

MEASUREMENT

A systematic way to assign numbers or names to objects and their features.

MEASUREMENT SCALES

- **nominal scale** = differences (categories)
↓
It assigns names or numbers to different categories.

only when there is a difference in the levels of a variable

Attributes are only measured

Categorical Variable

Example:

Variable: gender → 2 levels male
female = you can assign names or number to these levels using nominal scale as they refer to different categories

- **ordinal scale** = you are assigning numbers that are meaningful in terms of quality.

Example: L ranked order

You assign a specific number of stars to the quality of food in a restaurant

Attributes can be ordered

- **interval scale** = you assign numbers that are meaningful in terms of differences

Example: L equal intervals but without an absolute zero you cannot calculate any ratio. You cannot conclude that 20°C is twice as warm as 10°C.

The difference between 10°C and 20°C is the same as the difference between 20°C and 30°C.

Difference is meaningful

- **Ratio scale** = there is an absolute zero point.

Similar to interval scale, but with absolute zero.

Example:

- reaction time

Absolute zero

- weight of people

L if you weigh 100kg, you are twice as heavy as someone who weighs 50kg.

Quantitative variables

The type of information provided by different scales determines the possible statistical operations and possible conclusions.

RELIABILITY

$$\text{OBSERVED SCORE} = \text{TRUE SCORE} + \text{ERROR SCORE}$$

Consistency of measurement.
It is a prior condition for validity.

OBSERVED SCORE: measured value

TRUE SCORE: real value

ERROR SCORE: difference between the measured value and the

real value. Affected by:
→ the closer a measured score or value is to true value, the more reliable the measurement is.

- test (situation)
- the person

How can you get closer to the true score?
By reducing the Error Score.

L random error.

Types of Reliabilities

- ① **TEST RELIABILITY** = A test aims to measure a relatively stable quality (like IQ). Usually determined by observing the correlation.

TEST-RETEST RELIABILITY

consistency in results you obtain everytime the test (e measure) is used.
you compare the first and the second measurement.

threat: exercise effect (practice effect)
influence on the results of the second test.

SELF-REPORT MEASUREMENT

L self-report measures in which a person is asked to report his own behaviour or mental contents.
inaccurate and unreliable!

INTERNAL RELIABILITY

Consistency in a pattern of answers in a measure that contains several items,

like in the test.

SPLIT-HALF RELIABILITY = you divide the test items into two arbitrary groups and correlate their scores.

CRONBACH'S ALPHA = you collect data from a large sample of participants and then compute all possible correlations among the items.

$$\alpha = \frac{N\bar{C}}{\bar{N} + (N-1)\bar{C}}$$

N = number of items
 \bar{C} = average inter-item covariance among the items
 \bar{V} = average variance

② INTERRATER RELIABILITY

measures the degree of similarity between ratings of two observers of the same behavior.

consistency: often very low
different observers make different estimates wrt the same type of behavior.

Independent and relying on some kind of scoring form containing a number of items.

③ EXPERIMENTAL RELIABILITY

represents the consistency between the results obtained in an experiment → it can be estimated by exact replication.

which is a replication of a previous experiment.

DESCRIPTIVE RESEARCH

Involves what behaviors occur and in what quantity and frequency.

OBSERVATIONAL RESEARCH

The process of watching people or animals

and systematically recording how they behave

or what they are doing.

L why?

- to describe nature — Darwin's Expedition

- to generate hypothesis (not to test them)

L example: Bystrander effect

It refers to the idea that people are less likely to offer help in an emergency situation in the presence of other as compared to when they are alone.

L observation of a murder

L DIFFUSION OF RESPONSIBILITY

Idea: the presence of others leads to a different evaluation of the emergency situation

how it can be justified?

PROBLEMS:

- ① people are not always good at observation

L threat to interrater reliability

- ② Low internal validity, the extent to which is possible to rule out other possible explanations for our causal claim.

