

**AI 507 –  
Research seminar  
AI & Society**

**#RS3 from research  
question to designs**

# Flashlight

# Flashlight

## Theories

Science →  
knowing

Theory →  
looking at

## Why

Prediction

Causality

## What

### Core terms

- Constructs (= most abstract)
- Concepts (= a little more concrete)
- Variable = what we actually measure

Categorical



Ordinal



Parametric



Independent versus  
dependent variables

Links between variables

## How

### Theoretical statements

- Hypotheses  $A \rightarrow B$  /  $A > B$
- Research questions  $A:B?$
- Propositions:  $A =$
- Assumptions:  $A \rightarrow B \mid A =$



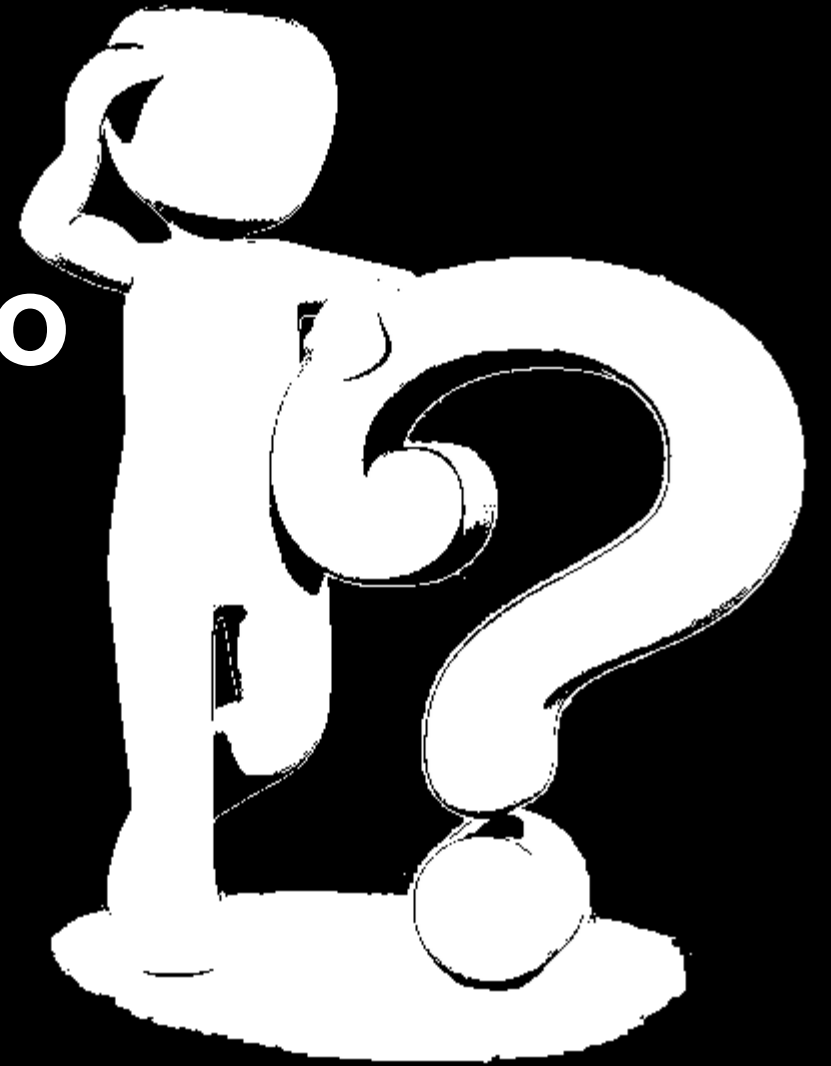
### Principle of falsification

### Theoretical linkage

### Operational linkage

How to find a research  
question

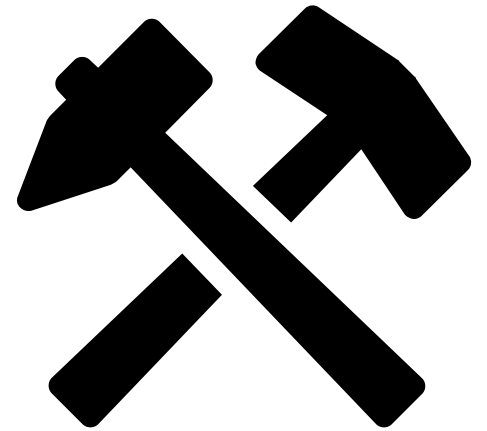
**Do you have questions to  
last Tuesdays' lesson?**



# Then I hope,

at the end of today's lesson, you will...

- ☐ ...know what the differences between qualitative and quantitative research methods
- ☐ ...have gotten to know four prominent methods for working with human data
- ☐ ... can make informed decisions to choose a methodological approach to answer your research question/ test your hypothesis



# Qualitative versus quantitative methods

# Think back to the research ideas you brainstormed last week

(Smith, 2024)

- Now: how do we answer our research question/ test our hypothesis?
- Two research paradigms

## Qualitative

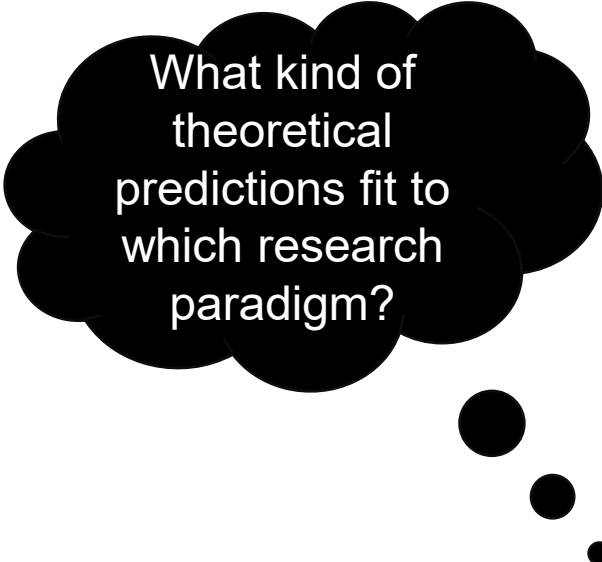
- Explores the how and why
- Focus on the subjective experience of the human data owner, the research subject
- Enables in-depth insights on topics that are not well understood
- Rooted in constructivist/ normative research traditions (→humanities)

## Research questions

## Quantitative

- Tests the "what," "how many," or "how often" using numerical, statistical data
- Focus on generalisable data and descriptions of the population
- Enables testing theoretical statements
- Rooted in positivist research traditions (→ natural sciences)

## Hypotheses



What kind of theoretical predictions fit to which research paradigm?

