

What Is Analytic Fluency?

Matthew E. Vanaman
University of Texas at Austin

“The training I got in capital* S *statistics was so far removed from the practice of data analysis that I would have been useless in a data analytics setting. If you needed a moment generating function, then damn it, I was your guy. But if you needed a research project taken from start to finish, I wouldn't have known my butt from a hole in the ground.”

- Experienced Analyst

“Analytic Fluency”

The “soft skills” of data analysis that go beyond formal training in statistics and which are thought to come from experience analyzing data “in the real world”.

Ok, but what are those skills?

“If you have to ask, you’ll never know.”

- Miles Davis

The Present Research

Goals

1. Make analytic fluency skills explicit.
2. Exhaust the analytic fluency concept.
3. Check our intuitions.
4. Develop a grounded theory of analytic fluency.

Method

Focus Group With Experienced Analysts

Reflect back to when you were first starting out. Now that you have had all your years of experience analyzing data in the wild, what advice would you give your less-experienced self?

Focus Group Sample Characteristics

- 4 Women, 3 men
- 44.3 years old (SD = 10.1)
- 5 White, 2 Asian
- 2 Hispanic

Focus Group Sample Characteristics

- 18.4 years of data analysis experience (SD = 9.4)
- 4 Academia, 3 Private Sector

Title	N (%)
Professor	2 (28.6)
Postdoctoral Researcher	1 (14.3)
Research Scientist	1 (14.3)
Senior Data Scientist	1 (14.3)
Senior Machine Learning Engineer	1 (14.3)
User Experience Researcher	1 (14.3)

Method

Analysis

Grounded theory:

- Open coding
- Axial coding
- Selective coding

Analysis

Grounded theory:

- Open coding
- Thematic (axial) coding
- Selective coding

Analysis

Example: Theory of Contingent Hope

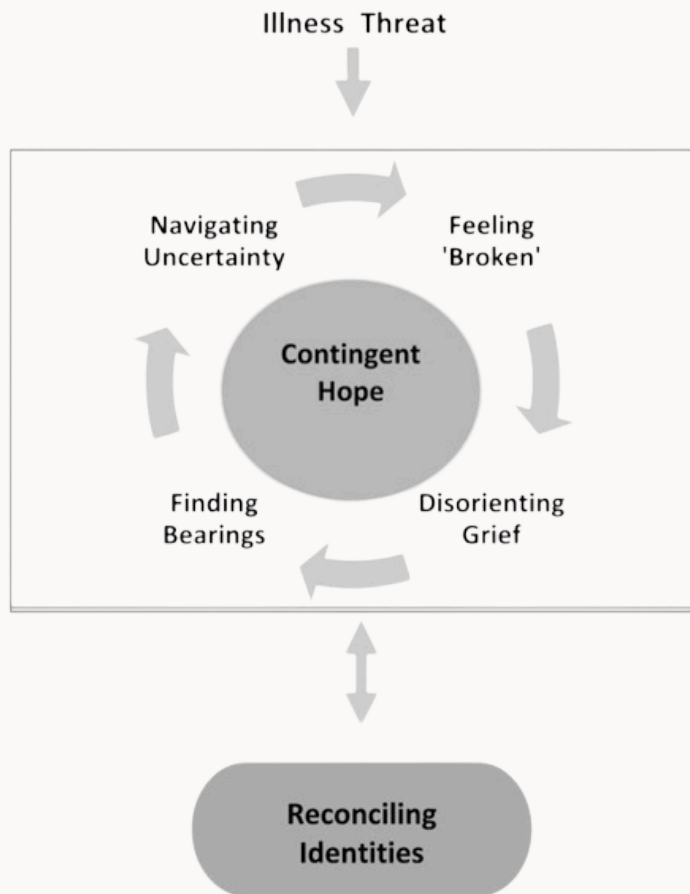


Fig. 2, Currin-Mcculloch and colleagues (2021)

Analysis

Raw Text

All nap and no hunt makes
Jack a chonk cat. All nap and
no hunt makes Jack a chonk
cat. All nap and no hunt
makes Jack a chonk cat. All
nap and no hunt makes Jack
a chonk cat. All nap and no
hunt makes Jack a chonk cat.

Analysis

Raw Text

All nap and no hunt makes Jack a chonk cat. All nap and no hunt makes Jack a chonk cat. All nap and no hunt makes Jack a chonk cat. All nap and no hunt makes Jack a chonk cat. All nap and no hunt makes Jack a chonk cat. All nap and no hunt makes Jack a chonk cat.

Open Code

Being
sedentary

Becoming
chonky

Analysis

Raw Text

All nap and no hunt makes Jack a chonk cat. All nap and no hunt makes Jack a chonk cat. All nap and no hunt makes Jack a chonk cat. All nap and no hunt makes Jack a chonk cat. All nap and no hunt makes Jack a chonk cat. All nap and no hunt makes Jack a chonk cat. All nap and no hunt makes Jack a chonk cat.

