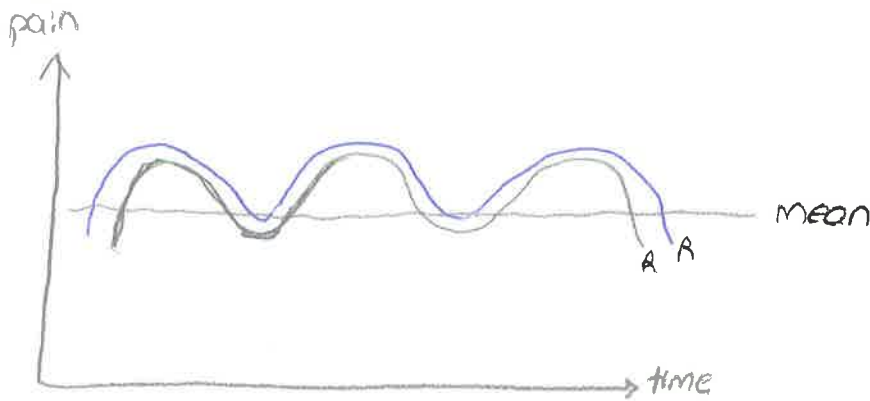
R	X	O ₁	quasi-experiment →	N/R	X	O ₁
R		O ₂		N/R		O ₂

- randomisation did not work.

✓ matching OR

✓ multivariate analysis

(5) Regression to the mean



it works!

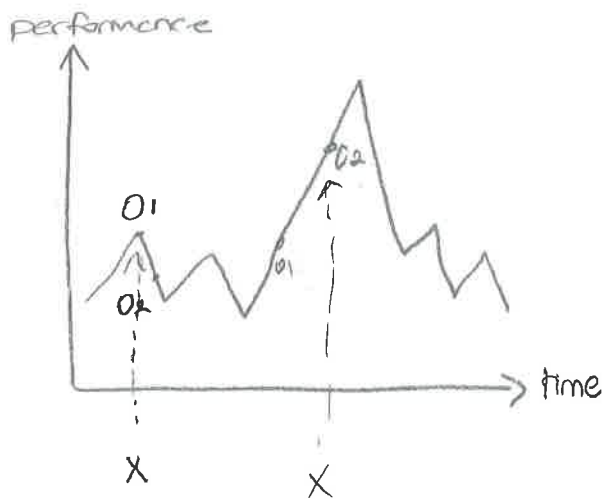
↳ is it valid?

