Meta-Science, Methodology, and Time

Kevin Munger

European University Institute (EUI)

October 15, 2025

A Heap of Stones

Science is built up of facts, as a house is built of stones; but an accumulation of facts is no more a science than a heap of stones is a house.
Henri Poincaré, 1902

An Edifice of Knowledge

Once upon a time, among the activities and occupations of man there was an activity called scientific research and the performers of this activity were called scientists. In reality, however, these men were builders who constructed edifices, called explanations or laws, by assembling bricks, called facts. *Forscher, 1963*

A Pile of Bricks

A Pile of Bricks

A misunderstanding spread among the brickmakers. The brickmakers became obsessed with the making of bricks. When reminded that the ultimate goal was edifices, not bricks, they replied that, if enough bricks were available, the builders would be able to select what was necessary and still continue to construct edifices. *Forscher, 1963*

And then it came to pass that a misunderstanding spread among the brickmakers. **The brickmakers became obsessed with the making of bricks.** When reminded that the ultimate goal was edifices, not bricks, they replied that, if enough bricks were available, the builders would be able to select what was necessary and still continue to construct edifices. *Forscher, 1963*

Methodology and Metascience

- I'm drawn to "methods"

Methodology and Metascience

- I'm drawn to "methods"
- Coursework, PolMeth, "methods" tracks

Methodology and Metascience

- I'm drawn to "methods"
- Coursework, PolMeth, "methods" tracks
- Crisis of faith: method applied to social media

Methodology and Metascience

- I'm drawn to "methods"
- Coursework, PolMeth, "methods" tracks
- Crisis of faith: method applied to social media
- My interdisciplinary position de-naturalizes methods

Methodology and Metascience

- I'm drawn to "methods"
- Coursework, PolMeth, "methods" tracks
- Crisis of faith: method applied to social media
- My interdisciplinary position de-naturalizes methods
- Philosophy of science: top-down, Jupiterian

Methodology and Metascience

- I'm drawn to "methods"
- Coursework, PolMeth, "methods" tracks
- Crisis of faith: method applied to social media
- My interdisciplinary position de-naturalizes methods
- Philosophy of science: top-down, Jupiterian
- "Recently discovered physics"

Methodology and Metascience

- I'm drawn to "methods"
- Coursework, PolMeth, "methods" tracks
- Crisis of faith: method applied to social media
- My interdisciplinary position de-naturalizes methods
- Philosophy of science: top-down, Jupiterian
- "Recently discovered physics"
- Metascience: bottom-up

Methodology and Metascience

- I'm drawn to "methods"
- Coursework, PolMeth, "methods" tracks
- Crisis of faith: method applied to social media
- My interdisciplinary position de-naturalizes methods
- Philosophy of science: top-down, Jupiterian
- "Recently discovered physics"
- Metascience: bottom-up
- Hard to understand the scope of our intuitions, exacerbated by eg academic Twitter

Methodology and Metascience

- I'm drawn to "methods"
- Coursework, PolMeth, "methods" tracks
- Crisis of faith: method applied to social media
- My interdisciplinary position de-naturalizes methods
- Philosophy of science: top-down, Jupiterian
- "Recently discovered physics"
- Metascience: bottom-up
- Hard to understand the scope of our intuitions, exacerbated by eg academic Twitter
- Implicit transdisciplinarity is often simply undisciplined

Methodology and Validity

- Methodology identifies threats to *validity*

Methodology and Validity

- Methodology identifies threats to *validity*
- Shadish, Cook & Campbell; a proliferation of *validities*

Methodology and Validity

- Methodology identifies threats to *validity*
- Shadish, Cook & Campbell; a proliferation of *validities*
- But whence "validity?"