Open Code

Being
sedentary

Becoming
chonky

Theme

Lifestyle of
Laziness

The Cat In Question



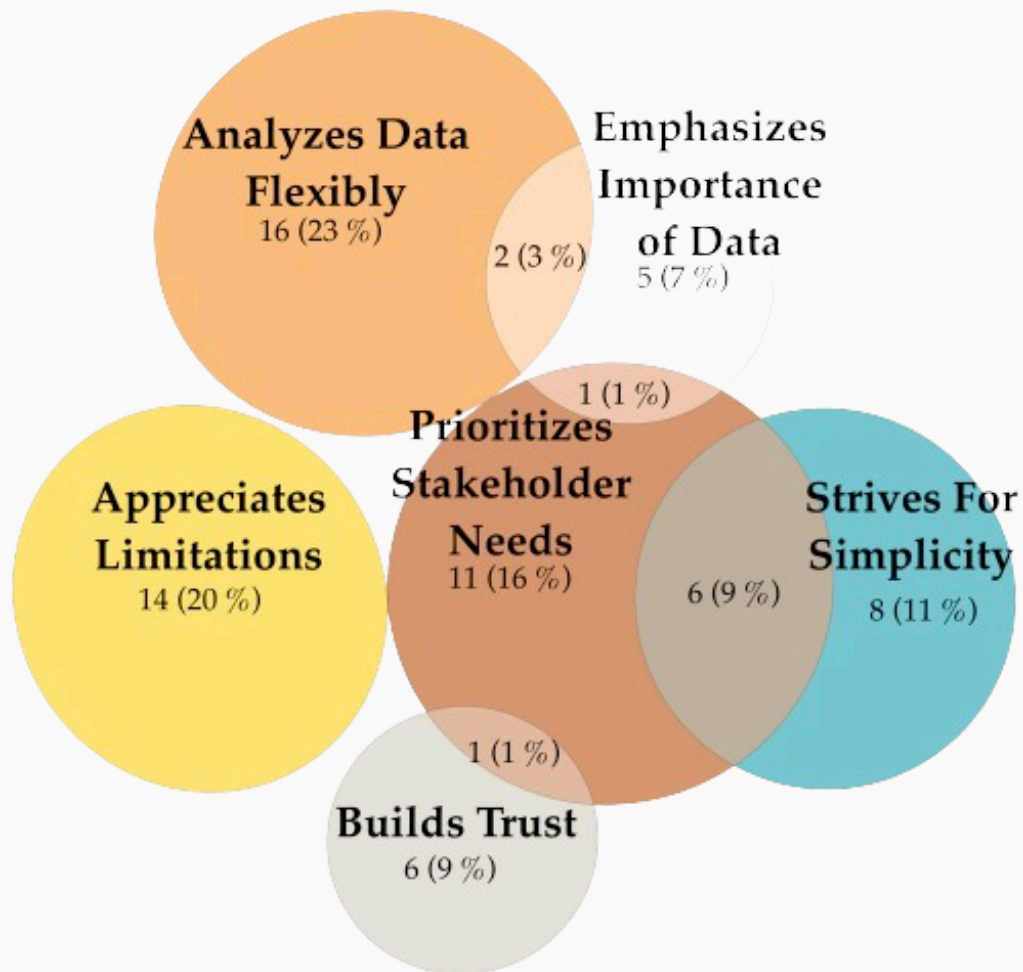
Results

Provisional Themes

Theme	Definition	Representative Quotation
Appreciates Limitations	Approaches analyses humbly, comfortable seeking outside perspectives, and is comfortable with or even enjoys uncertainty.	<i>Make usable knowledge and have a careful interpretation...[that's how I] overcome analysis paralysis.</i>
Strives For Simplicity	Prioritizes good over perfect, clarity over complexity, and strong fundamentals over unnecessary rigor.	<i>[I think] someone straight out of college might know how to do perfect science, but kind of [go] overboard with methodology.</i>
Analyzes Flexibly	Comfortable with the fluidity and subjectivity of analysis. Makes creative use of multiple approaches, software, sensitivity checks, and data processing and sourcing.	<i>If you don't get drastically different answers [from different analyses], then you're somewhat confident that [the] assumptions aren't driving your outcome.</i>
Prioritizes Stakeholder Needs	Centers entire analysis on needs of consumer(s). Allows those needs to shape data collection, analysis, visualization, and reporting.	<i>You [don't want your analysis to be] overkill [and] not understood by the people involved.</i>
Emphasizes Importance of Data	Internalizes a "garbage in, garbage out" philosophy. Patiently digests data, measurements, study design, and documentation. Considers multiple data sources.	<i>Study design, recruitment, data management, and things like that [are] where your projects go to die.</i>
Builds Trust	Never oversells conclusions, always communicates uncertainty, and proactively documents, communicates, and contextualizes decisions.	<i>The typical research paper [masks] important things like variability and exceptions to the rule. Great visualization shows what's under the hood.</i>

Note. Brackets indicate edits for readability.

Provisional Themes



Some Provisional Theorizing

Attack of the Quinians!

“Our statements about the external world face the tribunal of sense experience not individually but only as a corporate body.”

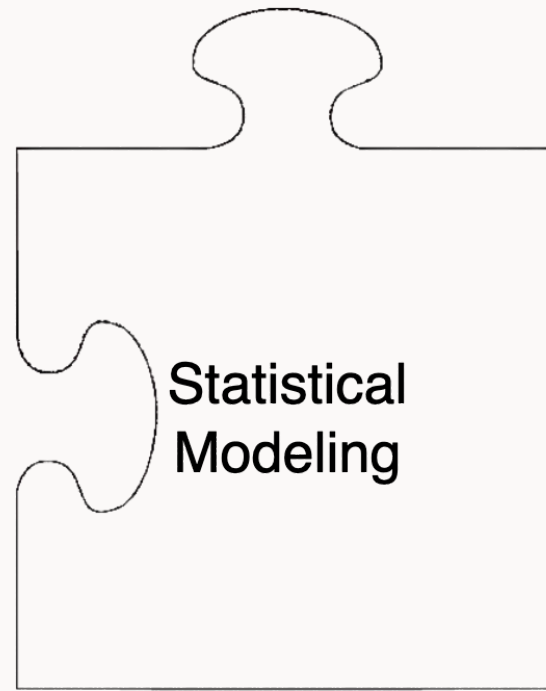
- W. V. Quine, *Two Dogmas of Empiricism*

Attack of the Quinians!

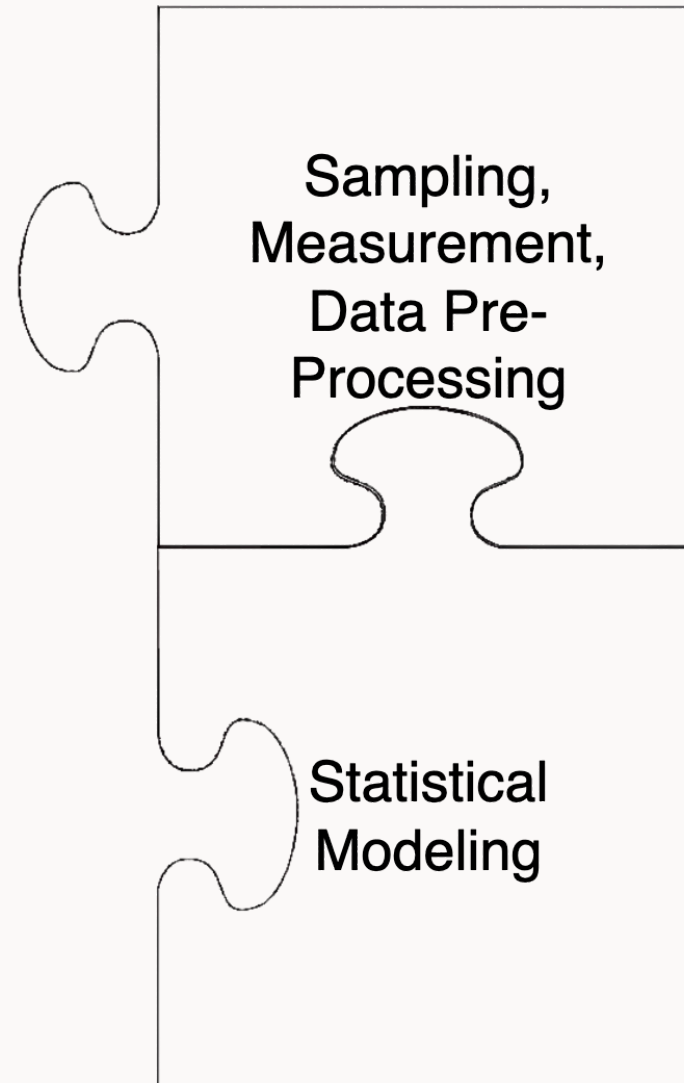
“All the steps of the research process that we go through — elements [of] design and sampling and measurement and analysis and all of those things — I think contribute to ... the replicability or reliability of the results that we have.”

