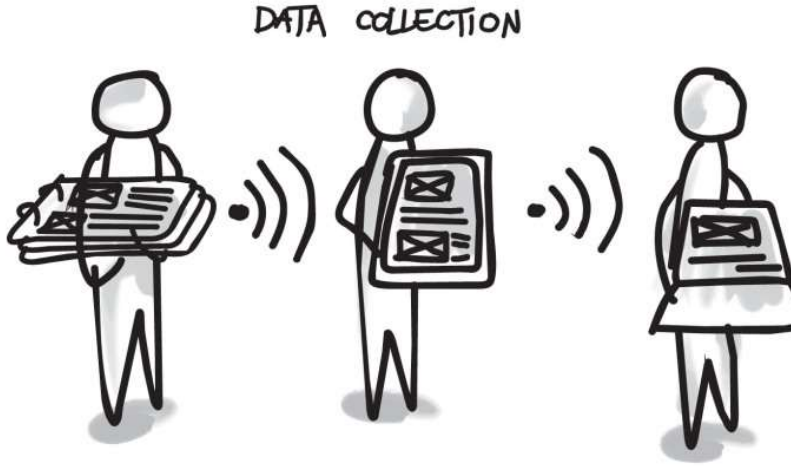
Universidad de Cádiz
Facultad de Filosofía y Letras
Applied Linguistics



Research methods in Applied Linguistics: Quantitative data collection

Bárbara Eizaga Rebollar

March 2017

2.2. Data collection

2.2.3. Variables

Variable ⟶ anything that has a quantity or quality that varies, e.g. *age can take different values*

for different people

for the same person at different times

City can take text values

New York Sidney London

2.2. Data collection

2.2.3. Variables (quan research)

Attribute → a specific value on a variable. The variable *sex/gender* has 2 attributes: *male* & *female*. Or, the variable *agreement* has 5 attributes:

1=strongly disagree 2=disagree 3=neutral 4=agree 5=strongly agree

Variables can be

- dependent — the variable the researcher is interested in, influenced by the independent one.
- independent — variable manipulated by the researcher to see if it makes the dependent one change.

2.2. Data collection

Example:

Objective	To find out how language use varies among social classes.
Subject	Presence/absence of [r] in postvocalic position
Dependent	Casual pronunciation of *Fourth floor*, emphatic pronunciation of *Fourth floor*.
Independent	The store(*Saks 5th Avenue* [highest ranking], *Macy's* [mid-ranking], *S. Klein* [lowest ranking], floor within the store, sex, age, occupation, accent.
Result	62% of Saks Fifth Avenue employees, 51% of Macy's employees, and only 20% of S. Klein's employees used all or some [r].
Significance of the study	The study revealed that social class of speakers influences the way they use the language.

Labov 2006

Confounding variable → extra variable the researcher failed to control or eliminate; it affects both variables (e.g. rhotacistic speakers).

2.2. Data collection

2.2.1. Quantitative data collection

Besides language tests,

questionnaires are the most common instruments in AL.

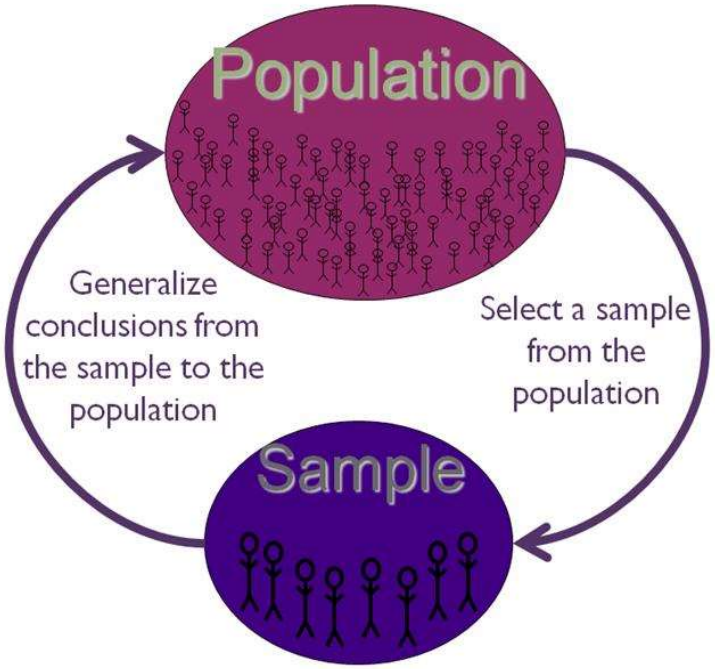
3 topics to develop an understanding of quan methods:

1. Ways of sampling participants.
2. Experimental & quasi-experimental studies
3. Collecting data via the Internet.

2.2.1. Quantitative data collection

1. Sampling in quan research

- *Population* ⟹ group of people the survey's findings are to be applied to, e.g. EFL learners in Chinese secondary schools.