R	O ₁	X	O ₂	} different + valid conclusion
R	O ₃		O ₄	

- The Connecticut Graduated on Speed

↳ because of regression to the mean

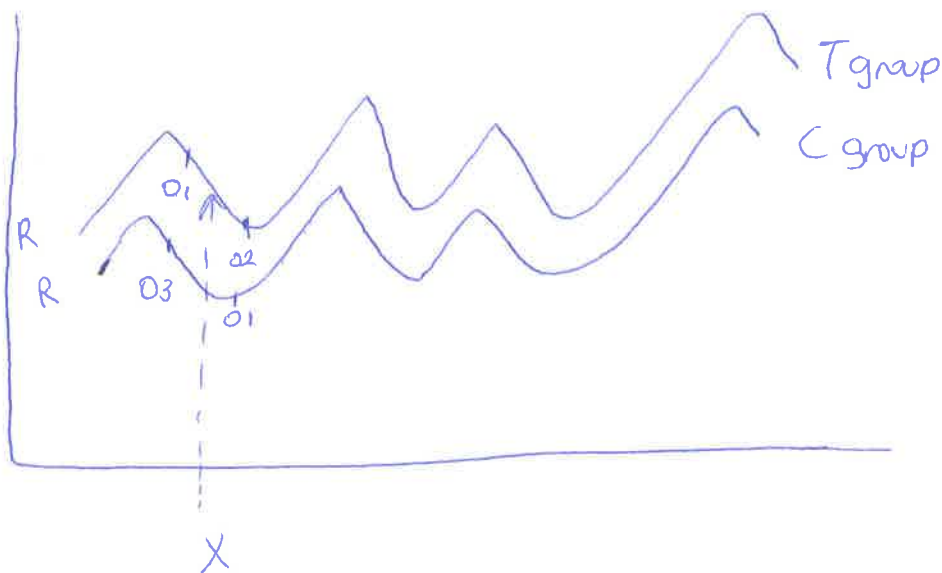
(6) Inherent Unstability



$$O_2 - O_1 < \phi$$

$$O_1 - O_2 > \phi$$

There are lots of external factors which may influence what you are observing



R O1 X O2

$$O_2 - O_1 = X + \text{history} + \text{maturation} + \text{testing} + i/u$$

R O3 O4

$$O_4 - O_3 = \text{history} + \text{maturation} + \text{testing} + i/u$$

$$[O_2 - O_1] - [O_4 - O_3] = X$$

cancel out
=

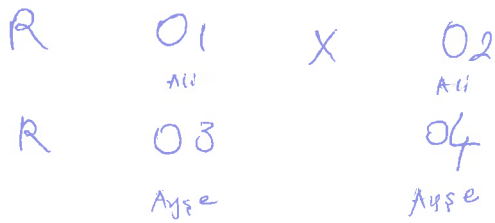
(7) Instrumentation = different scales / measurements used for different instruments
sets of observations

non-experimental:



If different supervisors
to measure performance use different
tests \Rightarrow instrumentation

experimental:

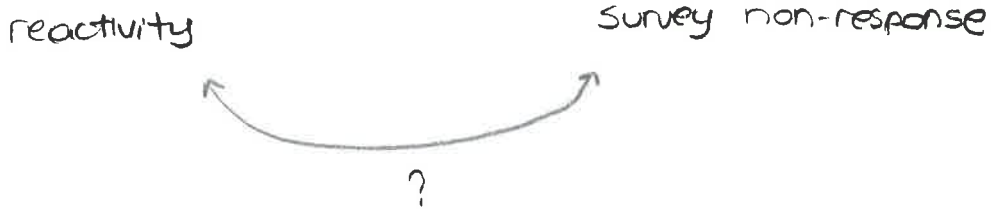


$$O2 - O1 = X + \text{history} + \dots + \text{instrumentation}$$

$$O4 - O3 = \text{history} + \dots + \text{instrumentation}$$

cancel out of
experimental
designs

SURVEY



reactivity \Rightarrow subjects react to being observed / studied

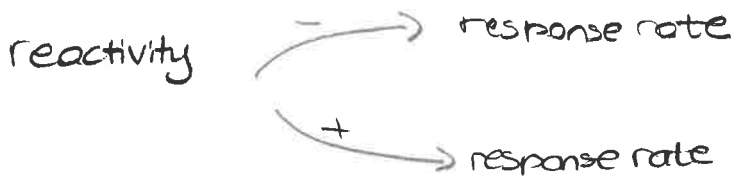
non-response \Rightarrow not responding to the survey instrument

- non response may have systematic or non-systematic reasons.

\downarrow
results in BIAS

\checkmark reactivity is a systematic reason of non-response

\checkmark Reactivity may bring about non-response.



\checkmark Reactivity may ^{also} increase the response rate.

example: service quality

\downarrow \downarrow
very very
satisfied dissatisfied



more likely to respond

downward bias

Survey → interviewing, asking, questioning, probing subjects;
(V) usually a large # of subjects

Survey → data collection instrument
(N)

✓ most likely to be a questionnaire form

✓ interview protocol → face-to-face
→ on the phone

✓ whatever the specific form it takes, it is usually "structured"
By "structured" we mean, "standardised".

- same set of items, in the same order...

Survey → questionnaire

non-response
quality + quantity

- booklet form

- cover page

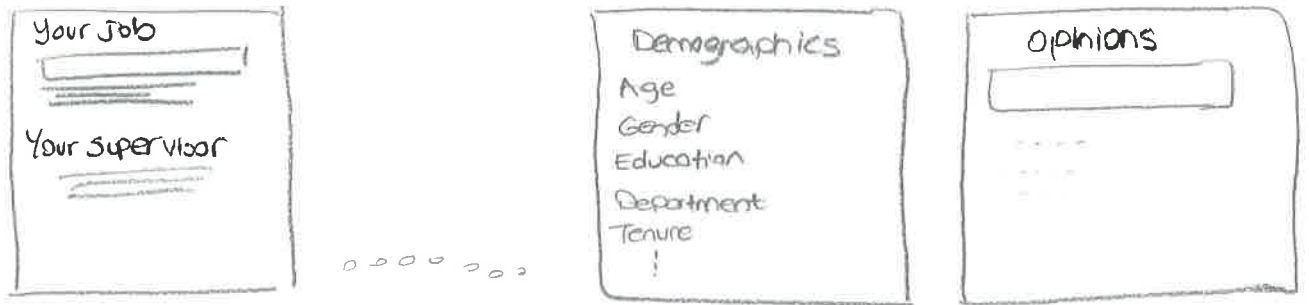
Employee
Attitudes
Survey
XYZ Co.
Ankara, 2014

