

# Grounded Theory

Wilhelm Arthur Sandberg Damsleth  
2024-03-06 / IN5000 Guest Lecture

# Agenda and Learning Outcome

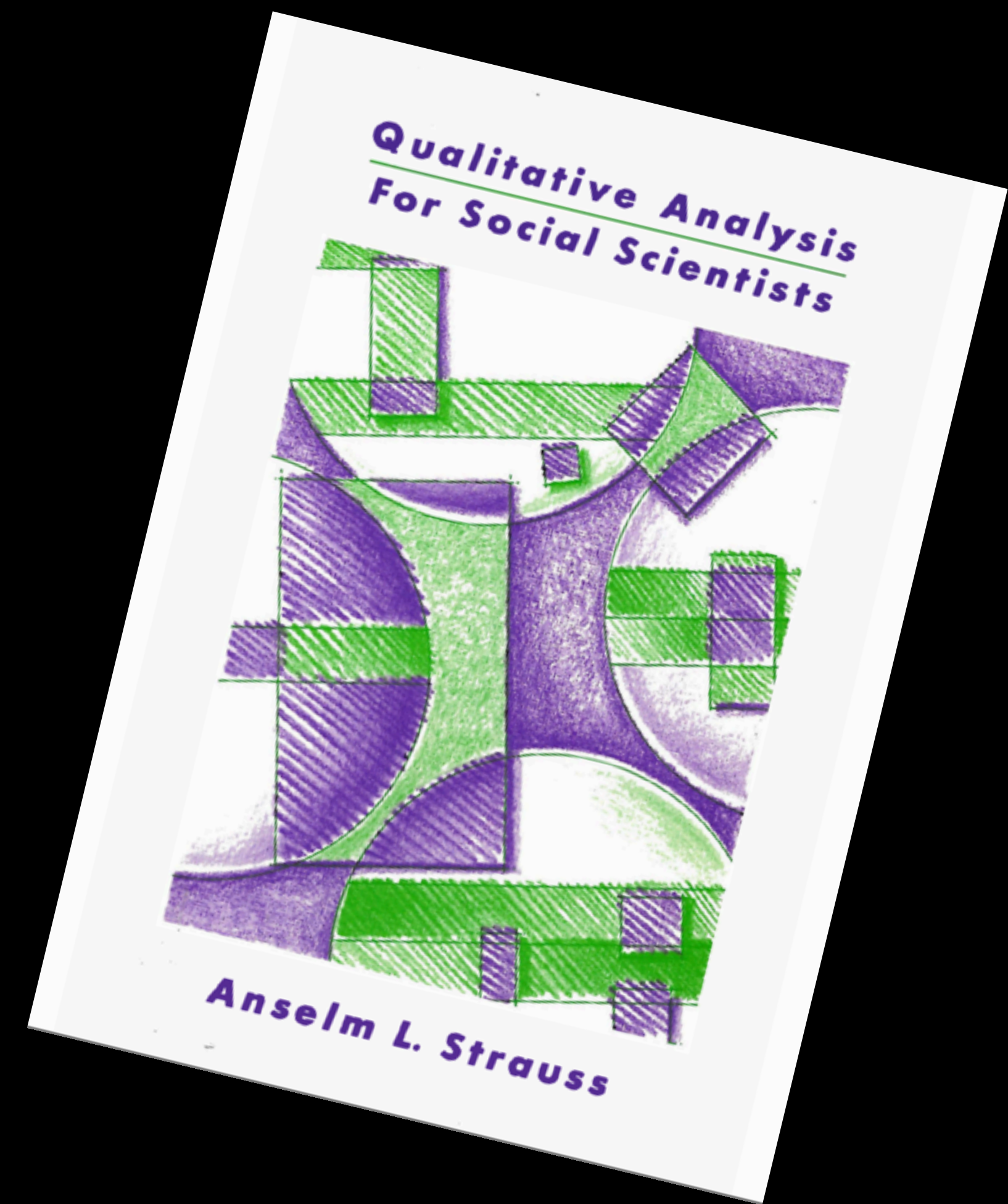
In this lesson, we will go through...

- Some theory
  - What is GT?
  - GT Pros&Cons
  - GT Primer (“crash course”)
  - Comments on codes
  - Comments on sources
- Some static examples
- Simple coding exercise

After the lesson, you will...

- Know about Grounded Theory as a methodology
- Understand the basic principles of Grounded Theory, its pros & cons
- Be able to attempt simple GT analysis of data