-*Sample* ⟹ group of participants the researcher examines
e.g. 3 Chinese secondary classrooms

⇓

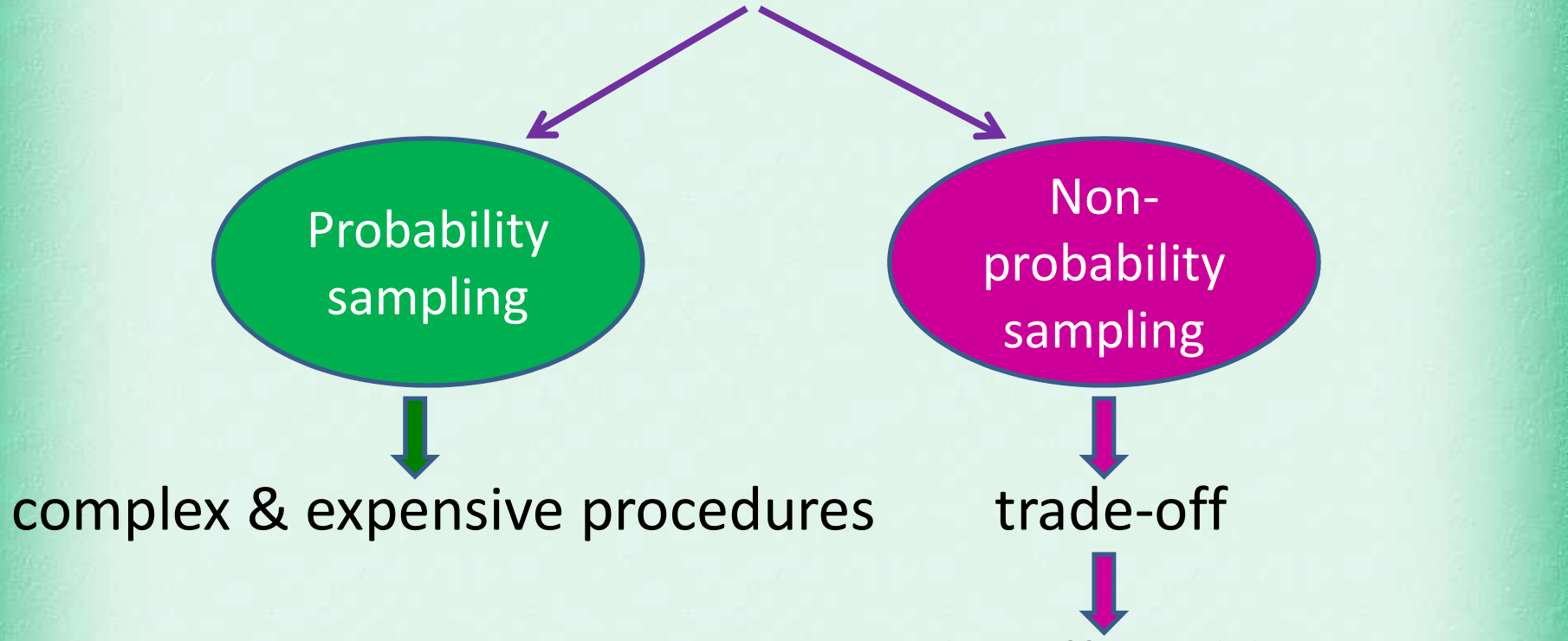
representative of the whole population

affects the strength of the conclusions from the result

- How do we select the sample? By *sampling procedures*

2.2.1. Quantitative data collection

1.1. Sampling procedures



2.2.1. Quantitative data collection

Probability sampling: number of scientific procedures

- *Random sampling*:

Selection

⇩

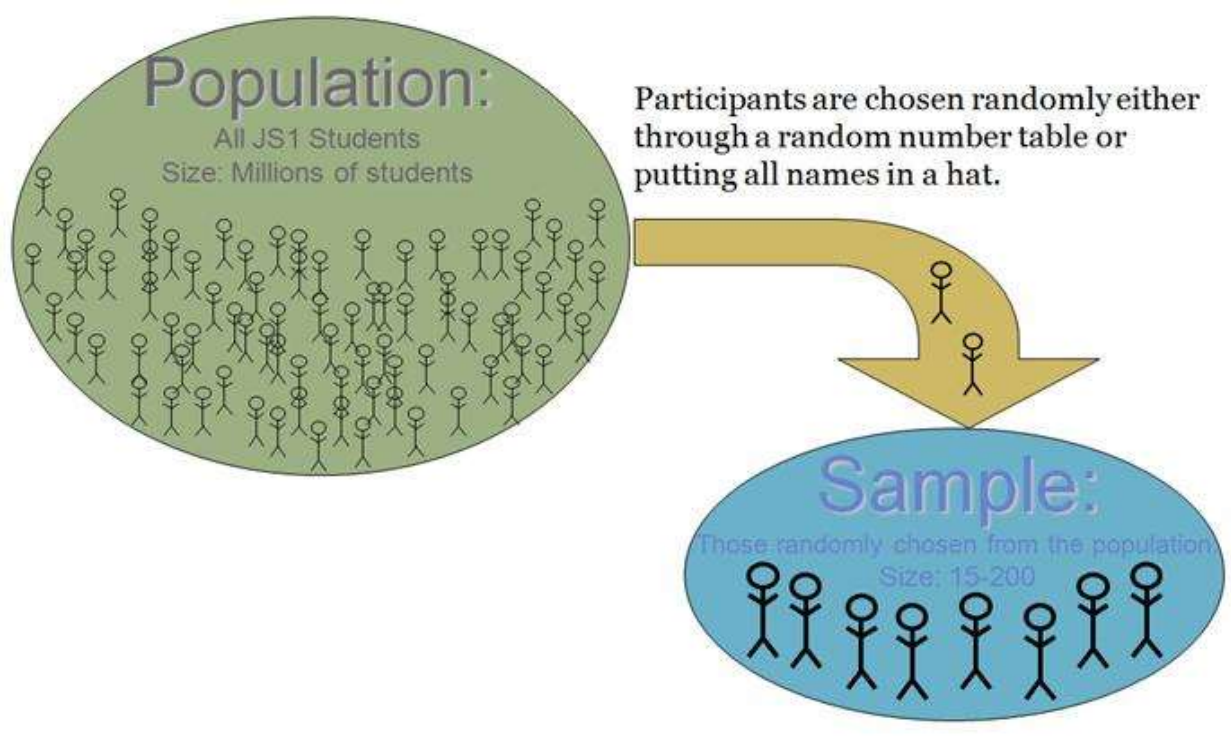
based on probability & chance

⇩

minimize the effects of subjective factors.

