Historical Methods and PEEBI Testimonial Structure for Abductive Studies in Strategy

Sandeep Pillai

September 5, 2025

sdpillai@tulane.edu

TULANE UNIVERSITY
Freeman School *of* Business

Summary

- We often pretend to do H-D; We mostly engage in Abduction

- Historical methods a necessary complement for most statistical studies:
  - Abduction requires judgments of multiple explanations
  - Cannot rely exclusively on frequentist statistics to make judgments

- PEEBI is a better structure for abductive papers

1. How we say we do research?

Hypothesis Deduction using frequentist stats

- Hypotheses based on prior theory and knowledge are offered
- Evidence of consistency is looked for in the data
- Consistency is then documented through claims about the predicted frequency of an empirical pattern in a population, conditional on a set of assumptions
- Outcome of a heated debate between John Maynard Keynes and Jan Tinbergen in the 1930's and 1940's over whether econometrics could be used to test economic theory, with Keynes suggesting it cannot and Tinbergen suggesting it can.

Challenge with H-D: Nature of claims

- Rejecting the null does not imply that H_1 is true
- Failing to reject the null does not imply that the null is true
- Forced to make truth claims based on hypothesis tests

Challenge with H-D using frequentist stats

- Limited claims as to the expectation of similar patterns in repeat samples from the same population
  - Not statements of the truth of any particular mechanism driving those patterns

- Interpretation of reported statistics as accurate predictions of a pattern in a population requires that the hypotheses, sampling plans, variables, and regression methods be specified *prior* to data collection
  - Else a frequentist statistic will understate the likelihood that a reported relationship is random.

- Developing a precise question worth answering and having the contextual understanding of the setting in which it will be tested *apriori* is hard!
  - Singular, non-trivial, RCT-derivable causes are rare

Challenge with H-D using frequentist: Summary

- Frequency claims only valid in very limited conditions
  - In Management, we rarely meet those conditions

- H-D forces us to present our work as if the entirety of the theory was imagined before we peeked at the data!

2. How we actually do research?

Abduction is usually what we do in practice

- Generate candidate explanations (H_1, H_2…, H_n)
  - Explanations: conjectures about cause-effect relationships deeply rooted in observed evidence and limited to a context
  - Theory: Cause-effect understanding that generalizes across many contexts
  - No room for explanations in H-D