Strauss, Anselm L. (1987)  
*“Qualitative Analysis For Social Scientists”*



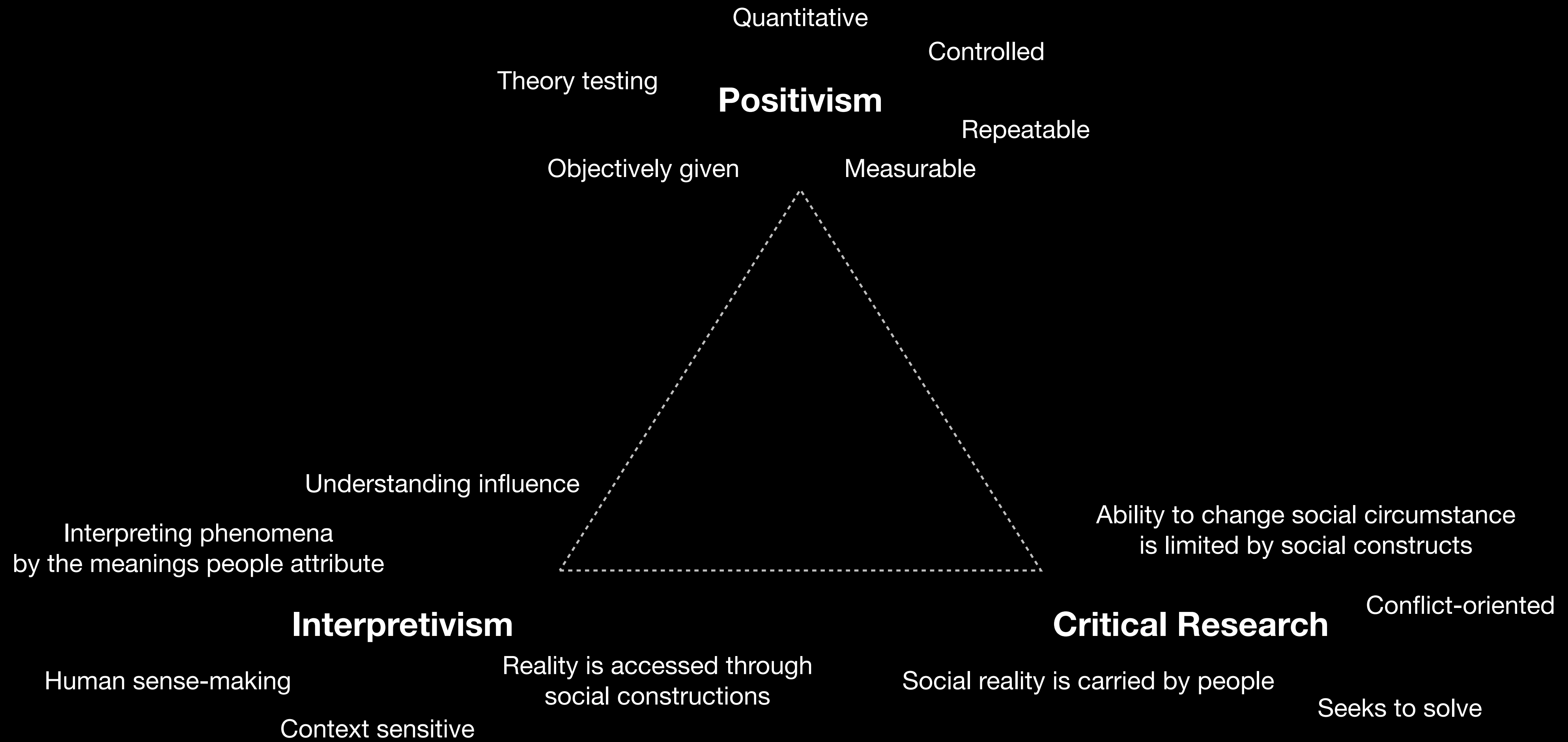
# GROUNDING THEORY IS

- A methodology and analytical approach for developing theory that is *grounded* in your data
- A generative process
- An opportunity to learn more than the sum of your data
- An opposition to *read-then-do-then-write* (Crang & Cook, 2007)?

# GROUNDING THEORY ISN'T

- Grounding Theory is not journalism\*
- Grounding Theory is not a quantitative analysis\*
- Grounding Theory is not an excuse

\* Grounding Theory –  
Quantitative Journalism?