L it generates frequency claims, which have a low internal validity

L how well the conceptual variable is measured or manipulated

③ THREATS TO CONSTRUCT VALIDITY

L observer bias (observer's expectations influence the outcome of the study)

solution: limit yourself to observable behaviours

L observer effect: a change in behaviour of study participants in the direction of observers' expectations.

L reactivity: a change in behaviour when study participants know another person is watching them.

L people quite often behave differently when they know they are observed.

L question order

Solution → UNOBTRUSIVE OBSERVATIONS

You use ways to avoid the participants to become aware of the fact they are observed.

L one-way screen/mirror.

L "WAIT IT OUT" strategy

Wait until the people no longer notice your presence

L or animals

L → UNOBTRUSIVE MEASURES

You are not observing specific behaviours but you are measuring the results of those behaviours.

SURVEY RESEARCH

SURVEY = A method of posing questions to people on the telephone, (post) in personal interviews, on written questionnaires, or via the internet.

QUESTION FORMATS → Open-ended questions

① forced-choice format:

people must choose an answer between 2 or more options.

② likert scale

rate a target object → score made of adjectives

③ semantic differential format:

leading question wrt to a specific concept → question should be neutrally phrased

L Example car accident

L double-barreled question (two questions in one)

L negatively worded questions: double negative makes questions more difficult for responding

L question order

To control this effect is important to include different versions of a survey with different orders.

L the use of response sets

Solution: inclusion of reverse worded items

Leave out the middle option (to avoid faking)

L how it can be justified?

L trying to look good or bad (faking)

L to respond in a socially desirable way

L people are not always capable to report accurately about their feelings, thoughts and behaviours.

L especially when they involve memories and past events.

DESCRIPTIVE RESEARCH (BOTH OBSERVATIONAL AND SURVEY) LEADS TO FREQUENCY CLAIMS

L often made wrt all entire population, but it still refers to a certain source.

POPULATION → SAMPLE

L generalization from the sample to the population of interest is possible if the sample is representative or unbiased.

All people in the population have

equal chance to be included in the sample

THREATS TO EXTERNAL VALIDITY

- Biased or unrepresentative sample

because of convenience: sample with only those who are easy to contact.

L self-selection: sample with only volunteers

L How do we obtain a representative sample?

① random sample

② different rewards:

L random sample and then people who refuse to take part in the study

L may be offered extra money (proffessor: no treatment going towards participant)

③ replication of the study with a different set of participants

THREATS TO STATISTICAL VALIDITY

- too small sample (relatively large margin of error)

L wrong conclusions

SURVEY RESEARCH

SURVEY = A method of posing questions to people on the telephone, (post) in personal interviews, on written questionnaires, or via the internet.

QUESTION FORMATS

① open-ended questions

② forced-choice format:

people must choose an answer between 2 or more options.

③ likert scale

rate a target object → score made of adjectives

④ semantic differential format:

leading question wrt to a specific concept → question should be neutrally phrased

L Example car accident

L double-barreled question (two questions in one)

L negatively worded questions: double negative makes questions more difficult for responding

L question order

To control this effect is important to include different versions of a survey with different orders.

L the use of response sets

Solution: inclusion of reverse worded items

Leave out the middle option (to avoid faking)

L how it can be justified?

L trying to look good or bad (faking)

L to respond in a socially desirable way

L people are not always capable to report accurately about their feelings, thoughts and behaviours.

L especially when they involve memories and past events.

DESCRIPTIVE RESEARCH (BOTH OBSERVATIONAL AND SURVEY) LEADS TO FREQUENCY CLAIMS

POPULATION → SAMPLE

L generalization from the sample to the population of interest is possible if the sample is representative or unbiased.

All people in the population have

equal chance to be included in the sample

THREATS TO EXTERNAL VALIDITY

- Biased or unrepresentative sample

because of convenience: sample with only those who are easy to contact.

L self-selection: sample with only volunteers

L How do we obtain a representative sample?

① random sample

② different rewards:

L random sample and then people who refuse to take part in the study

L may be offered extra money (professor: no treatment going towards participant)