# “Quali-quant wars” (Smith, 2024)

- Fundamentally different function in the research process – not necessarily better or worse

	Qualitative	Quantitative
When	Explorative, hypothesis generating, theory-building	Confirmatory, Hypothesis testing, theory testing
Whom	Few theoretically-sampled respondents	Many representatively sampled
What	Summarizing, interpreting, categorizing data	Statistical analysis of data (Frequentist or Bayesian)
How	Results expressed in words	Results expressed in numbers

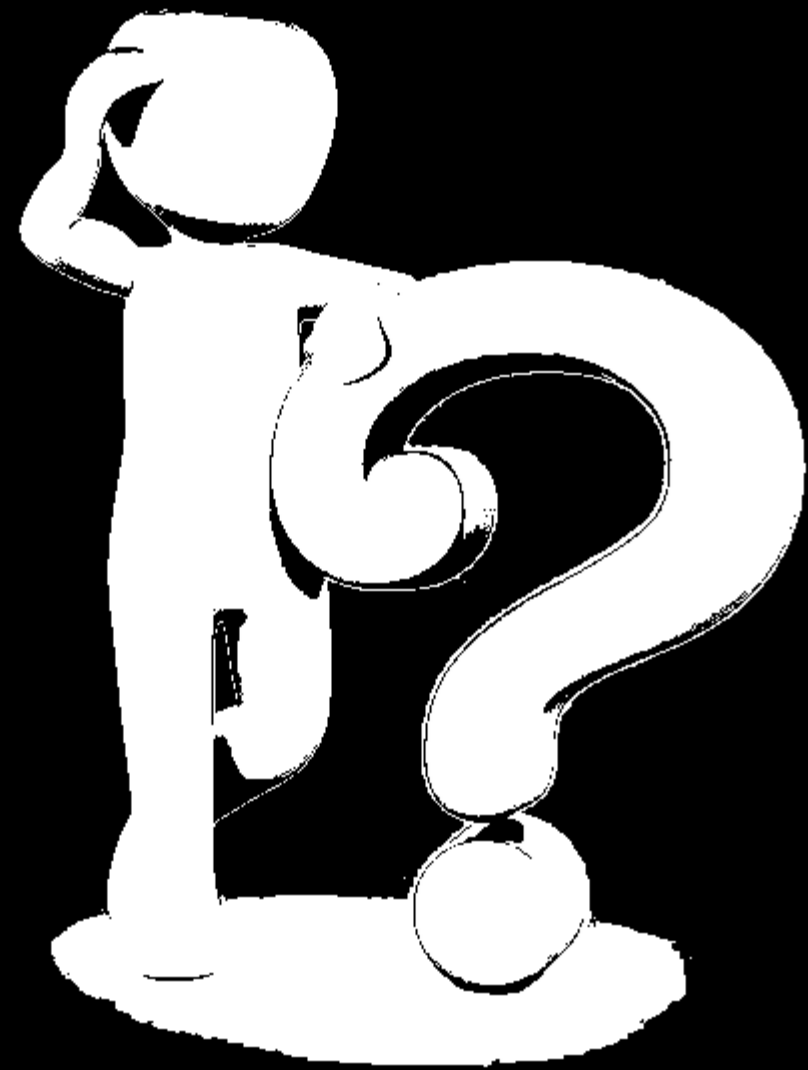
- Mixed-methods combine the best of both worlds (“pragmatic approach”)



**“It is tempting, if the  
only tool you have is a  
hammer, to treat  
everything as if it were  
a nail”**

(Maslow, 1966)

**Questions?**



**But how do you get the  
data?**

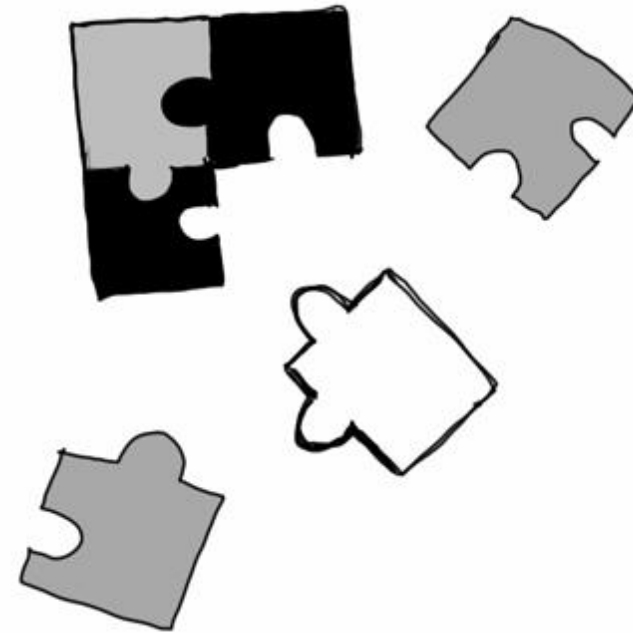
# Four prominent methods of data collection

## Quantitative

- Interviews
- Manual qualitative content analyses

## Quantitative

- Online-questionnaires (“surveys”)
- Experiments



# Interviews



# Interviews (Knott et al., 2022)

- Follow a familiar logic of human interaction: people talk with each other, interact and pose and answer question
- Provide a space for extended conversations that allow the researcher insights into how people think, how they feel, and what they believe
- Asymmetric: Interviewer poses questions – interviewee answers them
- Despite the “natural feeling”: posing the right questions and being a good interviewer is a skill on its own!