⇩

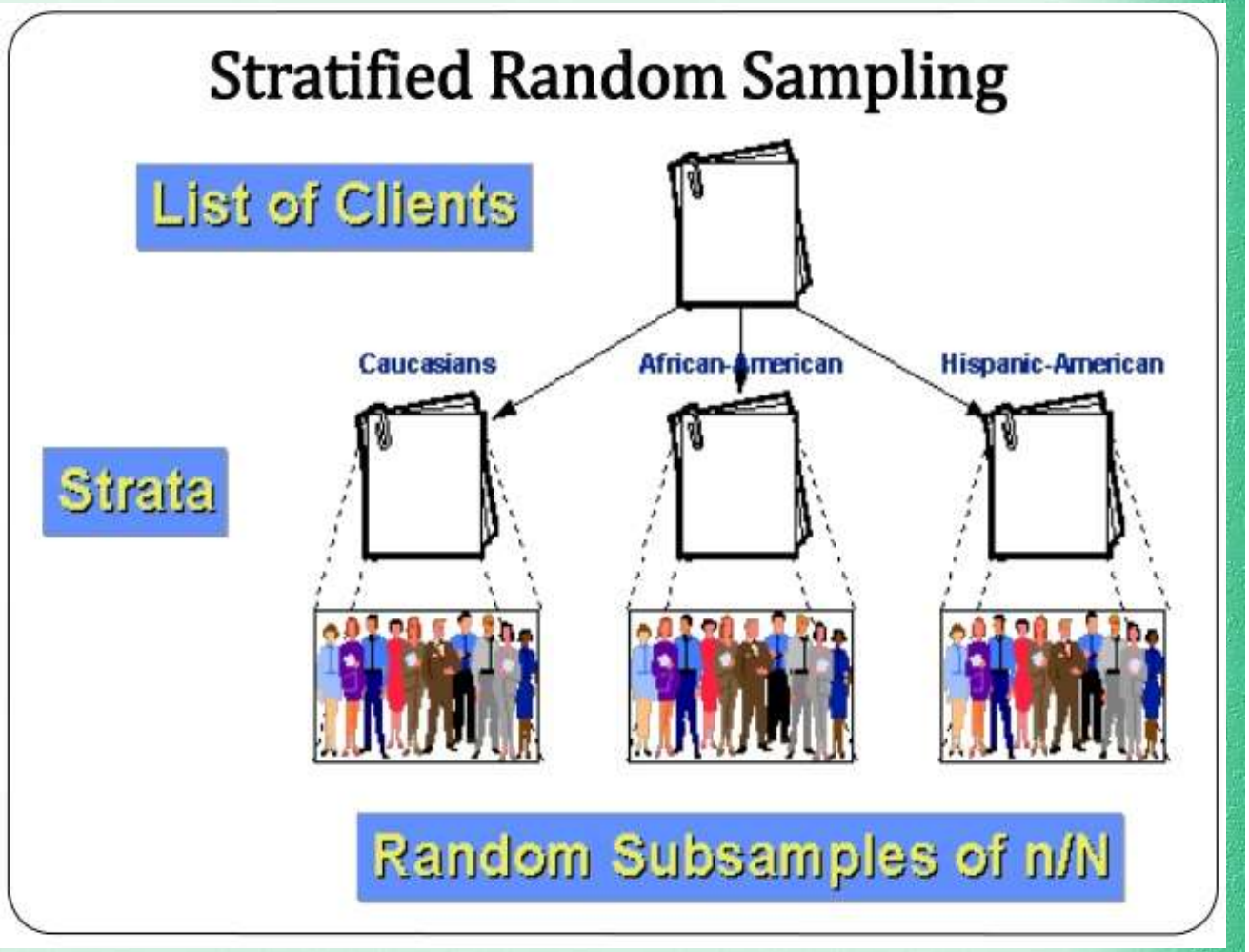
A large sample with **subjects with similar characteristics** to those of the population as a whole.



2.2.1. Quantitative data collection

- *Stratified random sampling*:

Combination of **randomization** & **categorization**

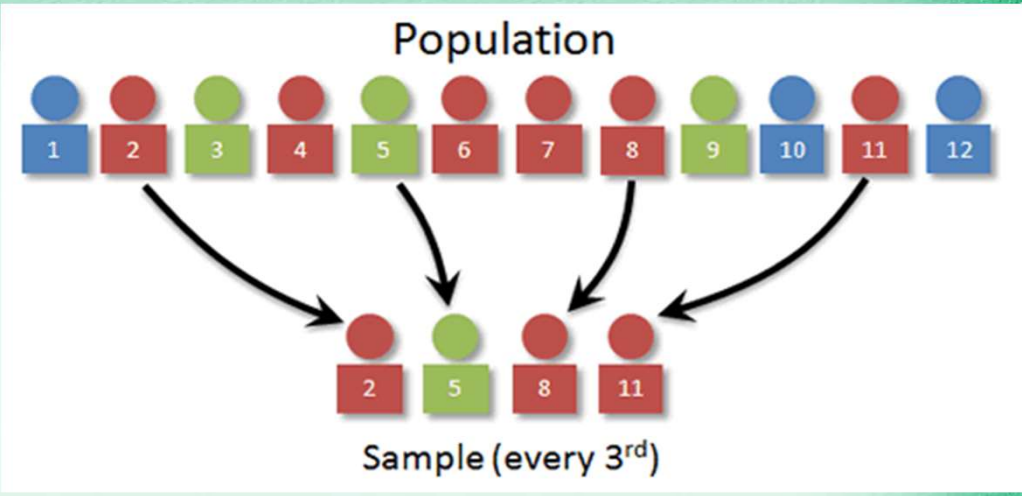


Population usu. stratified on more than 1 variable.