Cover page

Dear Participant,
Thank you for
participating in this
research.
Purpose/
Selection /
risks/confidentiality
Thank you

Name, E-mail, phone
Introduction pg. 3

Please answer all questions.
Contact information provided below.



pg. 4

✓ scale-by-scale

✓ mixing items of multiple scales (or subscales)

↳ organizational commitment scale
(18, 24 item versions; 3 dimensions)

- 12 pt TNR

↳ larger for blue-collar workers

- double spacing

- items numbered or bulleted

Questionnaire Administration

quantity-quality of data

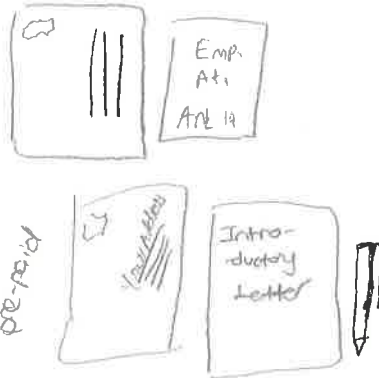
in-person by the researcher

by mail

electronic

individually ^{second best}
(not supervised)

groups of ^{first best}
subjects
(supervised)



- web page
(http://...)



- meet the subjects
- give them the questionnaire form
- briefly inform...
- ask them to respond when it is convenient for them

- go see the management, ask for cooperation
- promise a technical report
- prearrange a room, groups of participants, schedule (during work hours)

✓ usually very low response rate

✓ very expensive

1st wave → 1000 → 150

2nd wave → 850 → 60

3rd wave → 700 → 10

✓ geographic coverage

- send e-mail
- give information
- ask cooperation
- send the link to the webpage

✓ very cheap

✓ very wide coverage
(subjects with the e-mail addresses)

✓ data entry / coding

↓ NOT AN ISSUE

no cost

no errors

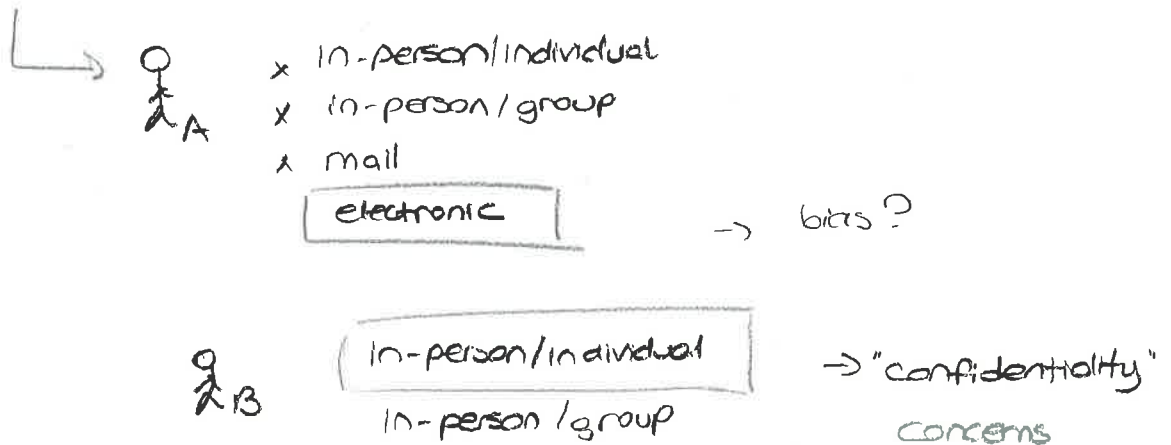
✓ very low response rate

✓ low quality of data

- If you mix methods, you are probably introducing bias.

mixing methods

introducing bias



— Non-standard methods may be introducing bias.