- Experienced Analyst

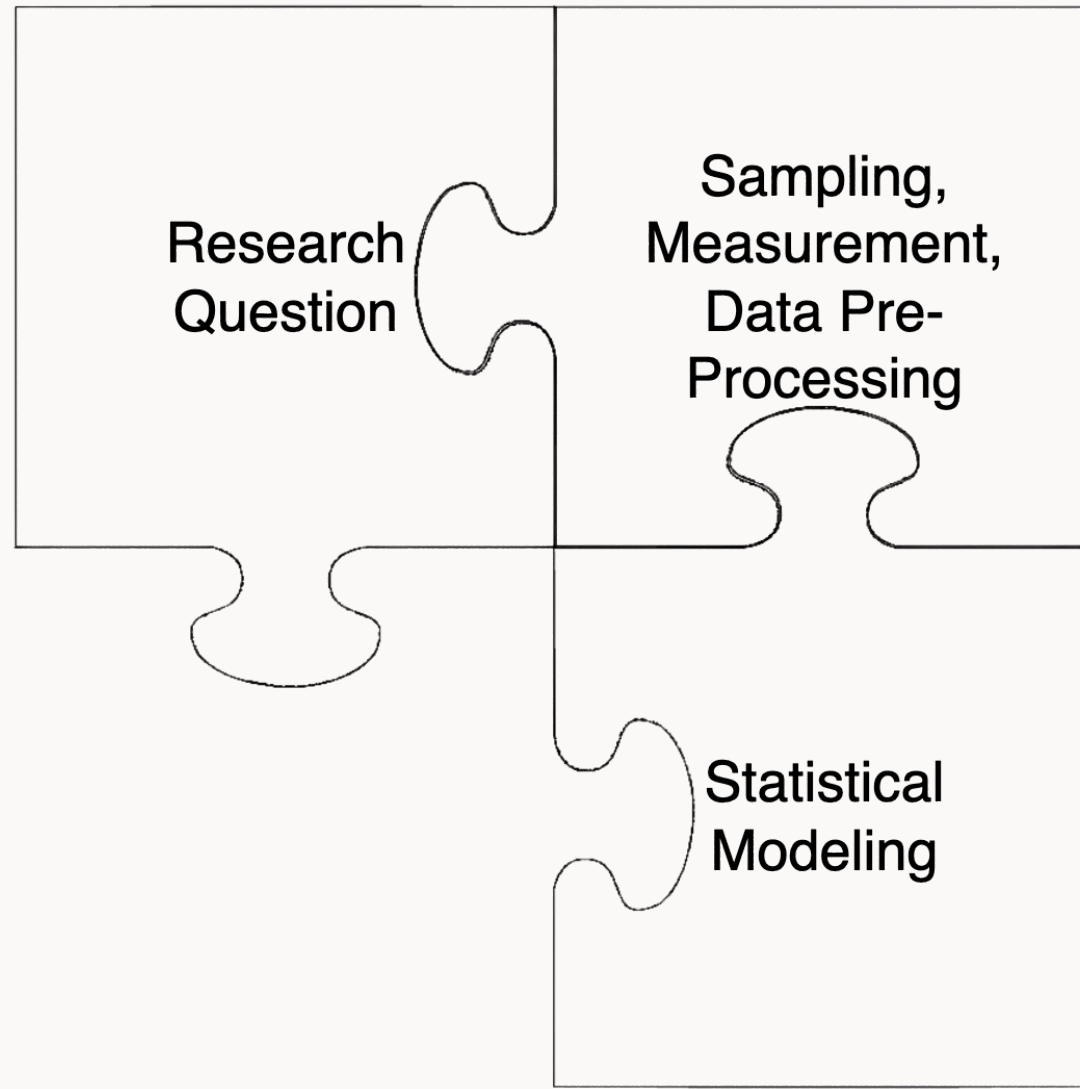
Attack of the Quinians!