## From the archive, 6 May 1977: Yours regretfully, R. Nixon

The Frost/Nixon interview is the closest we will get to a trial, but it is unlikely to have changed public opinion



David Frost interviewing former US president Richard Nixon, 1977. Photograph: John Bryson/Time Life Pictures/Getty Images

# Interviews

(Knott et al., 2022)

- Research interviews do not try to “reveal” misdeeds – but they also want to create an atmosphere where things can be said
- Differ in their level of structure
  - **Highly structured:** All questions are defined and standardized in advance – often including offering answers (how much do you like... not at all / somewhat/ very much)
  - **Highly unstructured:** Narrative, free floating approach
  - **Semi-structured:** Using a topic guide to ensure that all relevant aspects are covered while allowing for free formulations, different orders of questions and additional topics

# Semi-structured interviews (Knott et al., 2022)

## Topic-guide

- Start with easy questions (“icebreakers”)
  - E.g., socio-demographic data (if needed)
  - Broad questions to the topic → Help the interviewee to “ease” into the conversation
- More direct questions on the research topic
  - Start with concrete questions (how do you do XYZ, what is your experiences with...)
  - Follow with more abstract questions (how you feel with regards too...)
- Close-up: Open topics, something you want to have more details on, something that you missed

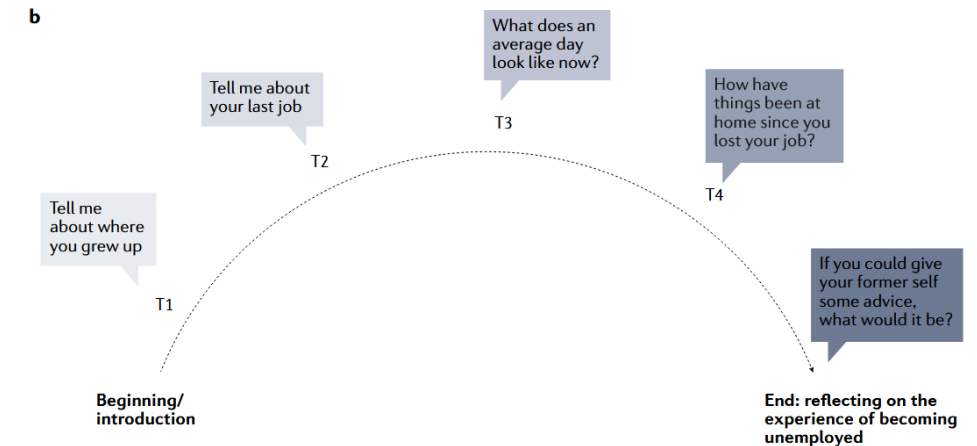


Fig. 1 | How to develop a topic guide arc. a | Elaborated topics the researcher wants to cover in the interview and example questions. b | An example topic arc. Using such an arc, one can think flexibly about the order of topics. Considering the main question for each topic will help to determine the best order for the topics. After conducting some interviews, the researcher can move topics around if a different order seems to make sense.

(Knott et al., 2022, p.3)



<https://www.pickpik.com/couple-guy-girl-love-dating-people-69120>

# Selecting interview partners

(Dworkin (2012, 2024); Glaser & Strauss, 1967; Francis et al., (2010), White & Cooper, 2022))

- Often guided by the principles of **grounded theory**
- Theory-building is grounded in the data
- Systematic data analysis aimed at generating theories
- Collecting data, analysing it, interpreting it are related to each other – a collection of interwoven methods
- Central for data collection: **Theoretical sampling**
  - Start with an initial data collection, identify first themes (“codes”) in the data and then collect further data till you reach **saturation**
  - **Saturation** = *theoretical* (new interview partners do not change the developed theory), *code saturation* (now new codes from new interview partners), *information saturation* (new interviewees do not provide new information)
- Often something between 20 to 50 interviews



Barney G. Glaser via  
[https://de.wikipedia.org/wiki/Barney\\_G.\\_Glaser](https://de.wikipedia.org/wiki/Barney_G._Glaser)



Anselm Strauß via  
<https://sociopedia.co/author/strauss-anselm>

# Selecting interview partners

(Fassinger, 2005; Patton 2002; 2015)

- Thinking about sampling requires thinking about inclusion and exclusion criteria: who or what do we want to hear from, and who or what do we not want to hear from.
- Qualitative researchers often aim to sample for
  - diversity of perspectives (maximum variation or maximum heterogeneity sampling)
  - typicality of perspectives (homogeneity sampling),
  - or something in between
- Different types do not allow for generalization but because heterogeneous people can plausibly provide a more complex picture of the phenomenon of interest
- Interviewees as experts for their own lived experience



# Conducting the Interview (Gläser & Laudel, 2010)

- Online or offline?
- One or two interviewers?
- Informed consent
- Record, record, record!
- Not too many interviews in a row
- Realistic time frames