✓ Informative sample → serves the purpose → limited
→ great

Incentive → **participation** → "selection bias"
→ **extent to which items are responded to**

What a good questionnaire looks like:

- (1) Scales = valid and reliable [standards: literature]
- (2) Short enough (short as possible)^{os}; Inclusive (include all necessary items so that you are ^{not} compromising internal validity of your scales)
- (3) Fits subjects' intellectual capacity (by design, by content)
↓ neat, readable ↓ avoid jargon

(4) No double-barreled items.

I like my supervisor and want to see him/her promoted

↳ There has to be 2 items rather than 1.

(5) There should be NO leading items => makes a particular response more likely

As prices are increasing, wages should be increasing (Y) ? N

↳ Wages should be increasing. Y ? N

(6) No emotionally-loaded items

I ~~really~~ like my supervisor Y ? N

My supervisor is ~~very~~ smart Y ? N

(7) Precautions against social desirability

Subjects' tendency to respond in a way that they think will please the researcher

Old employees should be fined for the sake

↓ of young ones.

peer

↓

bad

Y ? (N)

- This item increases the risk with social desirability bias.

→ Senior employees should be released of their positions/duties ...

Semantic Equivalence

(8) Avoid recall-dependent items

I was happy at two years ago.
work

Y ? N

problems with recall-dependent items.

(I) memory loss

(II) what we remember about the past determined by what we experience today

---> much more fundamental issue

(9) No unclear items / expressions / words

I feel good

Y ? N

{ What do you
mean? }

(10) Make no references to "risky", "sensitive" issues

- religion

- politics

- drugs

- sex

(11) No double-negatives

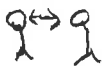
I do not think I do not know much about this topic. Y ? N

↳ I think I know much about this topic

Y ? N

INTERVIEW

- usually
face-to-face



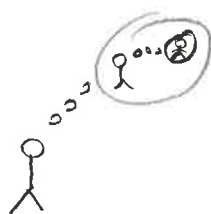
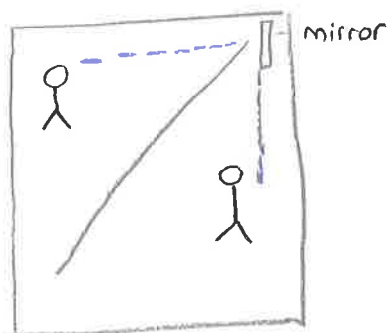
subject researcher

"reactivity" is important!

↳ especially when
carried to face-to-face

"act a script"

↓
rules, understanding how should
I behave



- When is interview most useful?

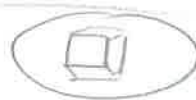
When you are interested in scripts = how your
subjects act in terms of the topic you are
interested in.



researcher



Subject



"legitimacy"