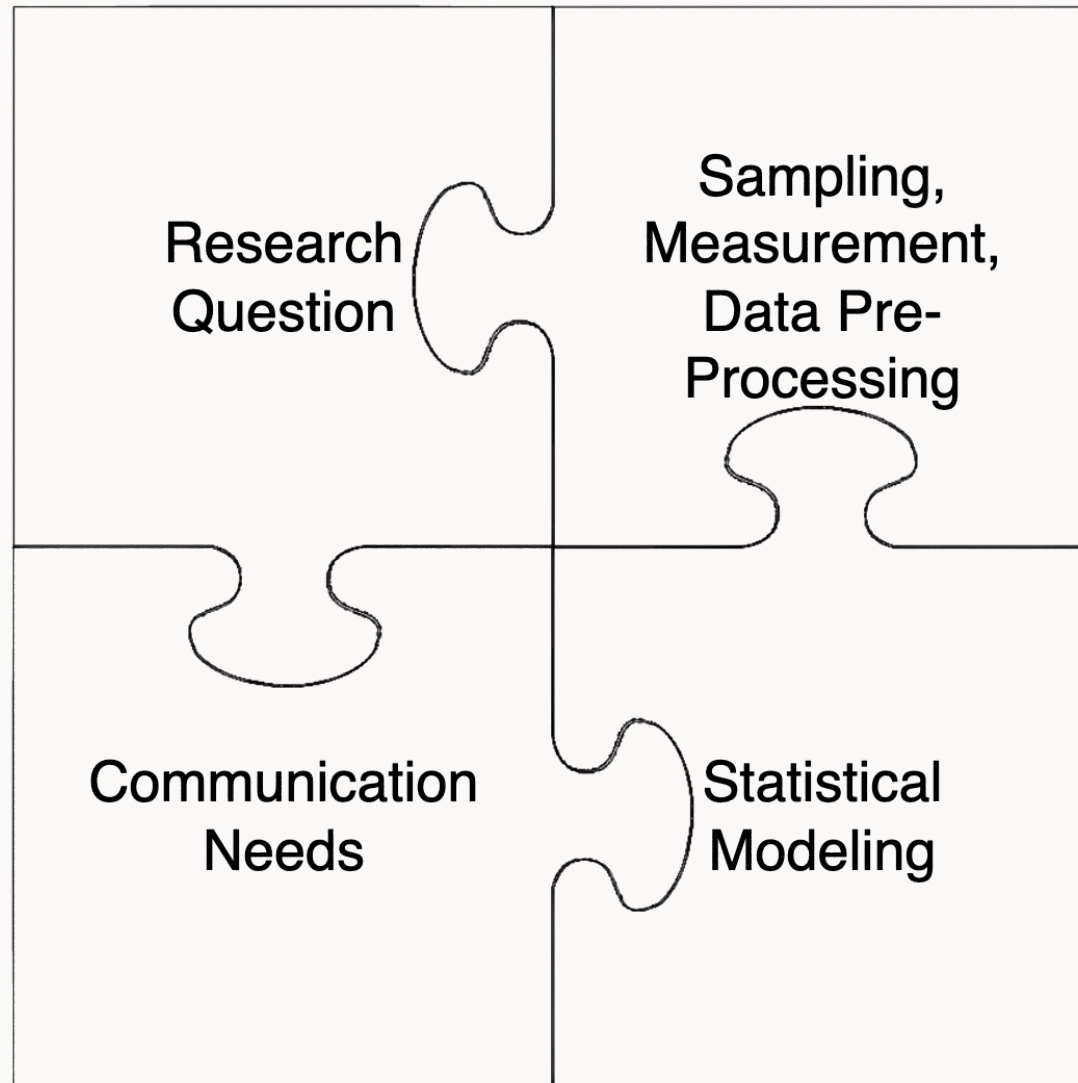
Attack of the Quinians!



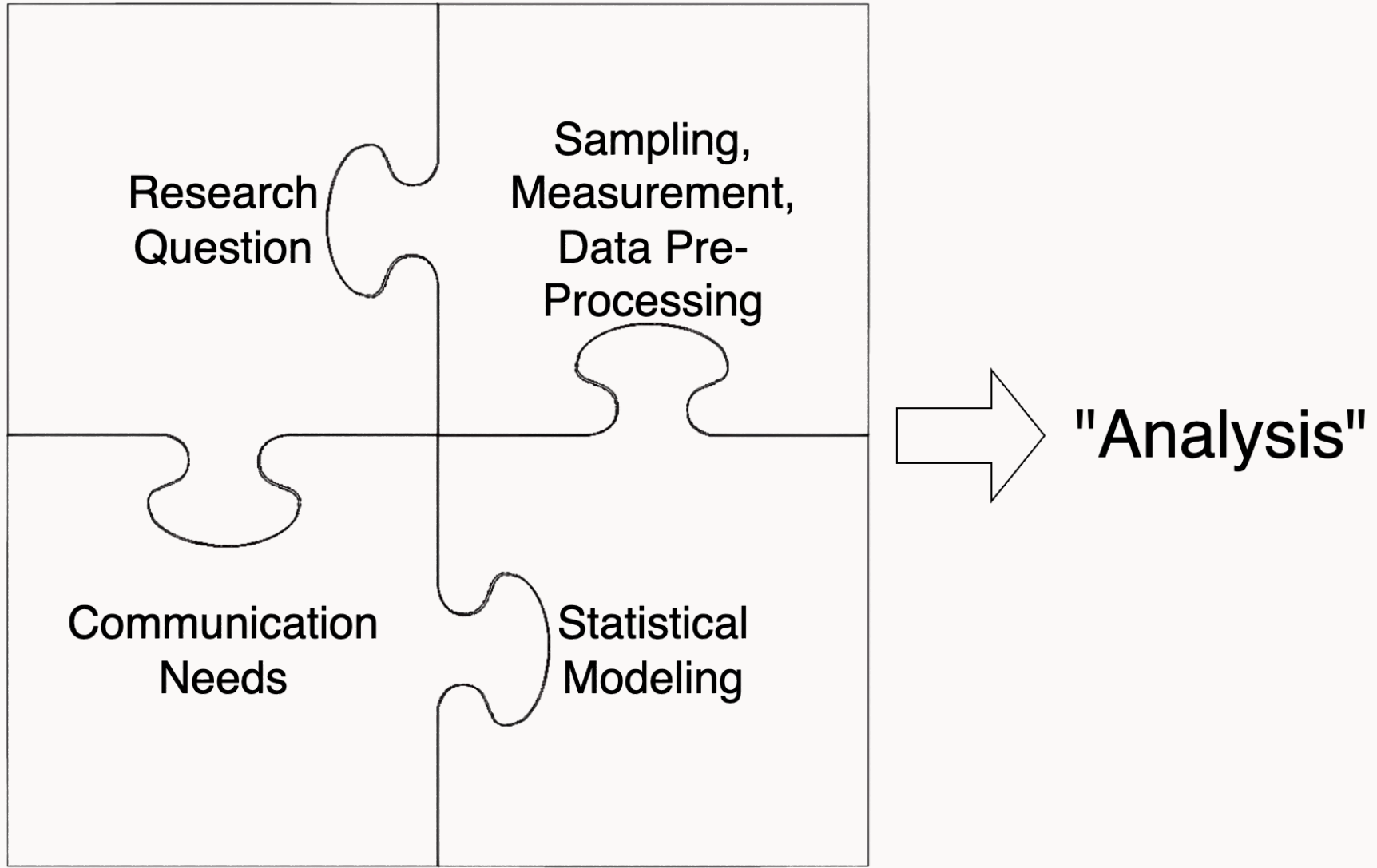
Attack of the Quinians!



Attack of the Quinians!



Attack of the Quinians!