## Afterwards:

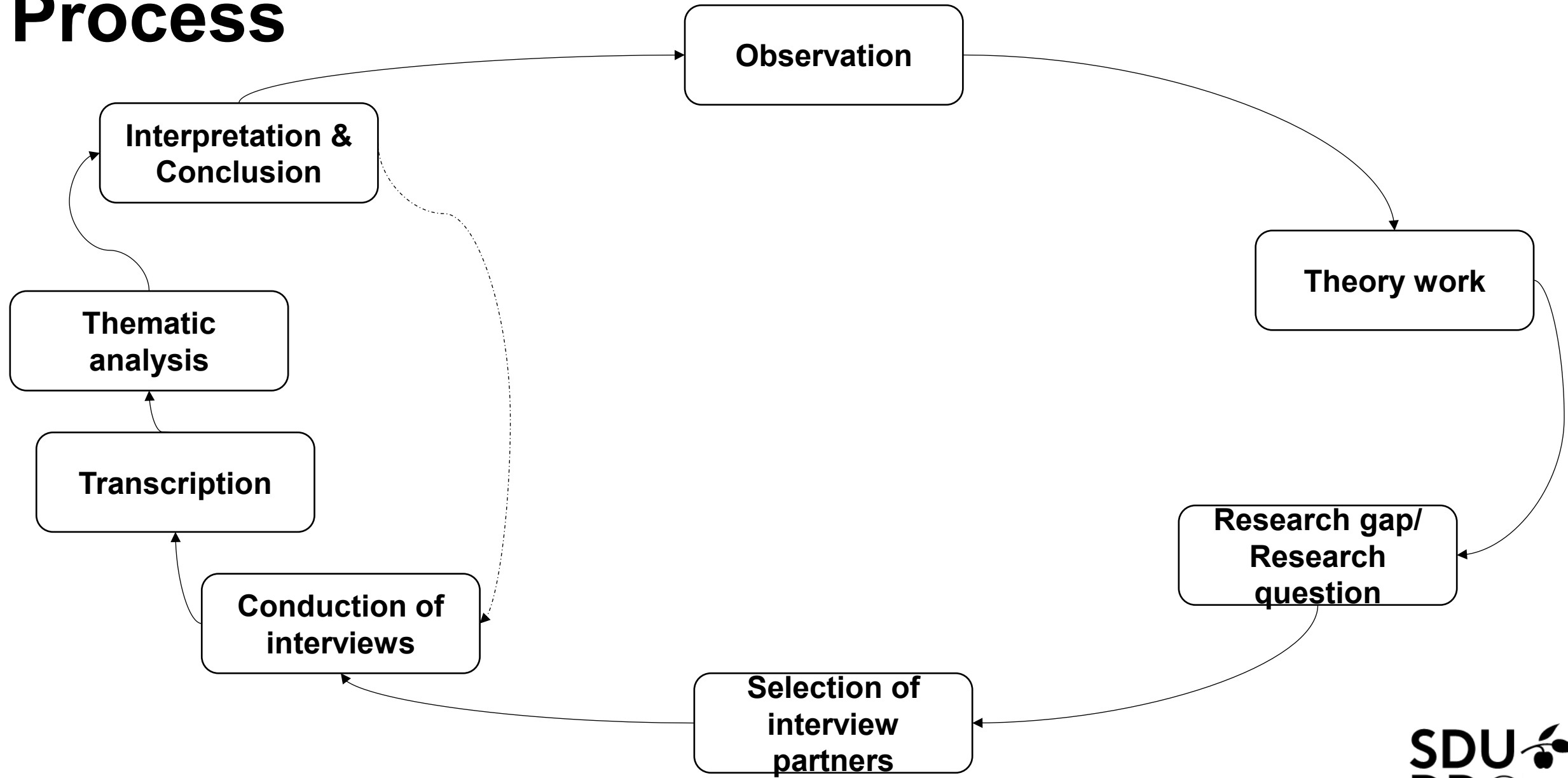
- Transcription
- Analysis

## AI-enhanced transcription at SDU (<https://libguides.sdu.dk/nvivo/transskribering>)

- Transcriber via Ucloud (has to be registered in advance, <https://docs.cloud.sdu.dk/Apps/transcriber.html>)
- SDU microsoft 365: 300 min./ month in Word



# Process



**Questions?**



**Time for a break**



# Content Analysis

# Content analysis

(Lacy et al., 2015; Mayring, 2000; Riffe et al., 2023)

- All human exchanges involve messages (content) - thus content analysis is particularly important for the study of communication
- Examples include interviews, social media posts, comment sections, news coverage, visual content, lyrics, advertisement etc.
- Systematic approach to reduce large amounts of (communicated) data to meaningful and intersubjectively reliable categories
- Central part: Assigning the data in a rule-based approach to categories
- Can be one part of the grounded theory process (then coding, theory, and recruiting are interwoven)
- Here: Based on pre-existing material and with a set research question or even hypotheses (i.e., more qualitative/ more quantitative forms)

# Content analysis

(Lacy et al., 2015; Mayring, 2000; Riffe et al., 2023)

## Qualitative content analysis

- Inductive thematic coding
- Developing of themes/ codes based on the text (sometimes starting with the interview guide as starting point)
- Refinement over the process of engagement with the material
- Codebook as the outcome of the process
- E.g., via NVivo

→ **Exploratory & theory-developing**

## Quantitative content analysis

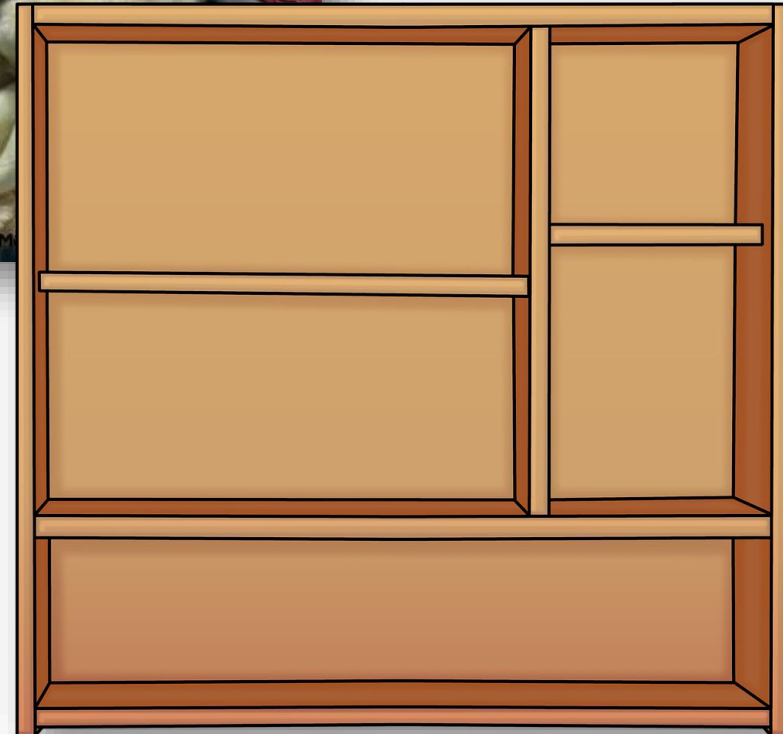
- The systematic assignment of communication content to categories according to rules specified in a coding protocol, and the analysis of relationships involving those categories using statistical methods
- Here you start with the codebook and then assign text to the categories in the codebook – often using 0/1 codes
- E.g., via spreadsheets

→ **More theory-testing**

# Qualitative-oriented content analysis

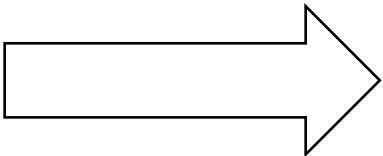
(Mayring, 2000, Riffe et al., 2023)

- Combines the technical expertise of quantitative content analysis (can handle large amounts of material) with a qualitative interpretative approach to capture latent meaning
- Strictly rule-based and therefore highly intersubjectively reliable
- Rules are based on psychological and linguistic theory of everyday text comprehension
- Codebook is developed inductively and then applied to the text



# The Codebook

Category	Definition	Anchor example	Boundary

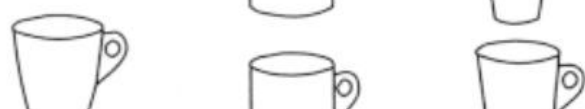


- Very rule-based approach
- Trained coders instead of majority votes (often in computer science)
- Can inform machine-learning/ AI-pipelines