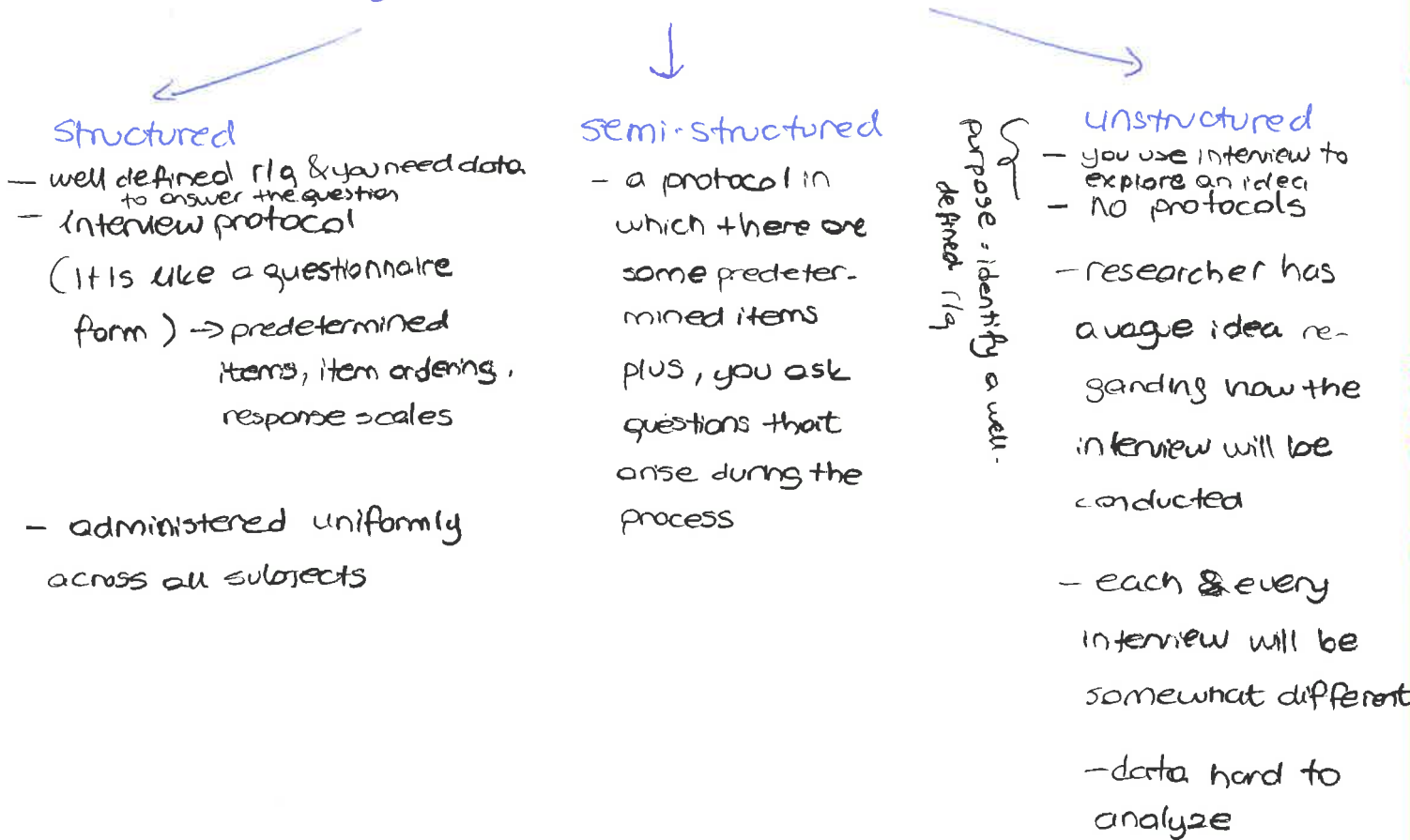
"whether a new practice
is endorsed by actors"

(approval, acceptance)

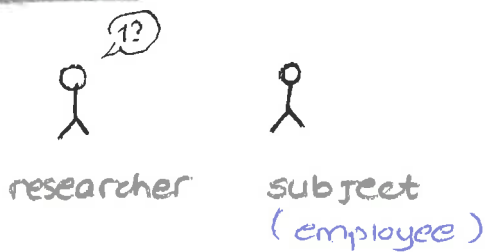
↳ then you are interested in
scripts.

- ↳ If you endorse a practice, you are likely to communicate that.
- ↳ Regarding issues about "government", "religion", "organizations", people may not be willing to communicate about them. The script they are acting in your presence may be different than what they actually think. Reactivity is an issue. It is a potential source of bias.

Types of Interviews



Unstructured:



power distance = low / high ? \longleftrightarrow supervisor - employee

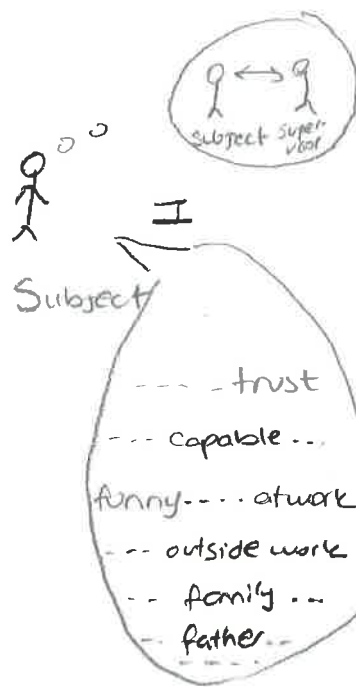
- cultural value
- the extent to which people consider power differentials normal

WHAT IS THE FIRST QUESTION ?