Random samples selected from all groups defined by the **intersections of the various strata**.

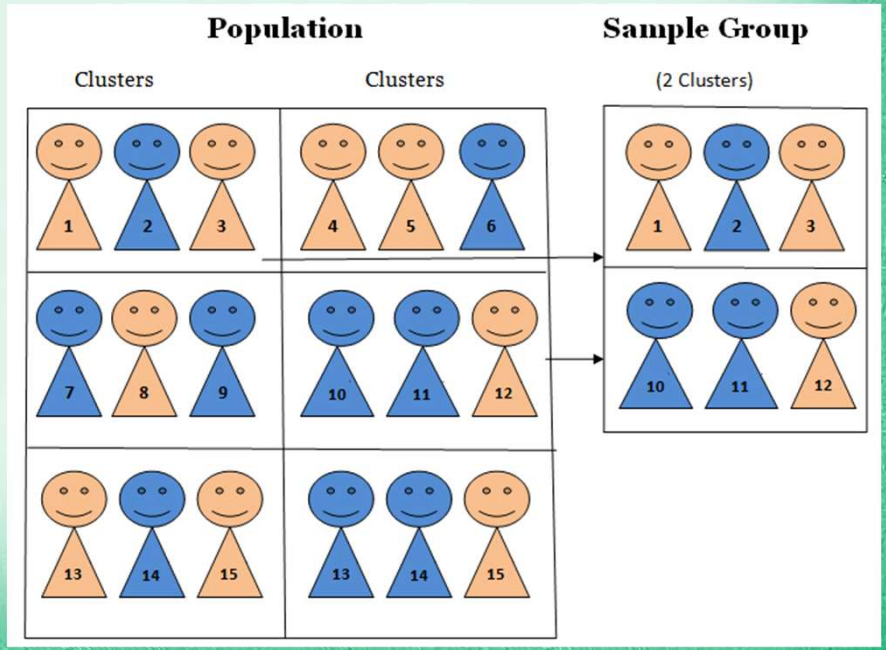
2.2.1. Quantitative data collection

- *Systematic sampling* involves selecting every nth member of the target group.



Used in anonymous surveys to make a random selection.

- *Cluster sampling* used when the target population is dispersed to randomly select larger groupings of the population (e.g. schools to examine all students in those groupings).



2.2.1. Quantitative data collection

Non-probability sampling applied by most actual research in AL.

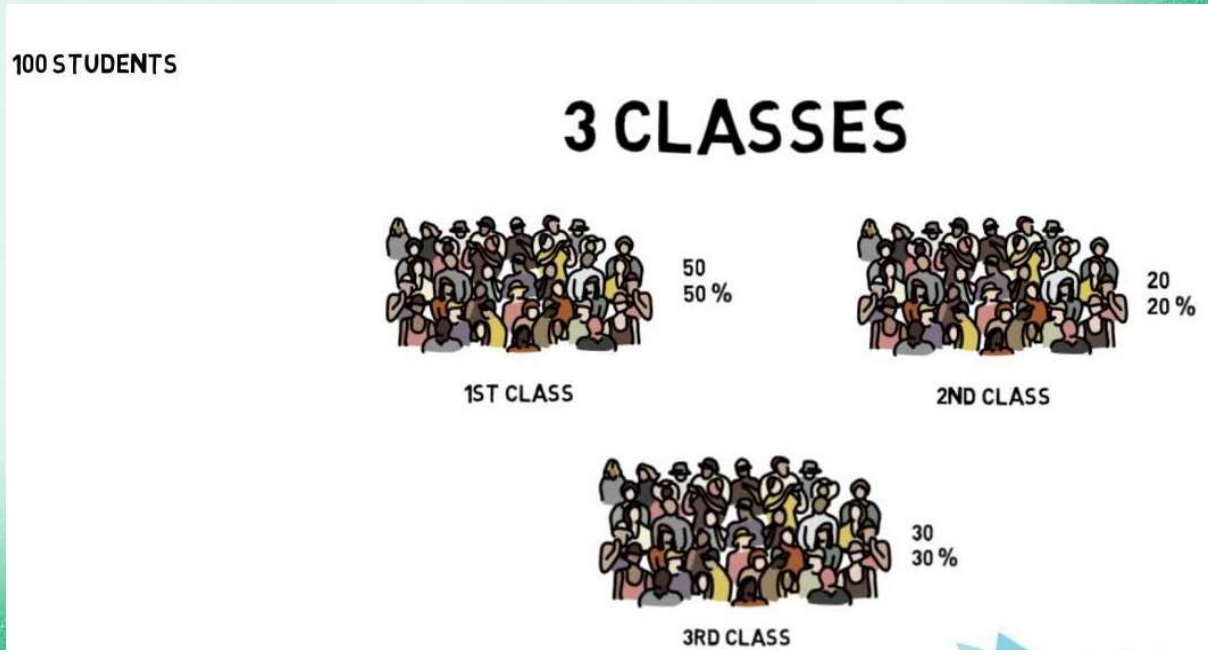
3 main strategies:

-*Quota sampling* :