Cup



Bowl

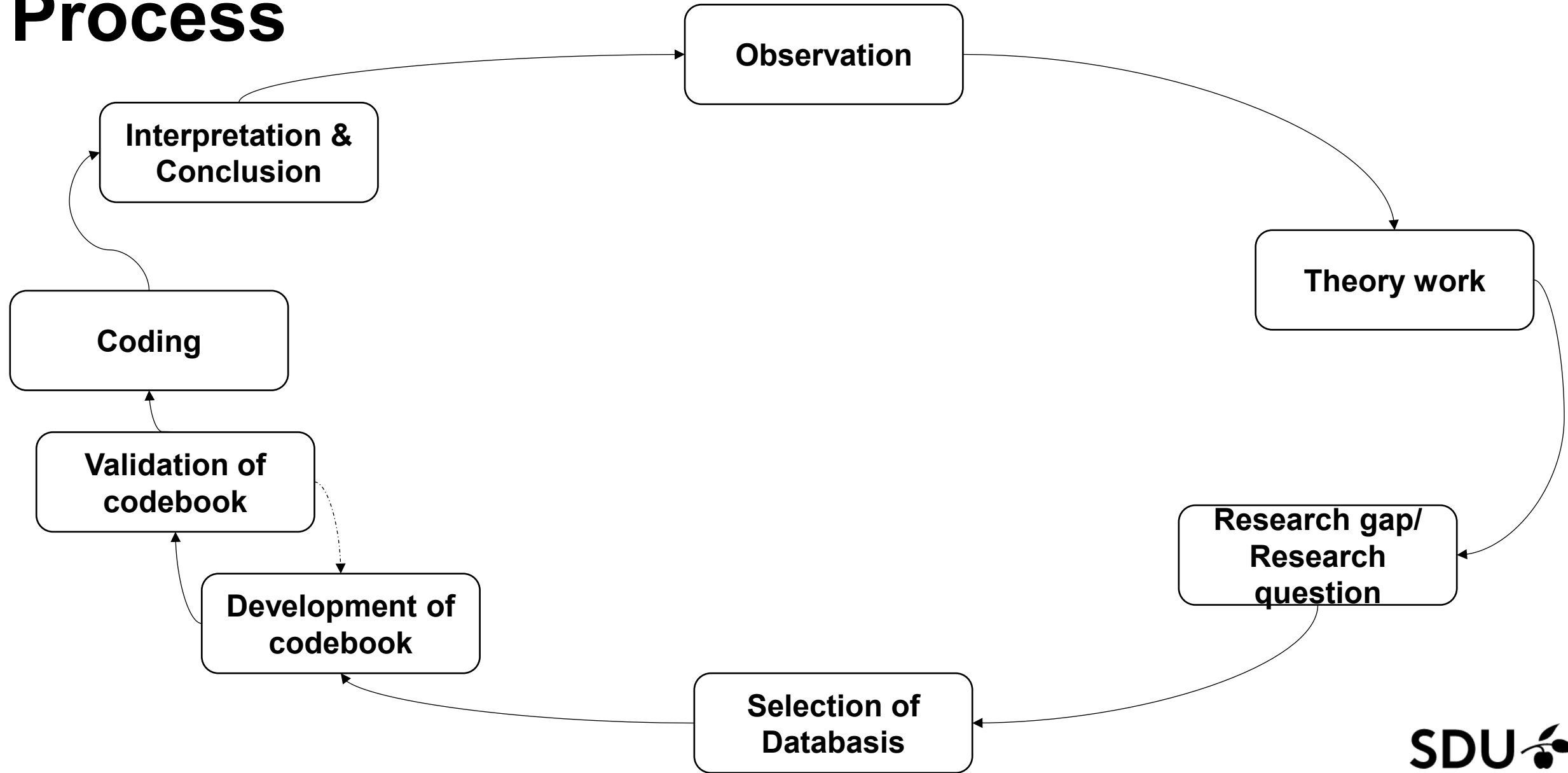


Glass

# Database

- Define the population of interest (e.g., Danish news coverage)
- Sampling
  - Convenience sampling/ snowball sampling (→ not representative, potentially biased)
  - Purposeful sampling (e.g., theoretical sampling in grounded theory)
  - Stratified sampling (e.g., articles from DR, TV2, Berlingske etc.), ensures that different subgroups are equally represented
  - Random sampling from the population (→ allows for generalisability)
- For qualitative approaches: often ~300 articles/ posts etc. – the more quantitative, the more can be coded
- If the codes should inform machine-learning models more is better

# Process



**Questions?**



# Surveys & Experiments

# Surveys

(Atkeson & Alvarez, 2018; Brady, 2000, Oberski, 2018, Zieser et al, 2025)

- Quantitative form of asking people for their opinions, feelings, behavioral intentions etc.
- Aims for results that are applicable to a wider population (→ random/ representative samples, avoidance of sampling biases)
- Administration of questionnaires that include measures for the variables of interest
- Nowadays often as either computer-assisted web-interview (great to avoid people not finishing) or self-administered survey (great for anonymity and more honest answers)
- Typically, closed questions (in contrast to interviews)
- Often ordinal or parametric variables (e.g., how much do you like... 1 = not at all, 5 = very much)
- Similar to interviews: The order of question matters for responses and prior questions can influence what people say next

# Surveys

(Atkeson & Alvarez, 2018; Brady, 2000, Oberski, 2018, Zieser et al, 2025)

- Formulation of questions is key: Avoid double-barreled questions and negatively formulated questions (do you agree that using this method is a bad idea or not? 1 = disagree, 5 = agree)
- Your participants must understand all the different words in your question and construct the meaning behind it
- Rule of thumb: Find someone else to blame – take measurements that have been tested before (ideally in your context)
- If possible, use direct questions ( “Is AI good for society”) instead of more complex agreement statements (“do you agree with the following statement: AI is good for society”) – but be aware of social desirability bias