- What is ~~power distance~~ ?
- How is the ~~weather~~ like ?
- Is it easy to reach your supervisor ? *At the beginning, you cannot be very specific.*
- How is your relationship with your current managers ? ✓
- Can you stand "power differentials" ?
- When your ~~supervisor~~ wants to do a task how many times does he have to tell you this ?
- Do you participate in decision-making ? *too specific*
- Do you feel ~~comfortable~~ working in this office ? *not relevant*
- Does your supervisor "empower" you ?
- Could you define the ideal ~~employee~~ - supervisor relationship ?
- How would you define your work-related relationship with your supervisor ? ✓

- Tell me about your relationship to your supervisor.

- You said "father" as you describe your relationship with your supervisor. what do you mean?



power-distance
employee-supervisor
relationship

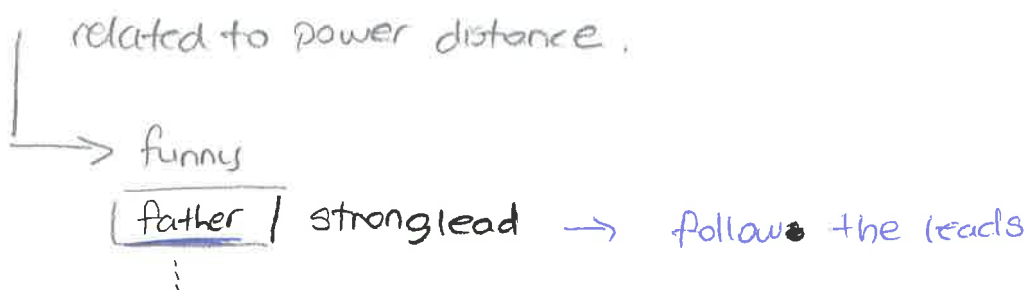
--- trust
--- capable...
funny--- at work
--- outside work
--- family ---
--- father ---

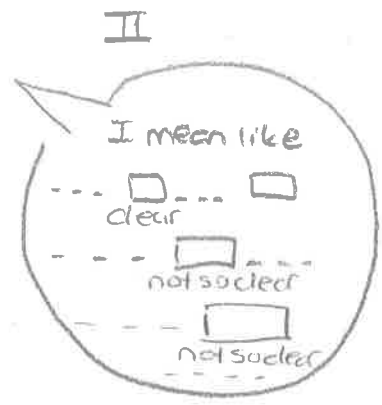
rule = Initial questions broad,
later questions more
specific / focused

✓ broad => still related to your r/q

- You cannot force subject talk about power distance. It may not exist in his relationship with supervisor. It may be irrelevant. Your first question has to be broad enough to let the subject talk about that relationship. You have to figure out if there is a relationship between power distance and employee-supervisor relationship.

- Do not intervene when the respondent answer your question. Listen and make notes. As you listen, look for "leads" regarding the power distance or something





There will be further leads.
 For the leads that are not so clear, ask for further questions & clarification.

— What if we did not see any leads?

Just push a little without saying anything about power distance.

If there is a pattern (no one mentions anything relevant about power distance), the relationship is not relevant.

Sources of bias in interviews (← + what makes a good questionnaire makes a good interview as well)

Subject

— Social desirability
 The subject believes the researcher has an idea of subject in his mind.
 The researcher should be careful about manners, question content.

But There is no solution to this.

Researcher

— Absence of rapport (researcher's inability to establish rapport)
 "Rapport" → mutual understanding and trust

Context

risky contexts
 ↳ workplaces
 "reactivity"
 — Subjects may omit smt that is significant because of not feeling comfortable

Researcher

to build rapport, give
information, ask for
consent; make a
number of promises
regarding confidentiality
(Gail journal :)

ethical misconduct = "retraction"

SCALING

⇒ influences analysis (especially
when it is related to the DV)

Your Supervisor	
	y ? N
	y ? N
	y ? N
	y ? N
	y ? N

→ response scale

(1) Nominal

→ "naming"

Gender: Male
Female

} categories, groups

(1) male (2) female

~~1 2~~

(2) Ordinal

→ "ordering"

(if it is nominal, you cannot do any mathematical
operations.)

My supervisor is a smart person

Yes > ? > N

✓ Responses can be meaningfully ordered.

(3) Interval

My supervisor is a smart person.