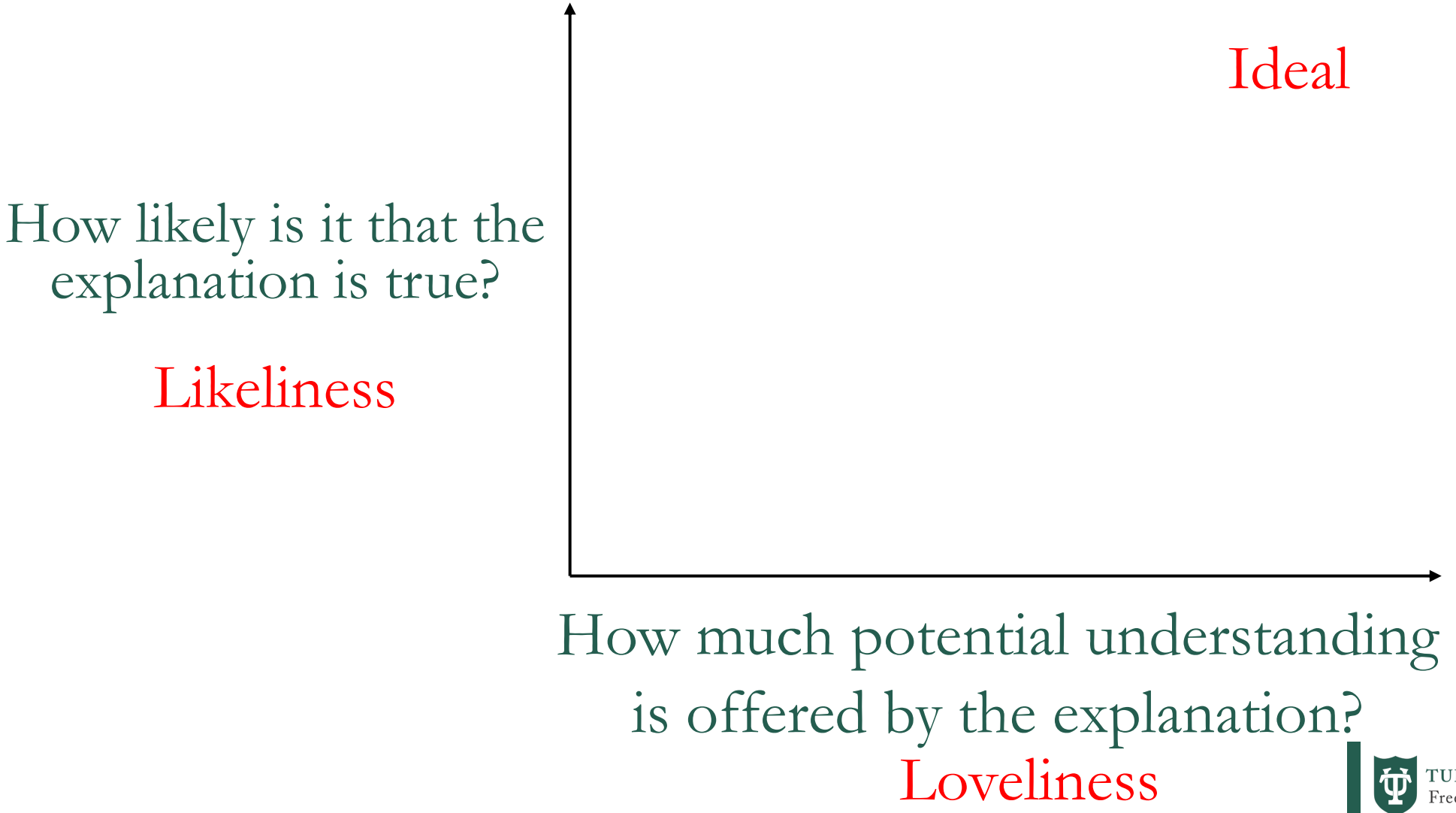
- Infer to the best explanation (IBE)
  - Select H_i

- New(ish) to Management lexicon, but not in practice

What is Abduction?

Deduction	Induction	Abduction
All humans are mortal (Theory) Sandeep is human (Data)	Sandeep is human (Data) Sandeep is mortal (Theory)	Socrates was mortal (Data) All humans are mortal (Theory)
Sandeep is mortal (Conclusion)	All humans are mortal (Hypothesis)	Socrates was human (Explanation)
-Theory and data necessarily imply the conclusion (no uncertainty)	-Ampliative: hypothesis goes beyond what is (logically) contained in the premises -Sandeep is pescatarian (theory) yields a different hypothesis	-Ampliative: explanation may be false even when data and theory are true. -Socrates could be a bird? animal? plant?

Inference requires ***judgments*** about qualities



L & L built on composite assessment of virtues

- Consilience: reconcile with known facts that are to be explained → Likely
- Precision: level of depth in the mechanism of explanation → Likely
- Parsimony: fewer assumptions, scope conditions, variables (simplicity) → lovely
- Generalizability: applicable to multiple contexts → Lovely
- Coherence: consistency with prior beliefs or theory → Likely & Lovely

How do we make judgments today?

- Loveliness = Generalizable, Fit with "my fav." literature
  - Bad-lot problem

- Likeliness = Statistical consilience
  - Same data can be interpreted as evidence for several, sometimes conflicting, findings
  - Each finding is a unique inference based on one set of assumptions
  - Frequentist claims are descriptions of fit in a sample, NOT predictions of patterns in future SAMPLES from the Population

- Useful, but insufficient to make judgements!

Claim: It is only possible to judge an explanation with a deep understanding of the context

3. Role of History

Historical methods complement statistics to improve judgement of L&L

- Go beyond "my favorite theory" and $p<0.5$
- Systematically uncover & tie together rich, contextual info. to get closer to the truth
- *"Though not proof of causation, correlation is a smoking gun; and history can often supply sufficient circumstantial evidence to convict"*

-Morck and Yeung, 2011, p.42

H.M.: collect, analyze, & interpret info about past

- Hermeneutics: actor's motivations, choices, perceptions
- Contextualization: root in actor's time and space, not researcher's
- Source criticism: creation? survival? missing? triangulation?