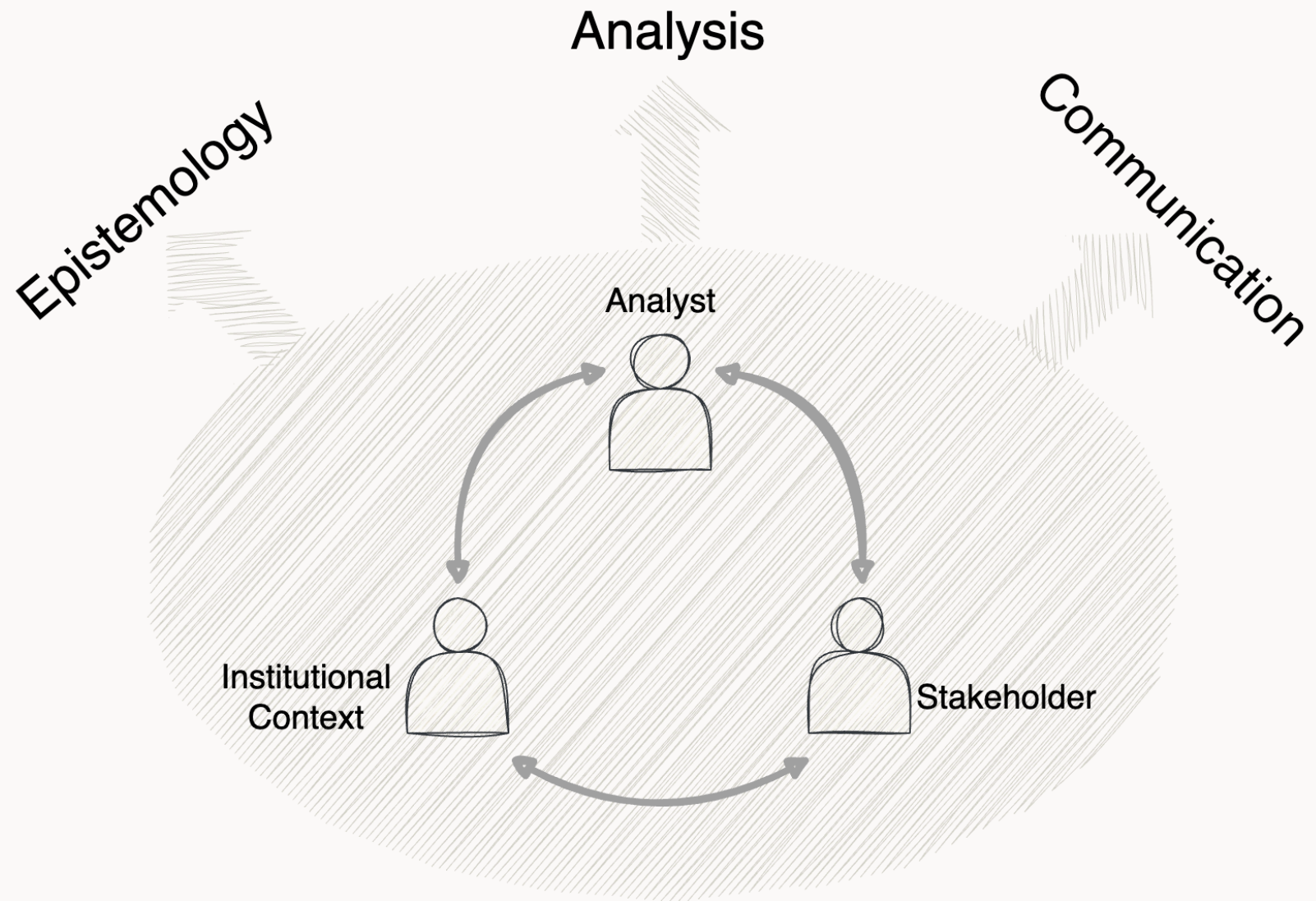
Social Construction of Data Analysis

Social Construction of Data Analysis

“Sociology of knowledge is concerned with the relationship between human thought and the social context within which it arises.”