	Absolutely disagree	Disagree	Neither agree nor disagree	Agree	Absolutely agree
	1	2	3	4	5
OR	-2	-1	0	1	2
OR	5	6	7	8	9

✓ Assumption: that the distance between adjacent responses are the same

✓ Starting points on interval scales are arbitrary. (For example, weather)

→ $4 > 2$ yes

→ $4 - 3 = 3 - 2$ yes

→ $4 = 2 \times ("2")$ No (Because the starting point is arbitrary; no absolute zero)

↓
0: does not exist!

(4) Ratio Scale

✓ It has an absolute zero.

Age = _____ "18" → "36"
Weight = _____ twice as much
Organizational size = _____

Archival Research and Secondary Data Analysis

* Secondary data → already collected, coded, stored

collected by some other person, some time ago, for another purpose

✓ General Social Survey → the U.S.
(GSS)

✓ European Social Survey (ESS)

} secondary data sources

There are also more local ones = University X

Government Agency Y

* Archive → a collection of documents

These documents have to be inspected for relevant info. and then relevant information should be coded into data.

✓ Secondary data are convenient (no need to spend a lot of time & money to collect data)

data

You should be still concerned about methodology &

↑↑

* methodology

* sample

Does the sampling method fit your purpose?

variables: constructs
(measurement)
↓ valid + reliable

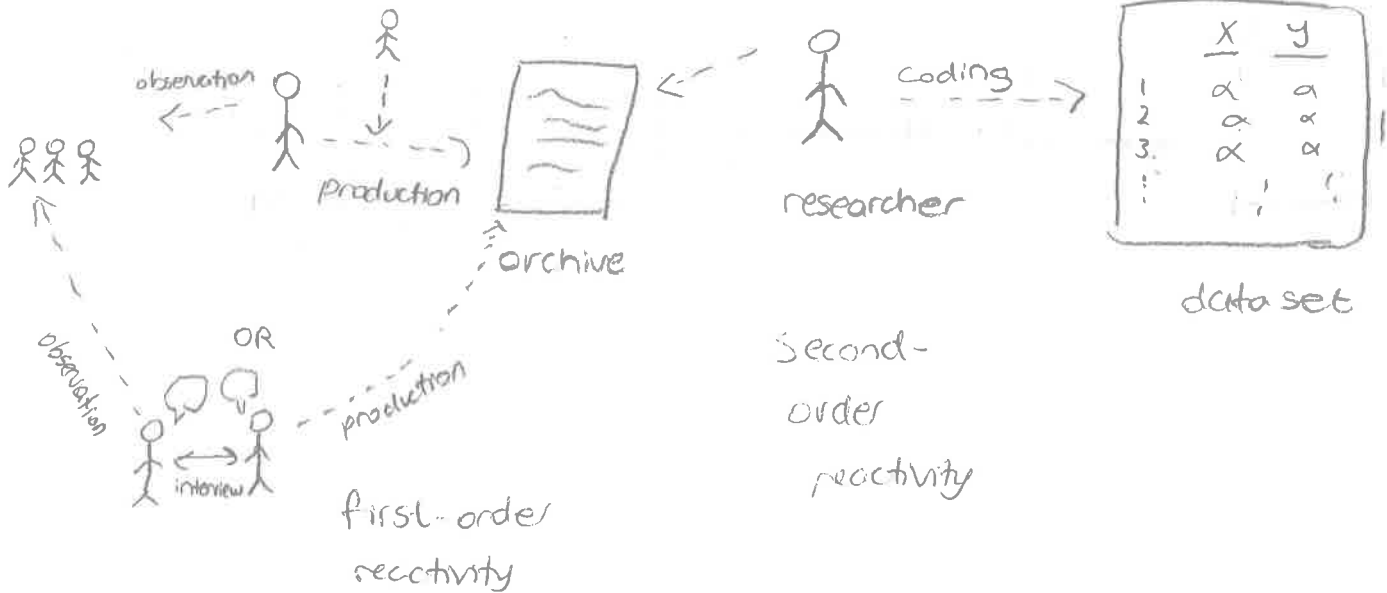
Interview:



reactivity

(Real people react to presence of other real people)

Archival research:



✓ Archives are the outcome of a social process.