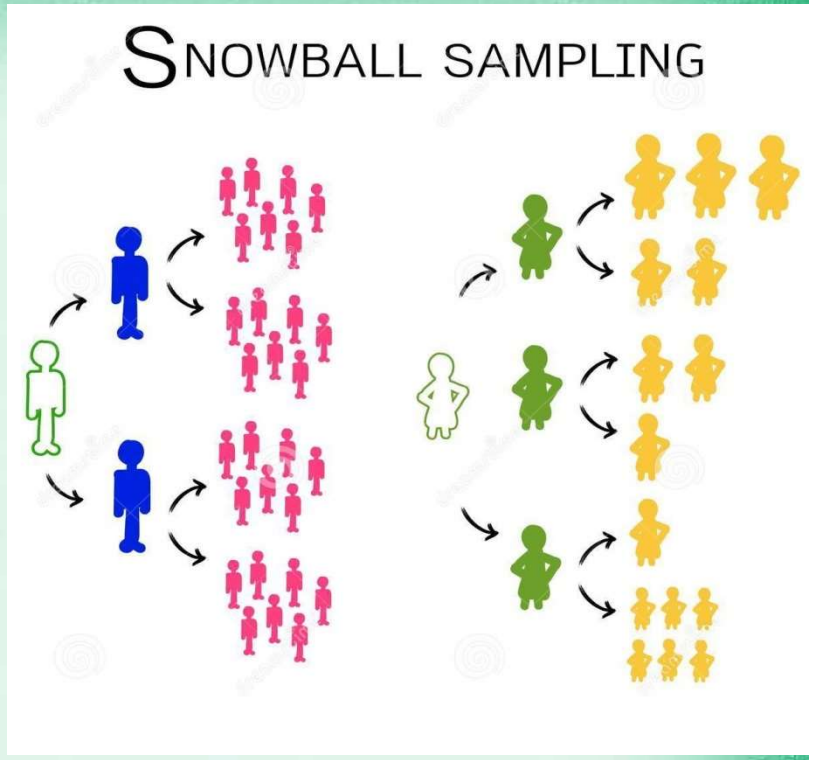
We start with a sampling frame & determine the size of the subgroups by the parameters in the frame.

⇩

No random element



2.2.1. Quantitative data collection



-*Snowball sampling* involves a *chain reaction*

Researcher identifies a few people meeting the criteria & then asks participants to identify further members.

Useful for unidentifiable memberships (e.g. gang members)

-*Convenience sampling*: members selected for the purpose of the study if they meet practical criteria (geographical proximity, availability, accessibility, etc.)

Most common sample in L2.

2.2.1. Quantitative data collection

1.2. How large should the sample be?

In the literature ➡ 1-10% of the population, a min. of 100

But the more scientific the sampling procedures, the smaller the sample size can be

↳ Opinion polls as small as 0.1%

Sample sizes for quan methods:

- ✓Correlational research ➡ at least 30 participants
- ✓Comparative procedures ➡ 15 participants per group
- ✓Multivariate research ➡ at least 100 participants

2.2.1. Quantitative data collection

Other things to consider:

- *Statistical consideration*: for a normal distribution, 30 or more people.
- *Sample composition*: If different subgroups (girls-boys), minimum size applies to the *smallest subgroup* in the sample.
- *Safety margin* for unplanned circumstances: participants' dropout, disqualified questionnaires or unexpected subgroups that need to be treated separately.
- *Reverse approach*: 1) we estimate the power of the expected results; 2) we determine the sample size.

2.2.1. Quantitative data collection

1.3. The problem of participant self-selection

Sample composition may be altered by participants:

- Researchers invite volunteers to take part in a study (often offering money to compensate for their time).
- The design allows for a high degree of dropout, so participants self-select themselves *out* of the sample.
- Participants are free to choose whether to take part in a study or not (online questionnaires).

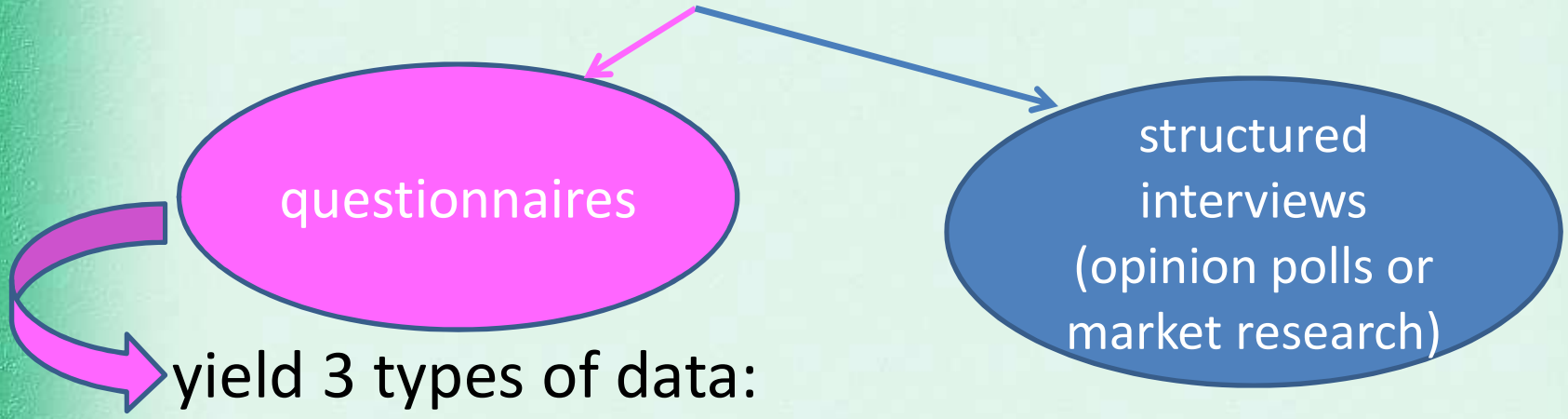
Self-selection is inevitable to some extent, but the sample may lose its representativeness preventing any meaningful generalizability.

2.2.1. Quantitative data collection

2. Questionnaire surveys

Surveys try to describe the characteristics of a population by examining a sample.

Survey data can be collected by



yield 3 types of data:

- ✓ *Factual questions*: facts about the participants (age, race)
- ✓ *Behavioural questions*: actions, life-style, habits & history.
- ✓ *Attitudinal questions*: attitudes, opinions, beliefs & values.