Judgements of Likely & Lovely using H.M.

- Hermeneutics: Cause and effect relevant to actors? Identify boundary conditions
  - E.g.: Silverman and Ingram, 2017; Pillai et al., 2019

- Contextualization: Uncovers facts to reconcile; Thick descriptions to judge generalizability
  - E.g.: Braguinsky & Hounshell, 2016; Rosenberg, 1963

- Source criticism: Triangulates across different sources; Re-interrogation of findings
  - E.g.: Khaire and Wadhwani, 2010

H.M. reveals costs, in terms of virtues, of L&L

- Contextualization identifies more facts: Increase consilience (likely), but constrains parsimony (lovely)

- Hermeneutics ensures explanations are rooted in actor's reality: Increase stat. consilience (likely), but constrains coherence (lovely)

4. How do we present the best explanation?

Challenge of testimony for an abductive paper

- It is possible for scholars to interpret the same premises and observations as supporting different explanations
- Abductive reasoning process is iterative in practice, with repeated interactions between data, theory, and explanations
- "Thicker" discussion of time, space, actors, sources
- Discussion of virtues

Our proposal: PEEBI

- **P**uzzle: A claim of contextualized facts of scholarly interest
- **E**vidence: Time, space, actors, data (facts), & analysis
- **E**xplanations: $E_1 \ldots E_n$
- **B**est: E* explains evidence better
- **I**nterpretation: Theoretical contribution

Section 1: Puzzle (Introduction)

- *Anchors* the reader with a statement of the paper's central contextual fact and a claim that they are interesting in light of prior understandings

- Overview of:
  - Set of explanations
  - Evidence that was used to evaluate explanations
  - Best explanation
  - Summarize potential theoretical contributions

- Purpose is to establish why the paper is:
  - *Lovely* by ascribing meaning to the phenomenon
  - *Likely* by summarizing explanations considered and the reason for selecting best

Section 2: Evidence

- Describe the time, space, and key actors of the study in sufficient detail

- Convince the reader that the evidence should be viewed as an accurate description of events & that the facts are relevant to the context.

- What can be ignored?
  - Present evidence to help the reader evaluate the explanatory virtues of the author's preferred explanation in comparison with other potential explanations

Section 3: Explanation

- Convince the reader that the bad lot problem has been reduced as much as possible

- In the main body present only those explanations that the author and/or the audience of readers will be expected to judge as potentially *lovely* or *likely*
  - Others in appendix

- Degree of background literature: elaborate reviews both unnecessary and impractical

Section 4: Best

- Present a preference ordering for explanations based on their *likeliness* and *loveliness*
- Though we seek both likely and lovely, as individuals, we differ in the relative weight we place on each
- Discussion of explanatory virtues - consilience, coherence, parsimony, generalizability, and precision - provides authors a systematic opportunity to discuss how they weighed the virtues
- Strongly suggest having a table with explanations and evaluation of its virtues (E.g. Kim et al., 2025)

Section 5: Interpretation

- Process of abstraction from the setting
- Detailed exposition that translates the explanation to theoretical understanding for the reader
- Boundary conditions: features of a context that are necessary for a mechanism to operate

Features of PEEBI

- Transparency of author judgments
- Focus on explanations (necessarily contextual)
- Reader can make their own judgments
- Closer to the epistemological structure of abduction

Characteristics of a bad abductive testimony

- Unconvincing Puzzle
- Too few explanations considered
- Insufficient interrogation of *all* the key considered explanations
- Poor justification of the best explanation using all explanatory virtues
- Unsatisfactory discussion of implications for higher order theory

Summary

- We often pretend to do H-D; We mostly engage in Abduction

- Historical methods a necessary complement for most statistical studies:
  - Abduction requires judgments of multiple explanations
  - Cannot rely exclusively on frequentist statistics to make judgments

- PEEBI is a better structure for abductive papers

THANK YOU

sdpillai@tulane.edu

TULANE UNIVERSITY
Freeman School *of* Business