— Peter L. Berger, *The Social Construction of Reality*

Social Construction of Data Analysis

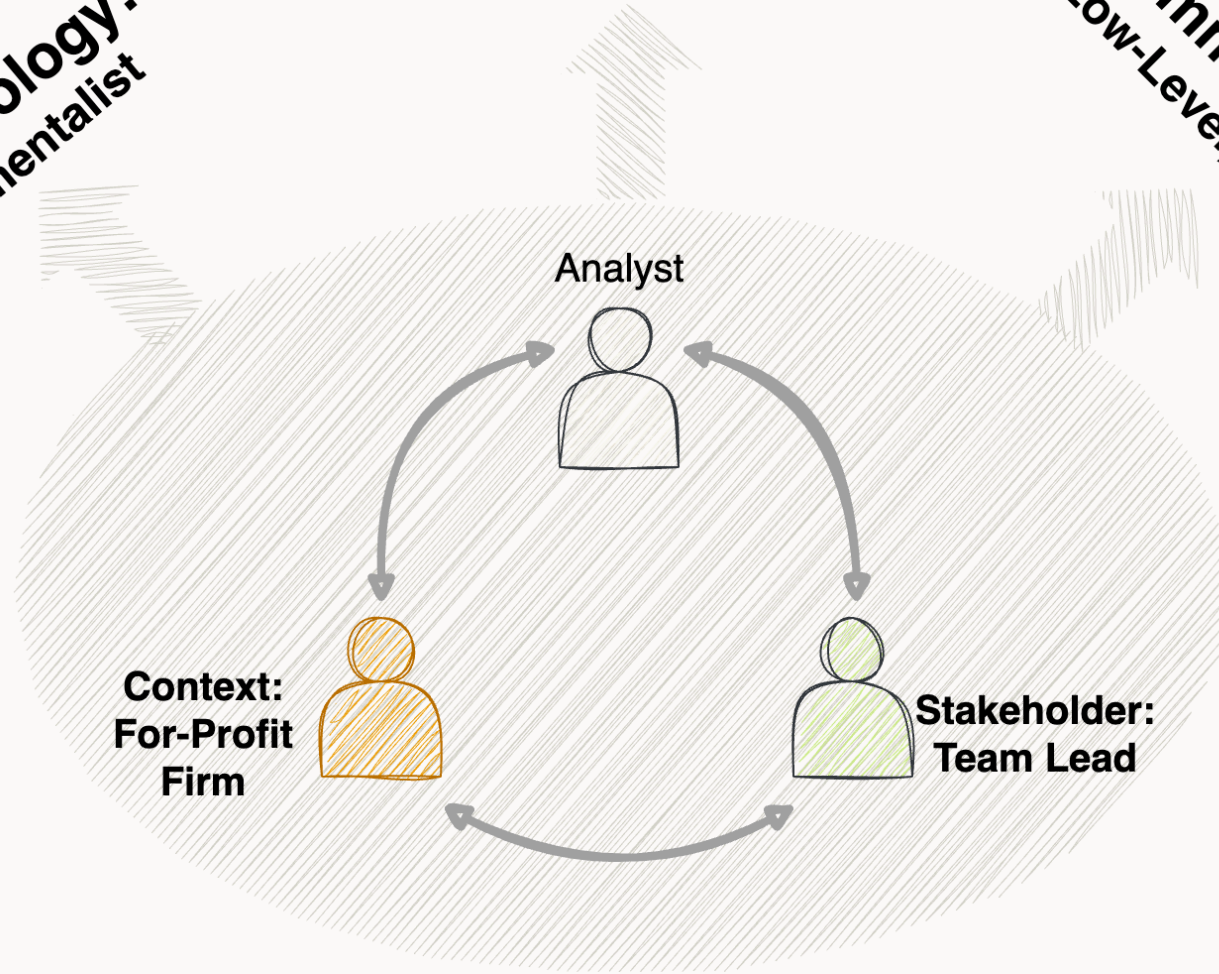


Social Construction of Data Analysis

Epistemology:
Instrumentalist

Analysis:
Predictive

Communication:
Low-Level, Technical

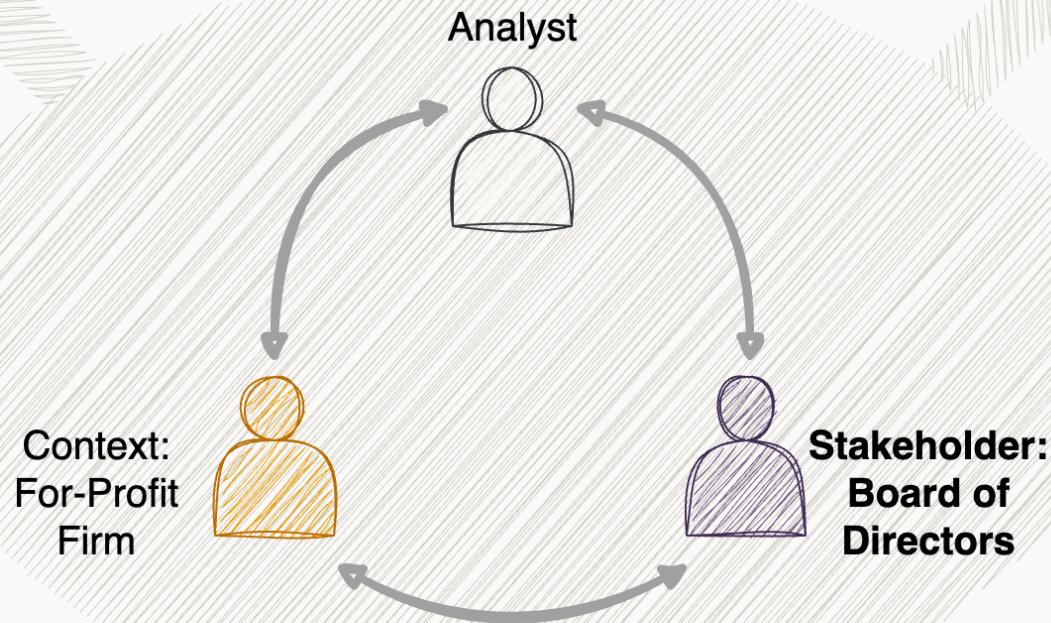


Social Construction of Data Analysis

Epistemology:
Instrumentalist

Analysis:
Predictive

Communication:
High-Level, Impact-
Focused

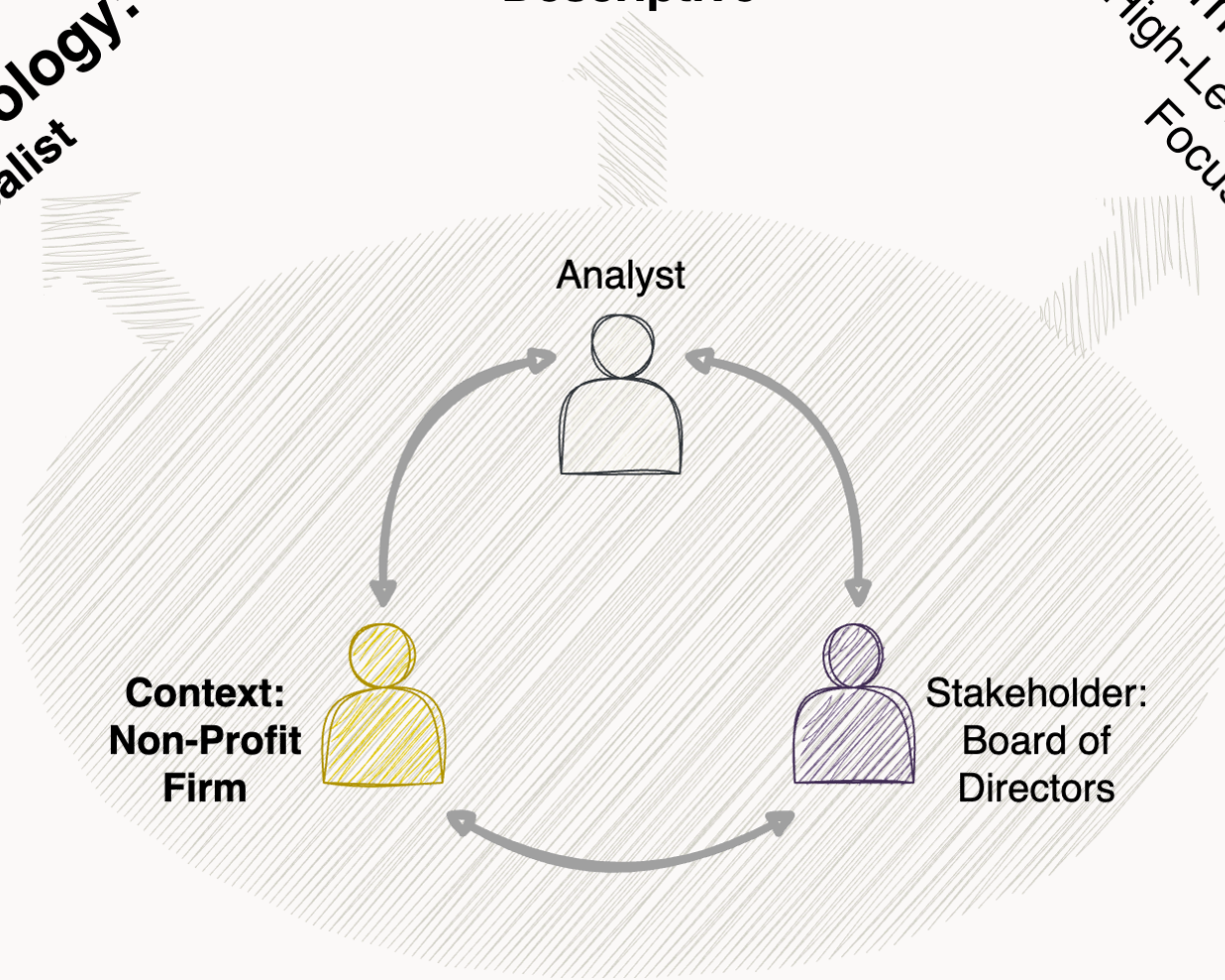


Social Construction of Data Analysis

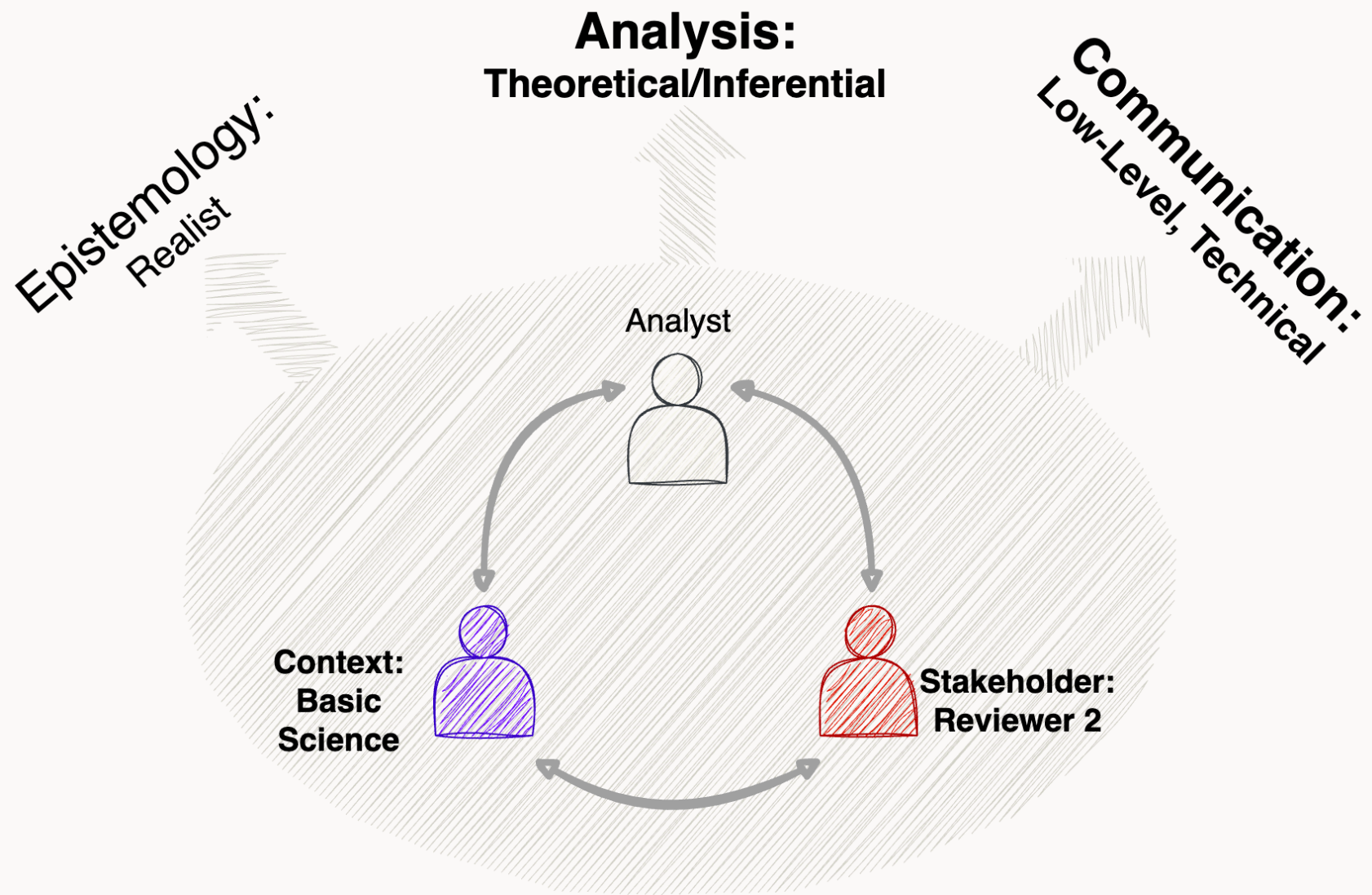
Epistemology:
Realist

Analysis:
Descriptive

Communication:
High-Level, Impact-
Focused



Social Construction of Data Analysis



Speaking of Reviewer 2

Do publication expectation punish good analysis?

Speaking of Reviewer 2

- Expectation 1: Analysis presented as only possible approach
- Expectation 2: Final decisions reported as if they were the only decisions
- Expectation 3: Creative or novel approaches are stigmatized

For Fun: Data Analysis As Aesthetic Experience

“I think probably the best memories I have are seeing a novel solution that somebody came up to answer a particular problem...seeing that solution was an aesthetic experience. You know, sort of marveling at how brilliant, how simple that was to solve what seemed like an almost insurmountable problem at the time.”