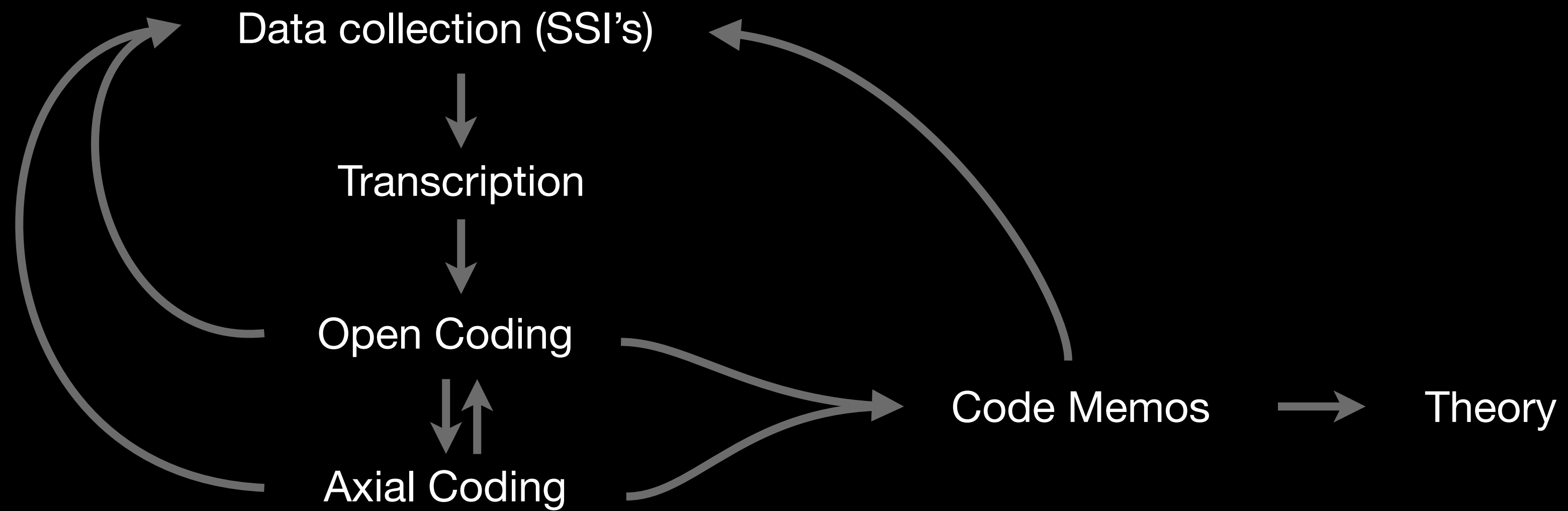


# GT Analysis: Main Elements

1. Concept-Indicator Model
2. Data Collection
3. Coding
4. Core Categories
5. Theoretical Sampling
6. Comparisons
7. Theoretical Saturation
8. Integration of the Theory
9. Theoretical Memos
10. Theoretical Sorting

(Strauss 1987, p. 23)

# GT Primer





# Transcribing

- Accurate and precise data forms the crucial base of GT analysis
- Accuracy – Nuances
- “Transcribing sucks”
  - Deal w/ it!
- Use a good template



“My first fully transcribed interview”  
1:26:56 – 25 pages – 11.009 words  
... one of many.