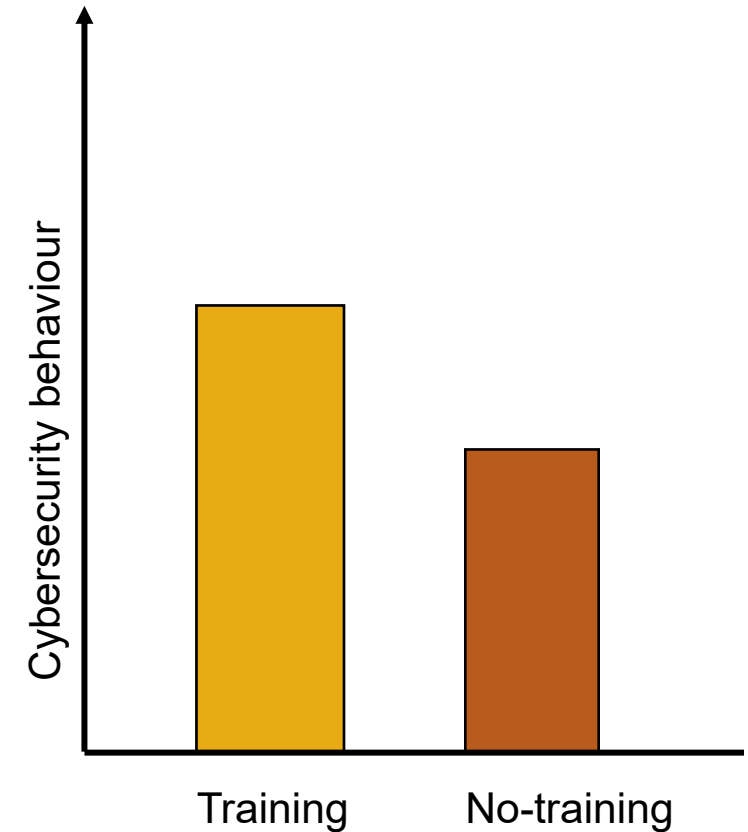


# Analyses

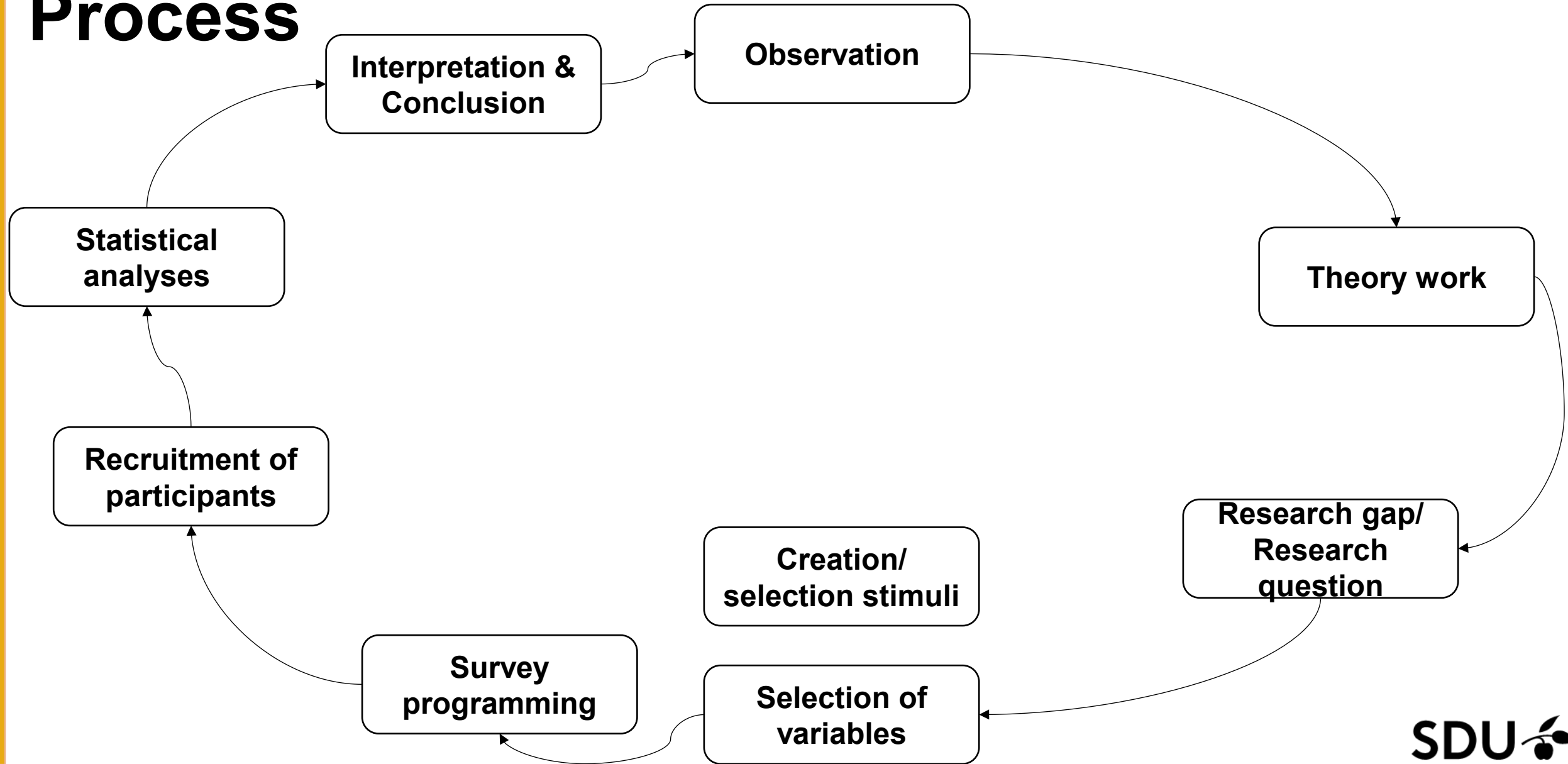
- Quantitative, often hypothesis testing
- Typically describing the distributions of the measures of interest (e.g., average AI-literacy, average age)
- Attention: Cross-sectional data does not allow for causal conclusions
- Often focus on correlations (e.g. older people are less literate when it comes to AI)
- Even if predictive models such as regressions are used: No final decision about causal direction!
- However, competing models can be tested to falsify causal assumptions (increase the plausibility that there is a causal process)

# Can be combined with an experiment (Field & Hole, 2023)

- Last week, we already talked about independent and dependent variables
- Experiments compare at least two groups that differ from each other on the independent variable: Experimental group and control group
- We then measure differences on the dependent variable between these two groups
- Participants should be randomly assigned to these groups
- Groups should be large enough to control for “confounding variables” and to detect the smallest effect size of interest
- If all of this is the case and we observe a statistically significant difference between the groups: causal hypotheses can be retained (if there is no difference – hypothesis is falsified)



# Process




**Where do I have to do with  
which method?**

# Case study: CPH 5 AI start-ups (<https://www.seedtable.com/>)

[illegible]


<https://www.easysize.me/>




Not sure about the size?

## Hassle-free size advice, automated.

EasySize




**Gulnaz K.** · 2


Helping online shoppers buy fashion that fits 

Kopenhagen, Hauptstadtregion, Danmark · [Kontaktinfo](#)

500+ Kontakte

 Cathrine Marie Lamm Nielsen ist ein gemeinsamer Kontakt.