✓ The social processes are sometimes rule-governed. (There is a strict methodology), for example; government keeping records of companies, labor unions...

rule-governed process

if there are standard rules administered uniformly across people and time, there is little risk of reactivity.

founding a new labor union: founders write a letter to the local government saying a labor union was established by A, B, C... on... date.

✓ Sometimes:

- reporting of social event like protests in a newspaper.

↳ ideology of the reporter, editor, boss, government

(Even if you are using archives, if you are using documents like them, you may end up with an issue of reactivity.

(news, a book about a field research

They produce sth knowing that someone will read

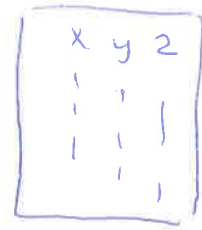
them (based on their ideology). { The author of the

reports may be responding to some future readership. } It is reactivity
or the subjects

- So archival documents are not free from reactivity

You have to be careful about the sort of the document.

Social
process



Content
analysis

data set



see what's in the
document



guided by the research question

(grounded...) → exploratory

→ needs, patterns,
issue

✓ Content analysis can be quantitative or qualitative.

It is usually a mixture of the two.

* Quantitative content analysis → count,
codes some standard
info into numbers,
calculate (based on the info.
provided in the documents)

They all take a numerical
form.

* Qualitative content analysis → looking for themes



reproducing text

* Fundamentally they are not different.

✓ Content analysis can be good or bad (sloppy)

Social processes -

- reactivity
- validity of documents → documents can be fake. You have to verify -the document & the piece of info contained in -the document.
- The documents may be preselected for presentation.
→ space ; political reasons ; administrative reasons
- document's physical quality

Easy-to-access files :

✓ annual reports of companies ; especially listed ones

↓
perhaps too much reactivity

{ an opportunity of research }

↓

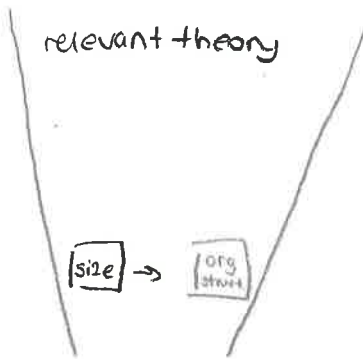
how are they presenting ---- ?
" CSR "

WRITING - SCIENTIFIC PAPERS

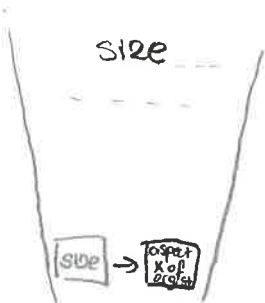
methodology

Theory-based
research questions
(strictly deductive)

↓
Is organizational
size related to structural
differentiation?
(structural contingency
theory)



OR

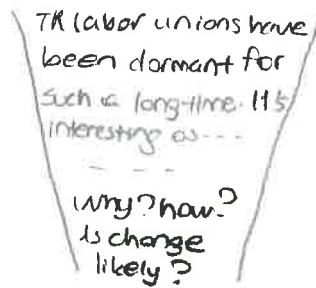


✓ general research
questions
(nomothetic)

✓ justification = contributing
to theory

Issue-based
(problem-based,
context-based)
research questions

↓
Can Turkish labor
unionism be
revitalized?
- Can still make use of
theory to justify
hypotheses.



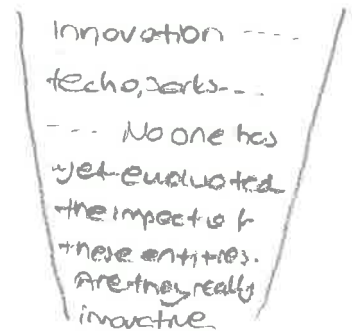
✓ Make use of diverse
literatures (literature on
labor unions, on social
movements, history of Turkey)
↳ you can identify bits of
theory

✓ justification: There is a
revitalization literature...

{ make references to
literature... } There is
smt missing in their
research. Try to address
smt missing in the literature }

Practical/applied
research question

↓
How technoparks
in TR are doing?
- can still make
use of theory to
identify relevant
DVs, IVs; to
explain why your
DVs and IVs are
related)



{ Antecedents of
performance,
performance itself }
These are the issues
that I will be
dealing in this paper.