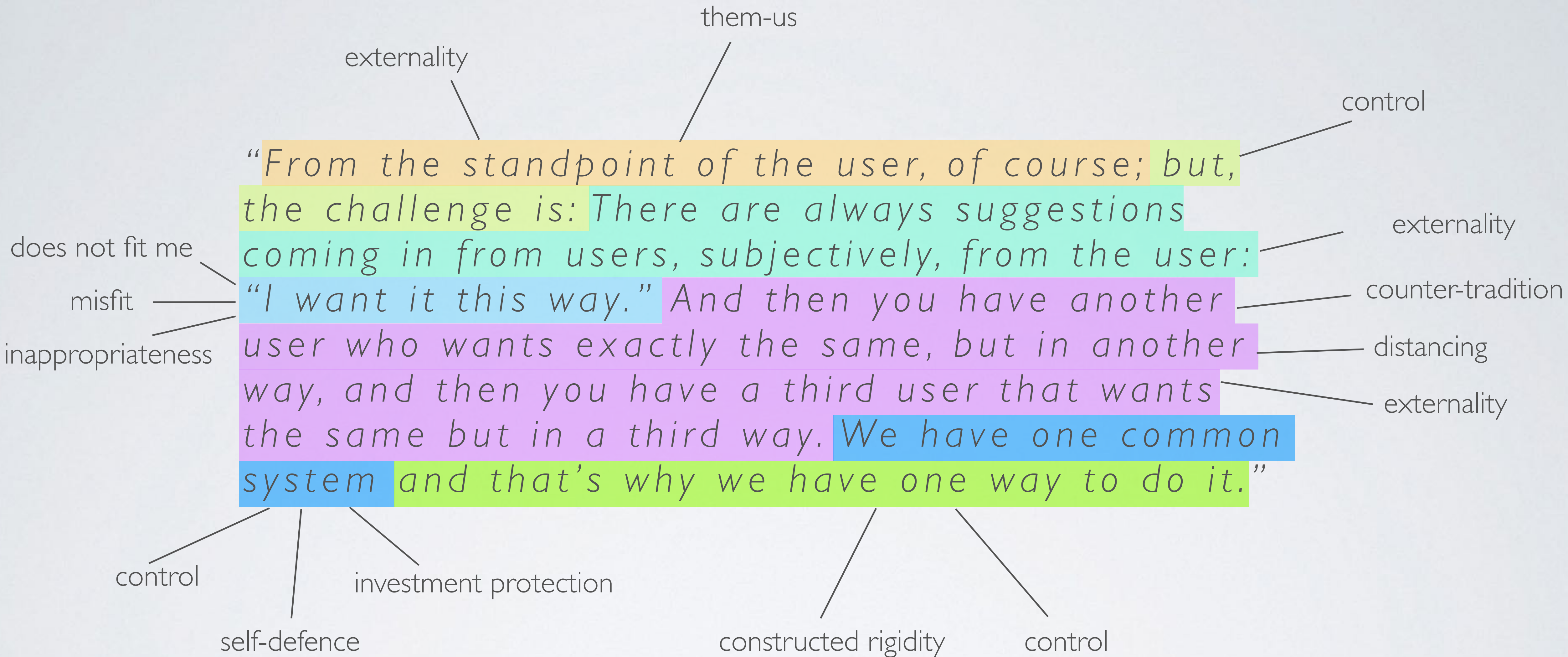


# Open Coding

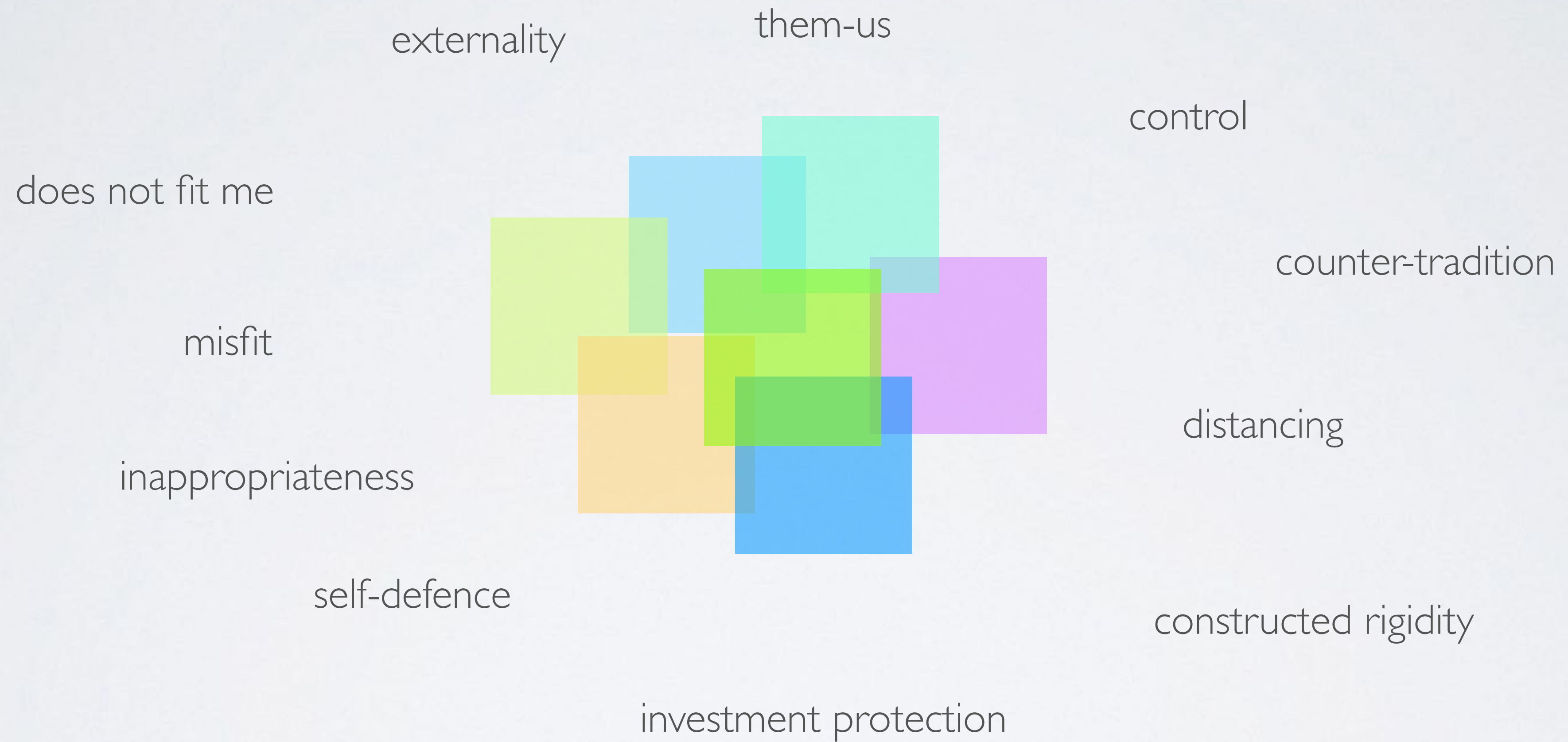
*“Coding. The general term for conceptualizing data; thus, coding includes raising questions and giving provisional answers (hypotheses) about categories and about their relations. A code is the term for any product of this analysis (whether category or a relation between two or more categories).”*

*(Strauss 1987, p. 20)*

- Analyze and assign codes to your data
- Use *constructed codes* or *in vivo codes*
- Coding paradigms (Strauss 1987, p. 27-28)
  - conditions
  - interaction among the actors
  - strategies and tactics
  - consequences



# Codes



	Codes in use	References coded
Interview 7	69	434
Interview 8	89	369
Total		803

- Coding is highly personal
- Coding paradigms (Strauss 1987, p. 27-28)
  - conditions
  - interaction among the actors
  - strategies and tactics
  - consequences

	Interview 7	Interview 8
Too difficult	17	19
Avoiding Microwork	11	11
Constructed Rigidity	20	2
Control	19	3
Working Around	9	13
Telephone	11	10
Manual Routines	11	9
Not my job	11	9
Time-consuming	8	12
Competency	14	5
Manual Automatization	14	5
Competencial inadequacy	6	11
Fails to Automate	8	8
Future System	14	2
Detective work	2	13
Faith in the Construct	13	2
Backstage, No Knowledge of	11	3
Bad UI	9	5
Compliance	7	7
ERP System by Name	12	1

# Axial Coding

(“Grouping” or “Categorizing” along Dimensions)

*Dimensionalizing: A basic operation of making distinctions, whose products are dimensions and sub dimensions.*

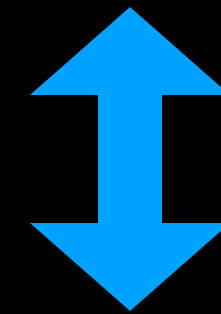
*Category: Since any distinction comes from dimensionalizing, those distinctions will lead to categories.*

(Strauss 1987, p. 21)

- Group codes into categories that have axial variability represented in the codes



*Are computer systems developed by upper management perceived as less relevant in the users job than those from colleagues?*