[Vernetzen](#)
[Nachricht](#)
[Zur Website](#)
[Mehr](#)


 Profil mit Premium verbessert


### Info

I'm an entrepreneur, passionate about making fashion more sustainable

### Im Fokus

Beitrag

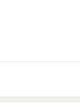
 After a stellar Black Friday/ Cyber Monday and holiday sales...



31 · 3 Kommentare

Beitrag

In the last few weeks, Shopify announced some major product...



Shopify Has A Plan For E-commerce Domination And It...  
forbes.com

43 · 2 Kommentare

Link

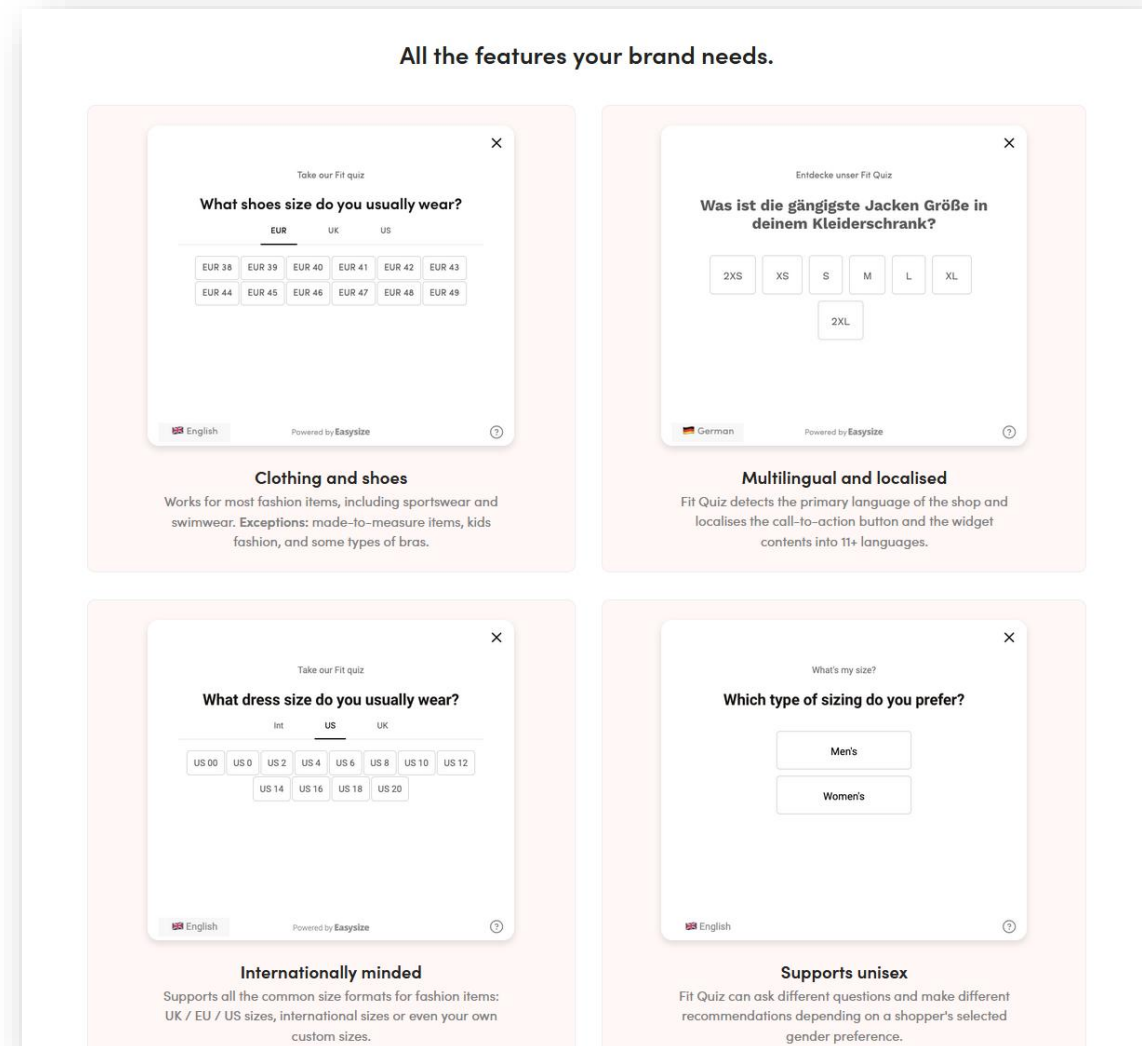
## EasySize

EasySize – making fashion more sustainable through data and AI. [easySize.me](#)

EasySize predicts the perfect size and fit for fashion online shoppers – for any brand, type of fashion and geo – to boost user confidence, increase sales and reduce returns. ...

# How human-data could play a role in the build-up phase

- Research question stage
    - How do customers describe their size?
    - Which concepts are familiar to them?
  - Hypothesis stage
    - Preferences predict fit
    - Fit reduces returns and increases profit
- Research questions stage
- How do people engage with the tool?