2.2.1. Quantitative data collection

2.1. Multi-item scales

How to word the items is very important, esp. when assessing non-factual matters (attitudes, beliefs, etc.).

Minor differences in how a question is formulated can produce different levels of agreement or disagreement:

e.g. *Do you think the US should [forbid/not allow] public speeches against democracy?*

How can we deal with this? ⟹ Multi-item scales:

1. Differently worded items focusing on the same target
2. Item scores are summed resulting in a total scale score
3. No individual item carries an excessive load

2.2.1. Quantitative data collection

2.1. Multi-item scales

Satiation control scale

- Once the novelty of learning vocabulary is gone, I easily become impatient.
- During the process of learning vocabulary, I feel satisfied with the ways I eliminate boredom.
- During the process of learning vocabulary, I am confident that I can overcome any sense of boredom.
- When feeling bored with learning vocabulary, I know how to regulate my mood in order to invigorate the learning process.

Tseng et al. (2006), *Self-regulatory Capacity in Vocabulary Learning questionnaire*

2.2.1. Quantitative data collection

2.2. Writing questionnaire items

Items usu. ask about very specific pieces of information (language preference) or give several options to choose from (closed-ended items)

Questionnaire data suited for quan, statistical analysis.

Questionnaires with open-ended items suited for qual and exploratory data, but unlikely to give rich descriptions of events or perspectives.

Closed-ended items are considered more professional, though most of them include some open-ended items.

2.2.1. Quantitative data collection

-Likert scales

- Most famous type of closed-ended items.
- A statement + respondents' degree of dis/agreement:

Spaniards are genuinely nice people.

Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree

Each response option is assigned a number for scoring purposes (e.g. 'strongly agree' = 5 … 'strongly disagree' = 1).

The scores for the items addressing the same target are summed up or averaged.

2.2.1. Quantitative data collection

Teachers' Perceptions of Second Language Learning and Teaching (Kvist 2014)

Background Questions

1. What is your gender? F M
2. Are you a native speaker of English? Y N
3. If so, of what English variety?
4. What is your country of origin?
5. In what country are you currently teaching?
6. In what other countries have you taught?
7. What age groups and levels are you teaching during the 2013-2014 academic year?
8. How many years of active teaching experience do you have?
9. Please list any academic degrees (with specializations) you have (if any):