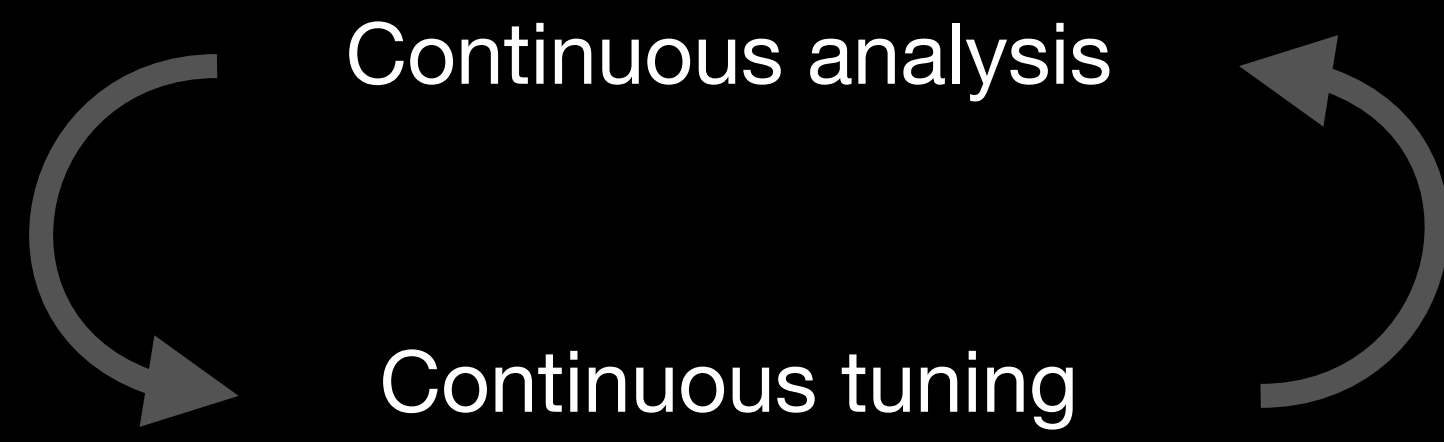


*Are computer systems developed by colleagues perceived as more relevant in the users job than those from upper management?*