For Fun: Data Analysis As Aesthetic Experience

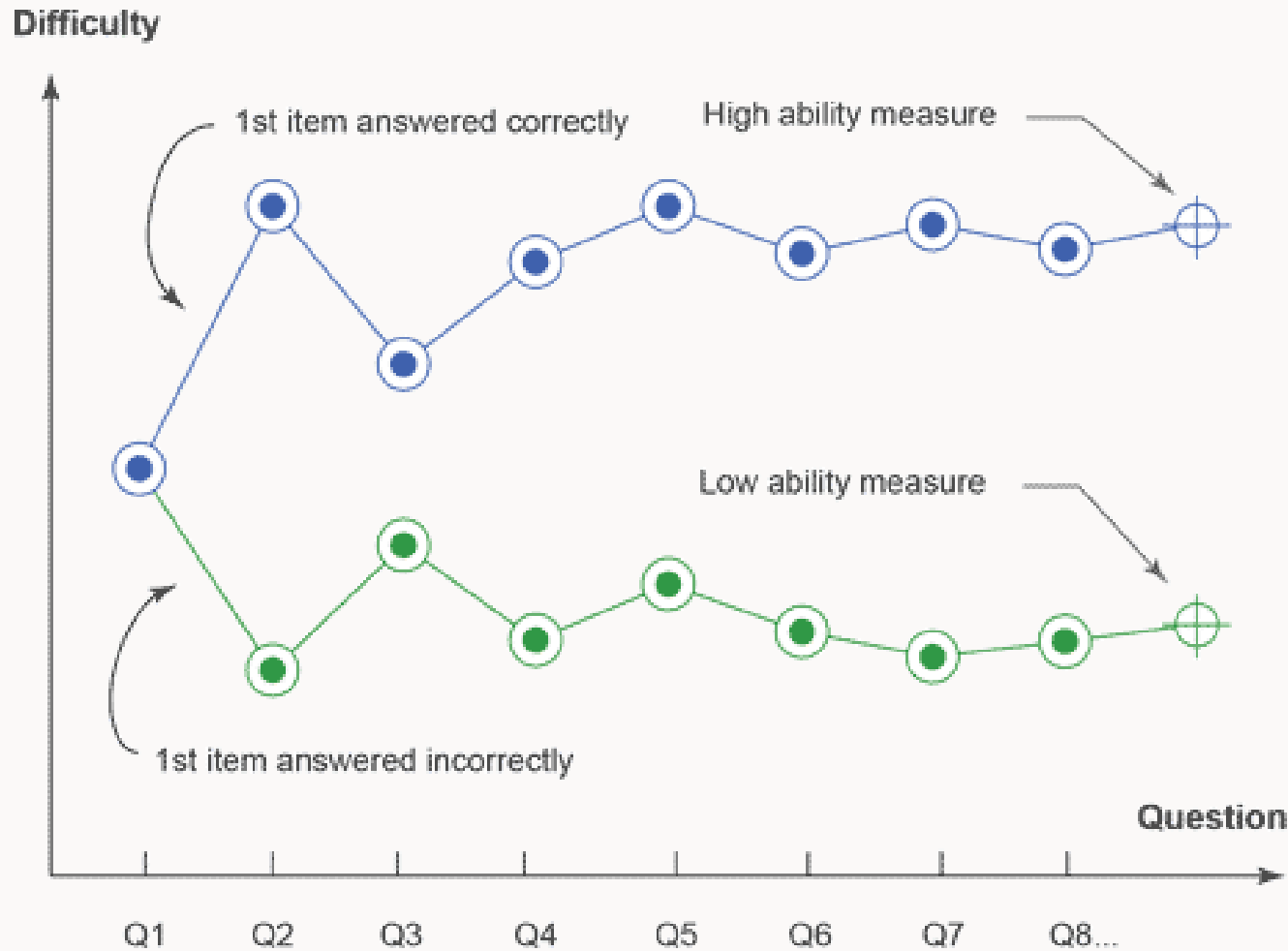
Interviewer: “And do you think that that is the kind of thing that an analyst will get better at over time?”

Participant: “Yes. It’s finding simpler and simpler solutions. Yes. Or at least more elegant solutions.”

Future

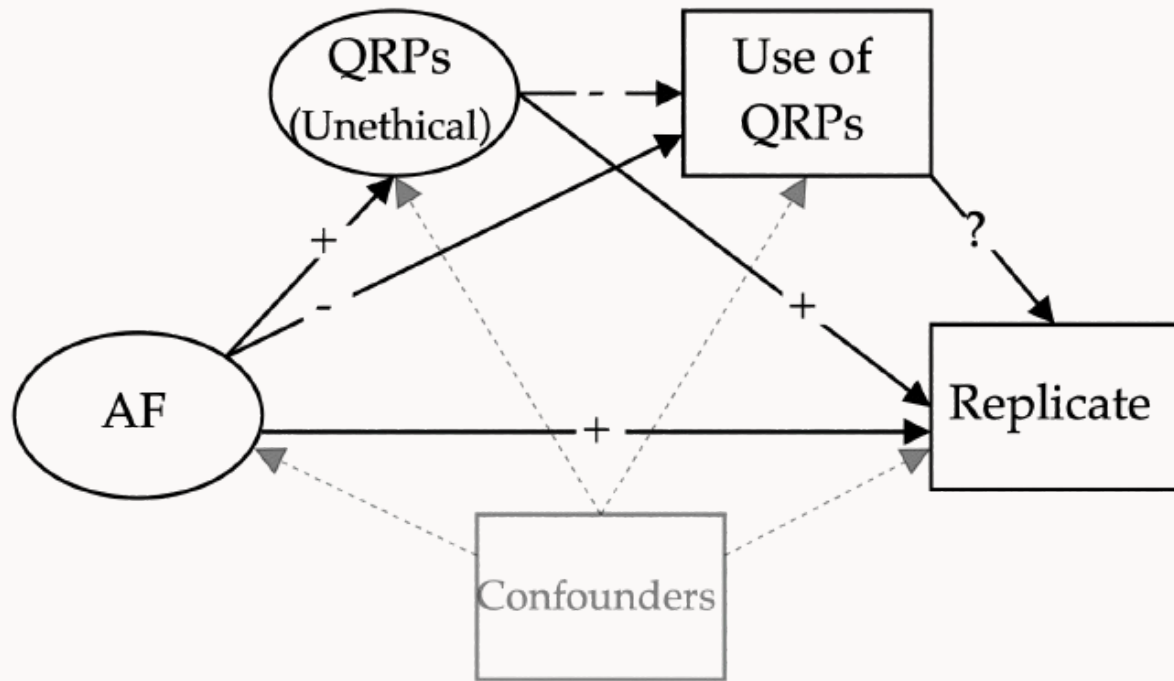
- Finalize our grounded theory
- Create analytic fluency assessment tool
- Identify its antecedents and consequences

Computer Adaptive Test Application



(Image from testpartnership.com)

Analytic Fluency and Scientific Credibility



Thank you!

Questions?

References

Currin-McCulloch, J., Walsh, C., Gulbas, L., Trevino, K., Pomeroy, E., & Jones, B. (2021). Contingent hope theory: The developmental exploration of hope and identity reconciliation among young adults with advanced cancers. *Palliative and Supportive Care*, 19(4), 437–446.

<https://doi.org/10.1017/S1478951520000656>

Quine, W. V. (1951). Two dogmas of empiricism. *The Philosophical Review*, 60, 20–43.