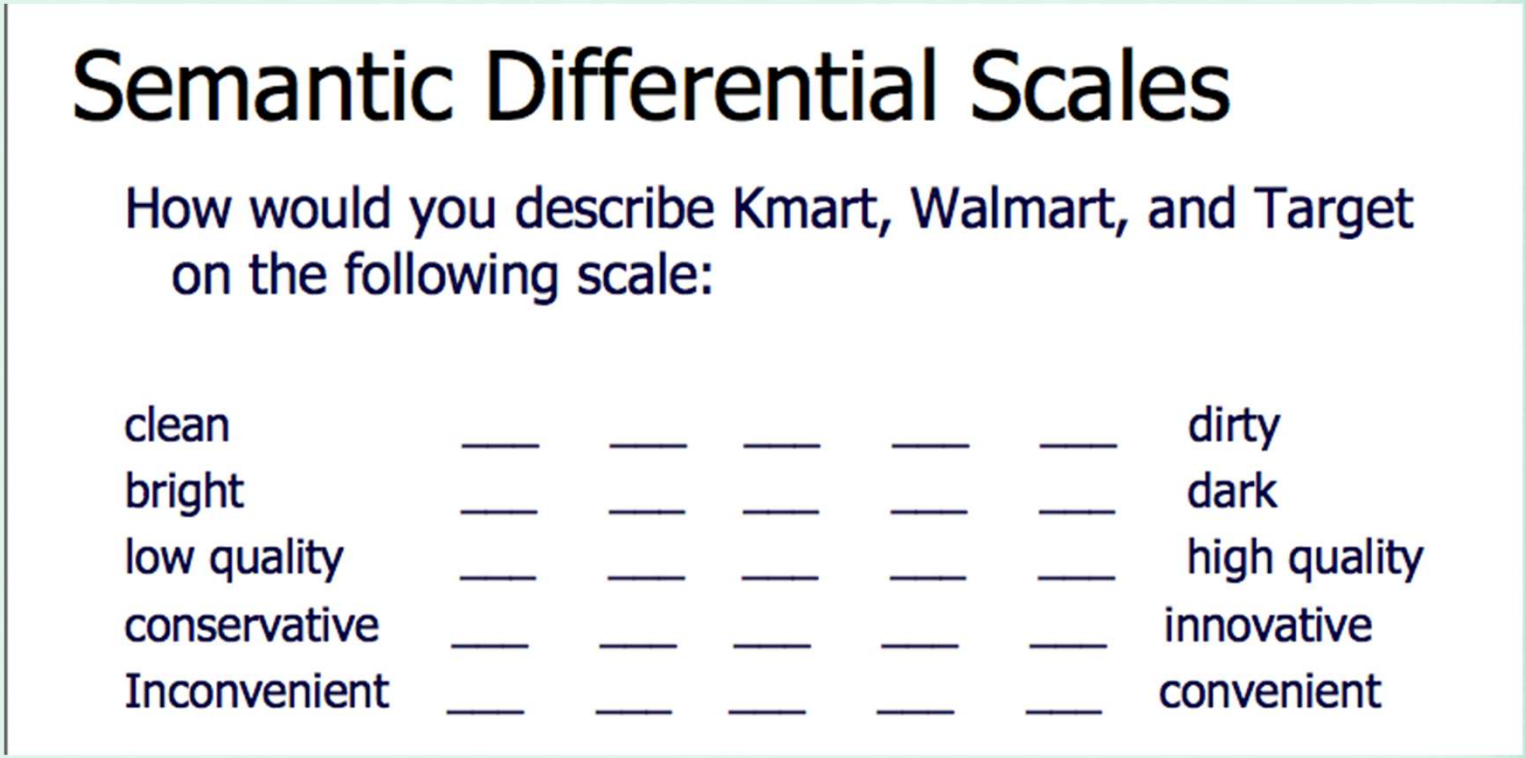
Section 2

Questionnaire Items		Strongly agree	Agree	Disagree	Strongly disagree
1	Adults learn a second language differently than children do.				
2	Making errors and discussion them can facilitate the process of language learning.				
3	Learners' beliefs about second language acquisition can either hinder or help them in their learning.				
4	When students do not learn what they are taught it is mainly due to them not being ready to learn a specific language feature.				
5	L2 students' errors should always be corrected immediately.				
6	Immersion programs are better than other forms of second language education.				
7	A focus on grammar in the second language classroom is important in helping students achieve accuracy.				
8	Learning a first language differs from learning a second language.				
9	Most students want teachers to correct their mistakes as soon as possible.				
10	Finding out what learners believe about second language acquisition can help teachers in their teaching.				
11	Students tend to copy each other's L2-errors when engaging in pair-work.				
12	Grammar should be taught in isolation in the second language classroom.				
13	3-year-olds have metalinguistic knowledge.				
14	Most learners of a specific L2 go through the same developmental stages.				
15	The younger the student is, the better it is with respect to learning a second language.				
16	Most L2 errors are caused by differences between the L1 and the L2.				

2.2.1. Quantitative data collection

-Semantic differential scales

- Elicit graduated responses by marking a continuum with a tick between two bipolar adjectives at the extremes.



- Multi-scores are computed by summing individual scores.
- Positive & negative poles alternate on the right and left.

2.2.1. Quantitative data collection

-Numerical rating scales

- A series of numerical ratings on a continuum. The rating continuum can refer to a wide range of adjectives.

- Very much used in everyday life, e.g. *assess the film you have watched on a scale from 1 to 5.*

2.2.1. Quantitative data collection

-Other closed-ended item types

- *True-false items*: In some cases, a polarized yes-no decision can be reliable, e.g. tests with children. Problem: oversimplification.

Which of the following terms indicates observable student performance? Check all that apply.

__Yes __No 1. Explains

__Yes __No 2. Identifies

__Yes __No 3. Learns

__Yes __No 4. Predicts

__Yes __No 5. Realizes

- *Multiple-choice items*: Popular in standardized L2 proficiency testing.
- *Rank order items*: order items by assigning a number to them according to participants' preferences.

2.2.1. Quantitative data collection

-Open-ended questions

Questions followed by a blank space for respondents to fill in.

They allow greater freedom of expression & richness than quan data.

They work very well if they contain some guidance:

- *Specific open questions*: facts about the participant, past activities or preferences.
- *Clarification questions*: Attached in multiple-choice item, e.g. *Please specify…*
- *Sentence completion*: 'The thing I like most about the course is…' ⇒ more meaningful answer than a question.
- *Short-answer questions*.

2.2.1. Quantitative data collection

-Item wording:

- ✓*Aim for short & simple items*, rarely exceeding 20 words & preferably written in simple sentences with 1 thought.
- ✓*Use simple & natural language*: clear & direct with no acronyms, abbreviations, colloquialisms, proverbs, jargon, or technical terms.
- ✓*Avoid ambiguous or loaded words & sentences*: nonspecific adjectives or adverbs ('good', 'easy', 'often'), universals ('not', 'never'), modifiers ('just', 'only'), phrases such as 'isn't it reasonable to suppose…?', 'don't you believe…?'
- ✓*Avoid negative constructions*: Responding to neg. constructions can be problematic ('not like' > 'dislike')

2.2.1. Quantitative data collection

- ✓*Avoid double-barreled questions*: No way of knowing which part of the question the answer responded to, e.g.

  'Do you always write your homework and do it thouroughly?'

- ✓*Avoid items that tend to be answered the same way by everyone*.

- ✓*Include positively & negatively worded items* to avoid respondents marking 1 side of a rating scale, e.g.

  ' Learning English is a **burden** for me'

  instead of ' I don't enjoy learning English'.

2.2.1. Quantitative data collection

2.3. The format of the questionnaire

Main parts

- *Title* to provide initial orientation & content expectations.
- *Introduction*

Purpose of the study	No right or wrong answers	Promise confidentiality or anonymity	Thank participants

- *Specific instructions*: Explain with examples how to answer the questions.
- *Questionnaire items*: Main body of questionnaire.
- *Additional information*: contact email, phone, findings.
- *Final 'thank you'*.

2.2.1. Quantitative data collection

Length

No more than 4-6 pages long & ½ an hour to complete.

Layout

- *Booklet format*: it has to *look* short to ease the reading.
- *Appropriate density*: pages shouldn't looked crowded.
- *Sequence marking* to create a sense of structuredness.

Item sequence: 4 principles:

✓ *Mixing up the scales* to prevent repeating same answers.

✓ *Opening questions*: interesting & relatively simple.

✓ *Factual questions* resemble bureaucratic forms & private matters ⟹ some **resistence** if near the beginning.

✓ *Open-ended questions* near the end.

2.2.1. Quantitative data collection

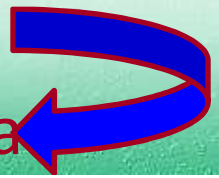
2.4. Developing & piloting the questionnaire

- *Drawing up an item pool* with both qual & quan data.
- *Initial piloting of the item pool*: feedback from colleagues.
- *Final piloting*: administering questionnaire to a group of respondents similar to target population (pre-test).
- *Item analysis*: missing responses, range of responses elicited by each item, & consistency of multi-item scales.