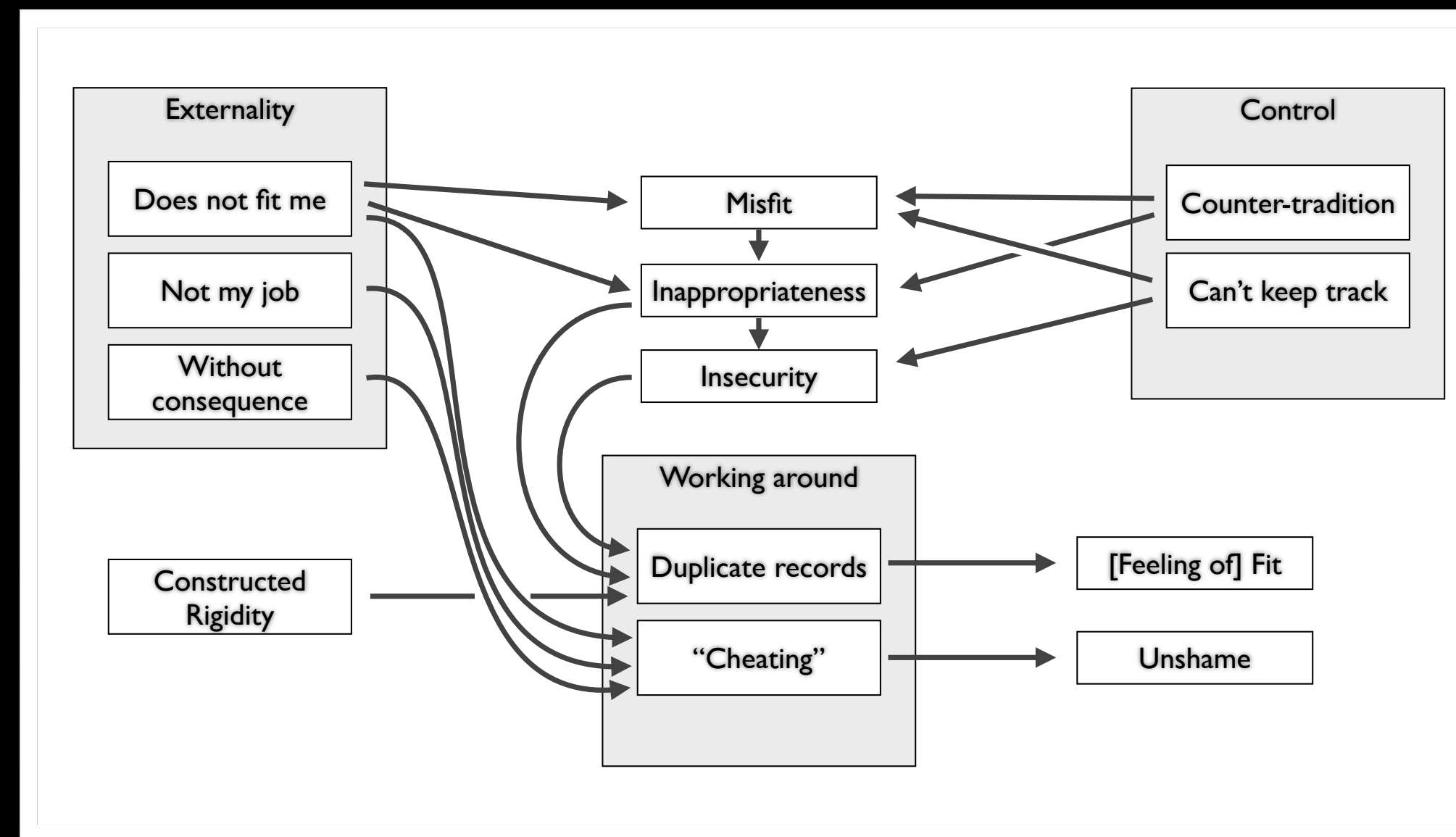


# Memoing and Analysis

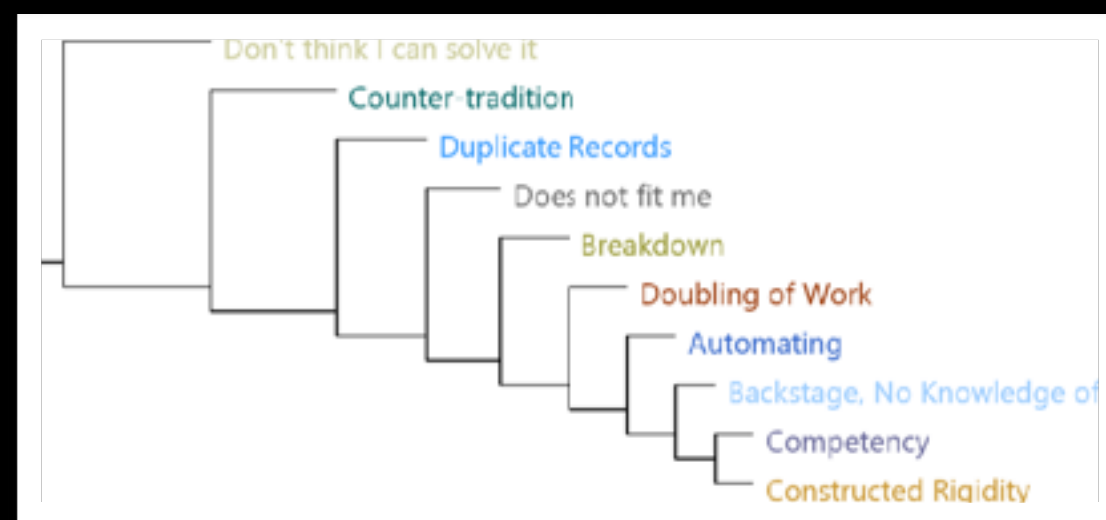
- When discovering a relationship between concepts (codes), categories and dimensions – write a memo!
- Code Memos are linked to one or several codes, categories or relationships
- Write Code Memos immediately when the idea(s) strike(s) you! Do NOT wait – the idea is fleeting, your data is not!
- Code Memos – or paper to publish? No form requirement – only structure required!