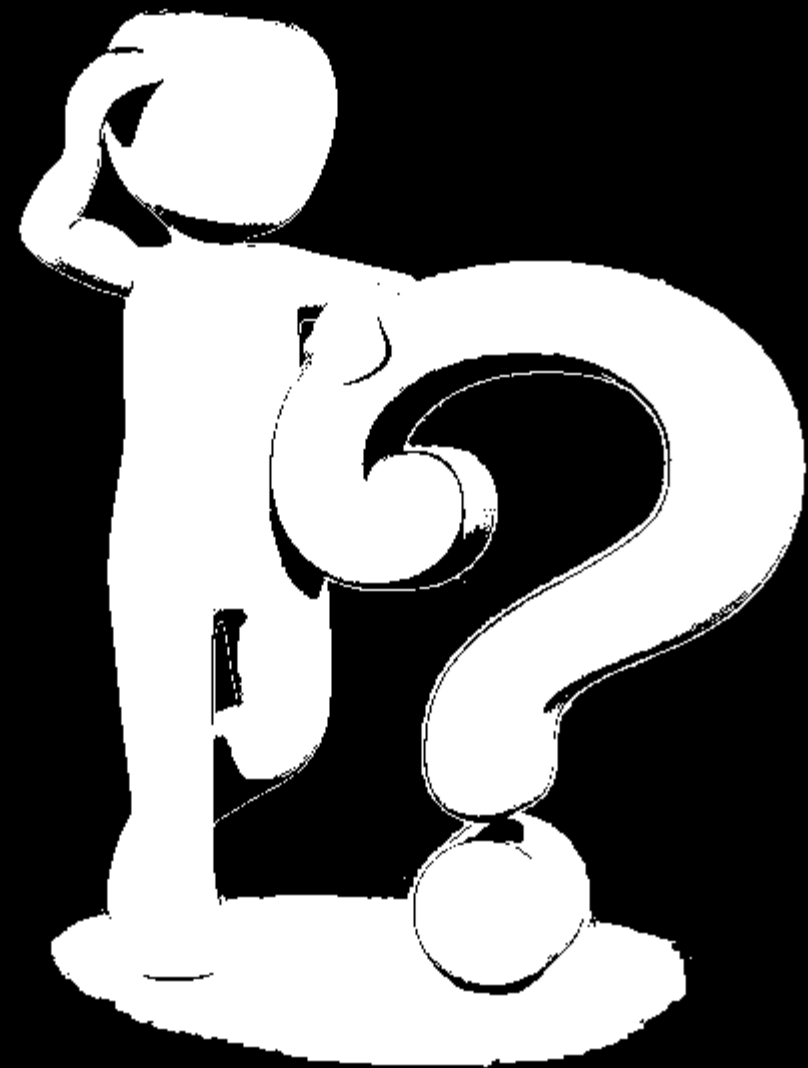


<https://www.easysize.me/>

# Now it's your turn

- Get back to your research ideas from last week
- Which methods would you choose and why?

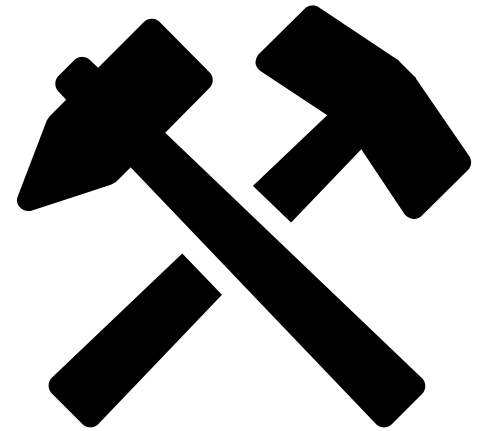
**Questions?**



# Then I hope,

You now...

- ☐ ...know what the differences between qualitative and quantitative research methods
- ☐ ...have gotten to know four prominent methods for working with human data
- ☐ ... can make informed decisions to choose a methodological approach to answer your research question/ test your hypothesis
- ☐ Bring your ideas next week – tomorrows' session is cancelled



**AI 507 –  
Research seminar  
AI & Society**

**Have a wonderful rest of  
the day!**

# References

- Atkeson, L. R., & Alvarez, R. M. (2018). Introduction to polling and survey methods. In L. R. Atkeson & R. M. Alvarez (Eds.), *The Oxford Handbook of Polling and Survey Methods* (p. 0). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780190213299.013.34>
- Brady, H. S. (2000). Contributions of survey research to political science. *Political Science and Politics*, 33(1), 47–57.
- Oberski, D. (2018). Questionnaire science. In L. R. Atkeson & R. M. Alvarez (Eds.), *The Oxford Handbook of Polling and Survey Methods* (p. 0). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780190213299.013.21>
- Knott, E., Rao, A. H., Summers, K., & Teeger, C. (2022). Interviews in the social sciences. *Nature Reviews. Methods Primers*, 2(1), 73. <https://doi.org/10.1038/s43586-022-00150-6>
- Dworkin, S. L. (2012). Sample size policy for qualitative studies using in-depth interviews. *Archives of Sexual Behavior*, 41(6), 1319–1320. <https://doi.org/10.1007/s10508-012-0016-6>
- Dworkin, S. L. (2024). Sample size policy for qualitative studies using in-depth interviews. *Archives of Sexual Behavior*, 53, 3701–3704. <https://doi.org/10.1007/s10508-012-0016-6>
- Glaser, B., & Strauss, A. (1967). *Discovery of grounded theory: Strategies for qualitative research*. Routledge. <https://doi.org/10.4324/9780203793206>
- Gläser, J., & Laudel, G. (2010). *Experteninterviews und qualitative Inhaltsanalyse [Expert interviews and qualitative content analysis]*. VS Verlag für Sozialwissenschaften.
- Francis, J. J., Johnston, M., Robertson, C., Glidewell, L., Entwistle, V., Eccles, M. P., & Grimshaw, J. M. (2010). What is an adequate sample size? Operationalising data saturation for theory-based interview studies. *Psychology & Health*, 25(10), 1229–1245. <https://doi.org/10.1080/08870440903194015>

# References

- Fassinger, R. E. (2005). Paradigms, praxis, problems, and promise: Grounded theory in counseling psychology research. *Journal of Counseling Psychology*, 52(2), 156–166. <https://doi.org/10.1037/0022-0167.52.2.156>
- Lacy, S., Watson, B. R., Riffe, D., & Lovejoy, J. (2015). Issues and best practices in content analysis. *Journalism & Mass Communication Quarterly*, 92(4), 791–811. <https://doi.org/10.1177/1077699015607338>
- Mayring, P. (2000). Qualitative content analysis. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 1(2).
- Patton, M. Q. (2002). Two decades of developments in qualitative inquiry: A personal, experiential perspective. *Qualitative Social Work*, 1(3), 261–283. <https://doi.org/10.1177/1473325002001003636>
- Patton, M. Q. (2015). *Qualitative research & evaluation methods: Integrating theory and practice* (4th ed.). SAGE.
- Riffe, D., Lacy, S., Watson, B. R., & Lovejoy, J. (2023). *Analyzing media messages: Using quantitative content analysis in research* (5th ed.). Routledge. <https://doi.org/10.4324/9781003288428>
- Smith, T. (2024). Qualitative and quantitative research | Social Sciences and Humanities | Research Starters | EBSCO Research [<https://www.ebsco.com/research-starters/social-sciences-and-humanities/qualitative-and-quantitative-research>]. EBSCO. <https://www.ebsco.com>
- White, R. E., & Cooper, K. (2022). Grounded theory. In R. E. White & K. Cooper, *Qualitative Research in the Post-Modern Era* (pp. 339–385). Springer International Publishing. [https://doi.org/10.1007/978-3-030-85124-8\\_9](https://doi.org/10.1007/978-3-030-85124-8_9)
- Zieser, M., Lorenz, T., & Voit, V. (2025). Conducting representative surveys and impact assessments: An overview of survey. In R. Holzmann (Ed.), *Rigorous and Comprehensive Evaluation of Financial Education and Literacy Collected Articles from 2024–2025* (pp. 241–263). Österreichische National Bank.