2.5. Strengths & weaknesses of questionnaires

- ✓ Efficient in terms of time, effort & financial resources.
- ✓ Fast & straightforward data processing, & versatile.
- ❖ Simple items results in superficial data.

Thin description of the target phenomena



2.2.1. Quantitative data collection

3. Experimental & quasi-experimental studies

3.1. Experimental study → quan data collection **design**

establish unambiguous **cause-effect** relationships to answer

- What's the reason for…?
- What happens if/ when…?
- What effect does sth have on…?

but difficult to untie interferences of several related factors

2.2.1. Quantitative data collection

E.g. Compare the effects of 2 coursebooks in L2

If we compare 2 learner groups using these books & find some differences, can we say it was due to the difference in the coursebook they used?

NO

group composition teachers learning environments

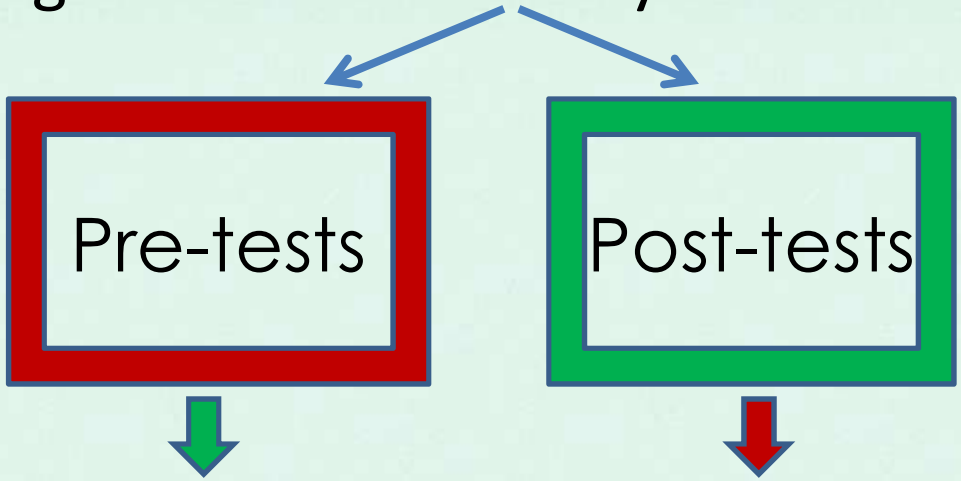
Solution: intervention study with at least 2 similar groups

experimental group control group

exposed to some conditions baseline for comparison

2.2.1. Quantitative data collection

Progress is measured by administering



before the intervention

after the treatment

Difference bet. survey & experimental study

relation bet. variables

1 variable altered to determine effects on other variables

Control group isolates specific effect of target variable

2.2.1. Quantitative data collection

3.2. Quasi-experimental designs

experimental designs without random assignment for comparisons → used in most **educational contexts**

to make causal claims, the effects of the group differences have to be taken into account

Improvements:
- avoid students' self-selection for treatment group
- minimize pre-test differences bet. treatment & control group by
  - using the covariance ANCOVA
  - matching participant in both groups

2.2.1. Quantitative data collection

3.3. Experimental studies in AL → less popular because

Many topics aren't related to intervention (e.g. gender differences, ethnolinguistic variation).

only a few variables can be altered at a time (e.g. in a class many factors & variables at the same time).

3.4. Strengths & weaknesses of experimental designs

✓achieve cause-effect relations & assess innovative education.

✓Pre-test-post-test design controls internal validity threats.

❖reduce external validity (generalizability): artificial context

❖Hawthorne effect.

2.2.1. Quantitative data collection

3.5. Strengths & weaknesses of quasi-experimental designs

✓external validity → authentic learning environments.

❖The *selection bias* → inequality of treatment & control groups. → →

→Outcome differences due to differences in pre-existing characteristics of the groups compared

❖Less effective in eliminating competing hypotheses

❖Large list of possible confounding variables.

❖In educational research → many independent variables ↓ unfeasible for exp. or quasi-exp. designs

2.2.1. Quantitative data collection

4. Collecting quan data via the Internet » Benefits:

- ✓ *Reduced costs*
- ✓ *Convenience of administration*: self-running
- ✓ *Automatic recording* of answers
- ✓ *High level of anonimity* but data security isn't guaranteed.
- ✓ *International access* » good for cross-cultural research
- ✓ *Access to dispersed or specialized populations*
- ❖ *Technical issues*: connection speed & installed software
- ❖ *No sampling strategy:* Self-selected participants difficult to generalize the findings

Compare 2 substrata of sample

Compare web-based results with a non-web survey

References

- Dörnyei, Z. (2007), *Research Methods in Applied Linguistics.* Oxford: O.U.P.
- Dörnyei, Z. & K. Csizér (2012), 'How to Design and Analyze Surveys in Second Language Acquisition Research', in A. Mackey & S. M. Gass, *Research Methods in Second Language Acquisition: A Practical Guide*. Malden: Wiley-Blackwell Publishing Ltd.
- Kvist, C. (2014), *Teachers' Perceptions of Second Language Learning and Teaching: A comparison of two groups of English teachers.* < http://lup.lub.lu.se/luur/download?func=downloadFile&recordOId=4810418&fileOId=4810435> [01-03-2017].

References

- Labov, W. (2006), *The Social Stratification of English in New York City* (2nd ed.). Cambridge: Cambridge University Press.
- Tseng, W-T., Z. Dörnyei, & N. Schmitt (2006). 'A new approach to assessing strategic learning: The case of self-regulation in vocabulary acquisition.' *Applied Linguistics* 27/1: 78-102.