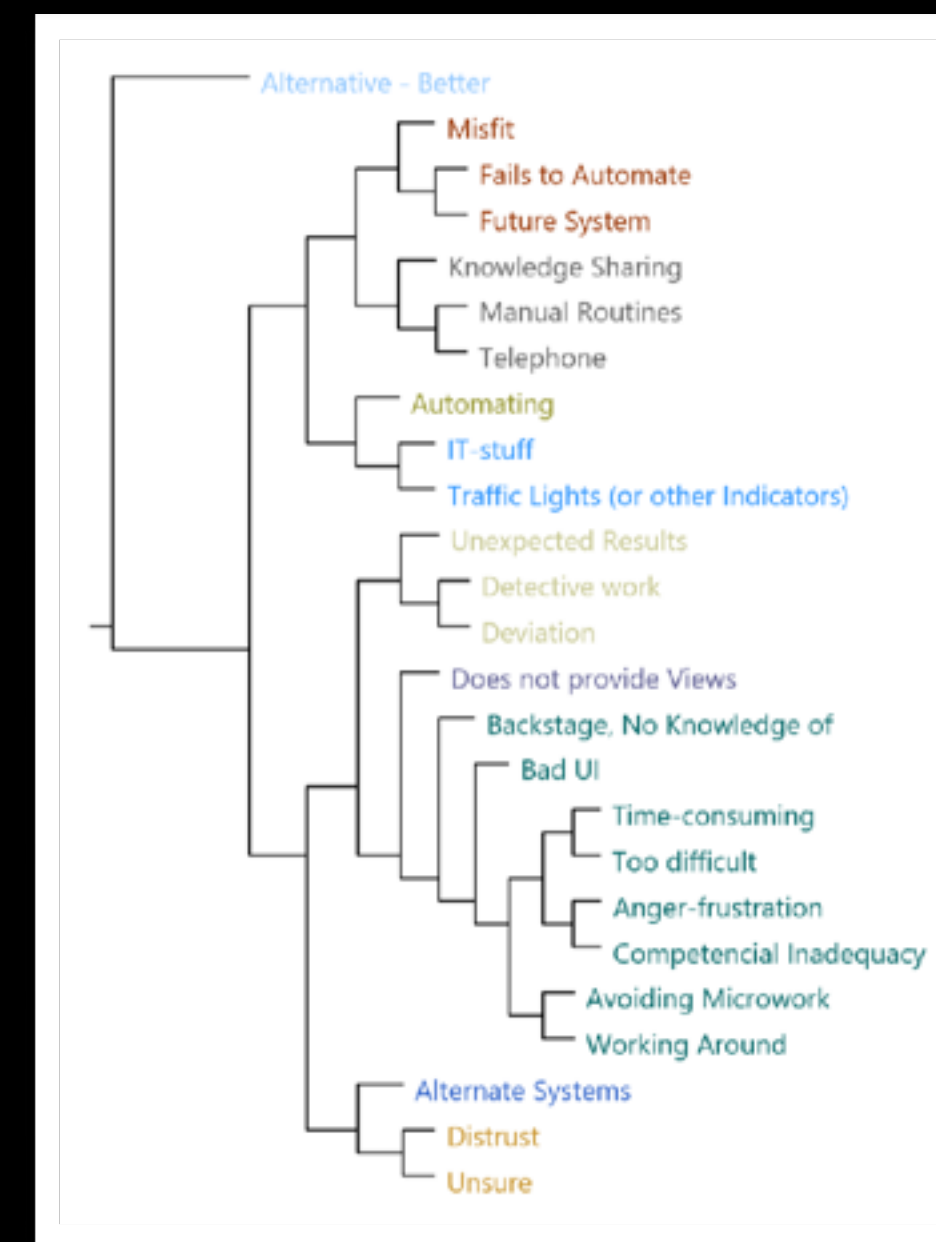
# Integrative Diagrams



# Programmatic Analysis



*Weak or no correlation*



*Strong correlation*

# Quick Comments on Codes

- Micro-coding
- Macro-coding
- Ten codes, a million codes?
- “Let the codes come to me and do not forbid them, for the Grounded Theory belongs to such as these!”
- It's easy to keep using the same codes – and dangerous
- Remember axial coding!

# MORE Quick Comments on Codes

- Codes will crystalize while coding.
- What happens to new codes you discover while coding?  
Should you go back and re-code the rest of the data in this light? Again? And again?
  - Keep going until the theory is saturated!
- You will get successively more codes with higher granularity as you code
- Go back? When is enough?

Do not group the codes too early!

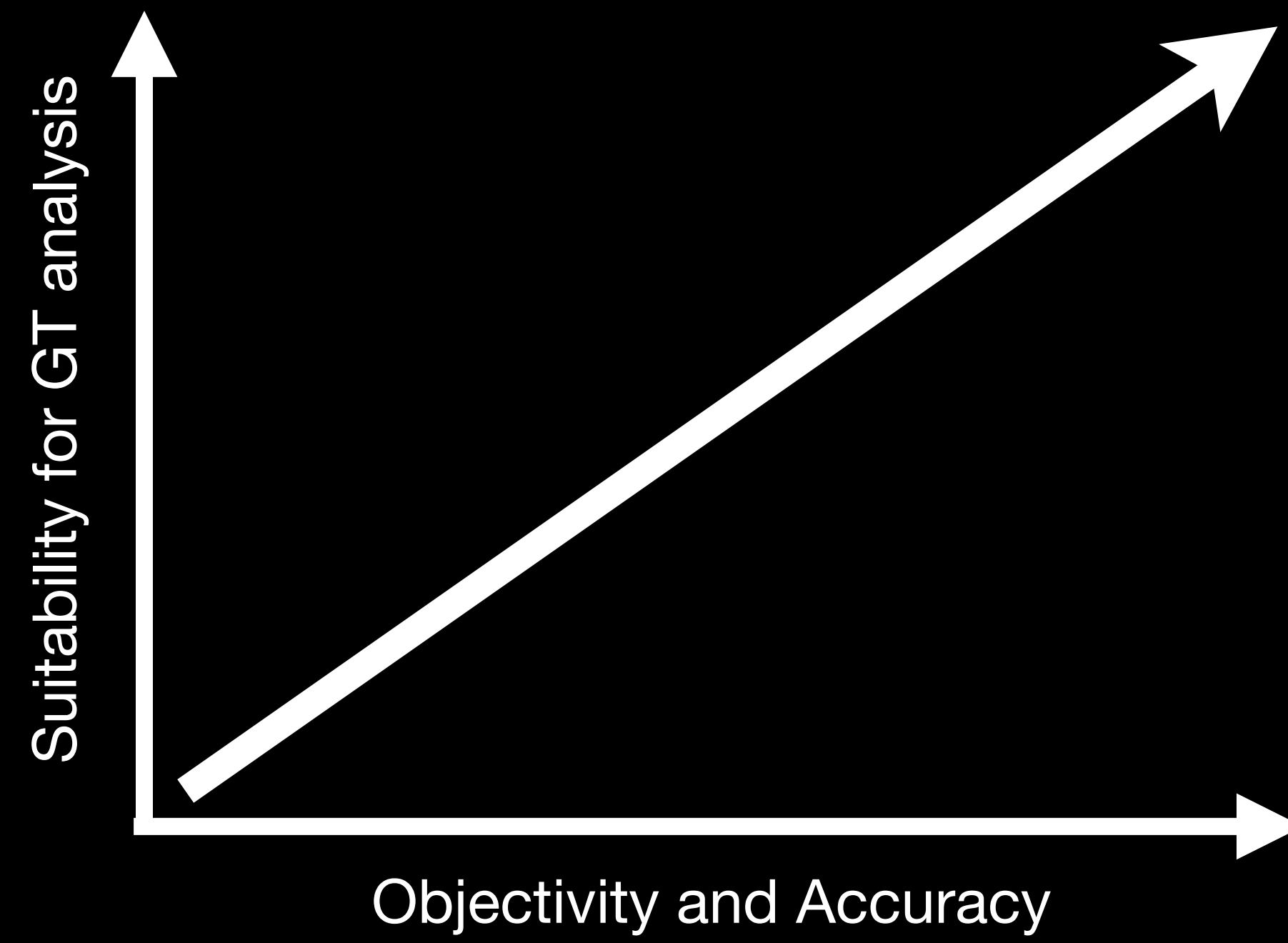
*and*

One occurrence of a particular code can be ten times as meaningful as ten occurrences of another code



# On Sources

Great Sources	Good Sources	Challenging Sources	Poor Sources
Interviews – transcribed!	Routine descriptions	Participatory/Passive Observation	Interview notes
More interviews – go back!	Job descriptions	Video	Abridged interviews
Academic texts	Internal documents, policy documents	Focus groups	
	Other unprepared texts	Prepared statements	
	Source code	Press releases, journalist work	



What about quantitative data?



You can use it – to support or contradict your findings!

# GT Pros&Cons

Pro	Con
(Usually) Great results!	Takes a lot of time
Grounded Data	Takes a lot of work
The discovery of more than the sum of your data	Can give false trails
Free styled, suitable for many sorts of outputs	Can't use all your results
Combines well with many other methodologies	Needs immersion
You don't need to know for sure what you're looking for	You can't know for sure what you're looking for

## NVivo

[English version  
of this page](#)

### Nvivo

- [Hjelp og  
veiledninger](#)

NVivo er en nokså kompleks programvare som kan brukes til å kode tekst, lyd, bilde og video. Du kan også bruke NVivo til å transkribere.

### Om NVivo

NVivo er en kompleks programvare for kvalitative forskere som blant annet kan brukes til å kode tekst, lyd, bilde og video.

NVivo er en programpakke med analyseverktøy for kvalitative forskere. Denne kan være til god hjelp i hele forskningsprosessen; fra problemformulering, organisering av ulike typer data, koding og samarbeid med andre forskere, til systematisk analyse/teoribygging og konklusjoner.

Hvis du har behov for å behandle et rikt datasett med ustrukturert informasjon (for eksempel intervjuer, logger, bilder og video), kan NVivo være hensiktsmessig. Programmet "automatiserer" en del manuelle oppgaver assosiert med kvalitativ analyse, som å klassifisere/organisere informasjon.

Claim:

In many ways, Grounded Theory can be understood as


***the data-driven generative analysis and reasoning around what meanings (language) humans attribute to structures (models)***



# Grounded Theory vs. LLMs

## Similarities

	Grounded Theory	LLMs
Data-Driven	collects and analyzes qualitative data to develop theories or explanations	process vast amounts of text data to generate responses
Iterative Process	researchers iteratively collect, code, and analyze data until they reach theoretical saturation	the model refines its responses based on continuous training on new data
Inductive Reasoning	develops theories based on observations and patterns identified in the collected data	learn patterns and generate responses based on the input data they are trained on

  
 Data bias?  
 Who decided what your LLM was taught?

## Key differences

	Grounded Theory	LLMs
Nature of Data	qualitative data obtained through interviews, observations, documents	primarily text data, including structured and unstructured text from various sources such as books, articles, and websites
Purpose and Domain	qualitative research methodology used to develop theories or explanations in social science disciplines	natural language understanding and generative tasks, such as translation, summarisation, or text generation
Automation vs. Human Interpretation	relies heavily on human interpretation and analysis of qualitative data to develop theories, requiring researchers to actively engage in coding and categorizing	automate language processing tasks using machine learning algorithms, reducing the need for human intervention in generating responses (roleplaying)
Scope of Application	specifically designed for social science research, particularly in fields such as sociology, psychology, and anthropology	broad scope, including natural language processing, text generation, and other AI-related tasks

## Which approach results in structured hallucination?

Caution: Don't try to solve GT with LLMs without first properly understanding GT!

*Green Field with Differences*

*Opportunities and applicability*

*Caution: Purpose and Ethics*



**You**

Have any scientists tried developing Grounded Theory using LLMs?



**ChatGPT**

As of my last update in January 2022, there hasn't been a widespread adoption of using Language Models (LMs) such as LLMs for developing Grounded Theory. Grounded Theory is a qualitative research methodology deeply rooted in human interpretation and analysis of data, particularly in the social sciences. While LLMs are incredibly powerful tools for processing and generating text, they lack the nuanced understanding and interpretative abilities that human researchers bring to qualitative analysis.

However, it's not entirely inconceivable that researchers may explore ways to incorporate LLMs into aspects of the Grounded Theory process. For example, LLMs could potentially assist in data preprocessing tasks, such as initial coding or identifying recurring themes in qualitative data. They might also be useful for generating initial hypotheses or assisting with literature reviews.

Still, the core aspects of Grounded Theory, such as constant comparison, theoretical sampling, and the iterative process of theory development, heavily rely on human judgment and interpretation, which may not be easily replicable by LLMs. Moreover, there are ethical considerations and concerns regarding the potential biases or limitations of using AI models in qualitative research, which would need to be carefully addressed.

While there may be exploratory efforts at the intersection of LLMs and qualitative research methodologies like Grounded Theory, it remains a relatively unexplored area compared to the extensive use of LLMs in other text-related tasks.

# Developments in GT since Strauss

- Integration of Technology (supporting work, not replacing it)
- Mixed-Methods Approaches
- Reflexivity and Transparency (researchers biases, perspectives, and the influence of their backgrounds)
- Online and Digital Environments (social media, online forums, or virtual communities)
- Critical Perspectives (importance of power dynamics, social context, and the role of language and discourse in shaping reality)
- Globalisation and Cultural Sensitivity (cultural context and diversity of participants)

**Everybody who does Grounded Theory  
will end up doing it wrong.**

And that's okay.

Wilhelm Arthur Sandberg Damsleth



**KEEP  
CALM  
AND  
CODE  
ON**