Methodology and Validity

- Methodology identifies threats to *validity*
- Shadish, Cook & Campbell; a proliferation of *validities*
- But whence "validity?"
- History of science de-naturalizes

Methodology and Validity

- Methodology identifies threats to *validity*
- Shadish, Cook & Campbell; a proliferation of *validities*
- But whence "validity?"
- History of science de-naturalizes
- Kindel dissertation traces enshrinement of "validity" theory to 1950s Harvard Psychology department and the economic monopoly on psychometric testing

Methodology and Validity

- Methodology identifies threats to *validity*
- Shadish, Cook & Campbell; a proliferation of *validities*
- But whence "validity?"
- History of science de-naturalizes
- Kindel dissertation traces enshrinement of "validity" theory to 1950s Harvard Psychology department and the economic monopoly on psychometric testing
- "Validity" is binary

Methodology and Validity

- Methodology identifies threats to *validity*
- Shadish, Cook & Campbell; a proliferation of *validities*
- But whence "validity?"
- History of science de-naturalizes
- Kindel dissertation traces enshrinement of "validity" theory to 1950s Harvard Psychology department and the economic monopoly on psychometric testing
- "Validity" is binary
- "Not to worry, officer, my ID is mostly valid"

Methodology and Validity

- Methodology identifies threats to *validity*
- Shadish, Cook & Campbell; a proliferation of *validities*
- But whence "validity?"
- History of science de-naturalizes
- Kindel dissertation traces enshrinement of "validity" theory to 1950s Harvard Psychology department and the economic monopoly on psychometric testing
- "Validity" is binary
- "Not to worry, officer, my ID is mostly valid"
- Publication, in-line citation are binary – and eternal

A Pile of Papers

For 20 years I thought my job was as a basic scientist. Publish papers and throw them over the wall for someone else to apply. I now realize that there's no one on the other side of the wall. Just a huge pile of papers that we've all thrown over.
Duncan Watts, 2019

- The "paper" both stored and circulated knowledge

The Tyranny of the .pdf

- The "paper" both stored and circulated knowledge
- The internet makes this no longer necessary!

The Tyranny of the .pdf

- The "paper" both stored and circulated knowledge
- The internet makes this no longer necessary!
- Highest-status "brick" is a 35-page pdf that advances both theory and empirics

The Tyranny of the .pdf

- The "paper" both stored and circulated knowledge
- The internet makes this no longer necessary!
- Highest-status "brick" is a 35-page pdf that advances both theory and empirics
- Obvious technical improvements for replicability, interoperability (.html! .ipynb!)

The Tyranny of the .pdf

- The "paper" both stored and circulated knowledge
- The internet makes this no longer necessary!
- Highest-status "brick" is a 35-page pdf that advances both theory and empirics
- Obvious technical improvements for replicability, interoperability (.html! .ipynb!)
- The "paper" reifies methodology and therefore validity, at the expense of generalizability / applicability

The dialectic of "methodology"

The dialectic of "methodology"

- Towards "more perfect papers"

The dialectic of "methodology"

- Towards "more perfect papers"
- Fundamentally *inward-facing*, leads to periods of hyper-refining one aspect of brick-making

The dialectic of "methodology"

- Towards "more perfect papers"
- Fundamentally *inward-facing*, leads to periods of hyper-refining one aspect of brick-making
- Rational choice, public choice, Rochester school

The dialectic of "methodology"

- Towards "more perfect papers"
- Fundamentally *inward-facing*, leads to periods of hyper-refining one aspect of brick-making
- Rational choice, public choice, Rochester school
- 1990s-2000s: cross-country regressions, statistical modelling

The dialectic of "methodology"

- Towards "more perfect papers"
- Fundamentally *inward-facing*, leads to periods of hyper-refining one aspect of brick-making
- Rational choice, public choice, Rochester school
- 1990s-2000s: cross-country regressions, statistical modelling
- Credibility revolution: identification, internal validity, careful design

The dialectic of "methodology"

- Towards "more perfect papers"
- Fundamentally *inward-facing*, leads to periods of hyper-refining one aspect of brick-making
- Rational choice, public choice, Rochester school
- 1990s-2000s: cross-country regressions, statistical modelling
- Credibility revolution: identification, internal validity, careful design
- Local answers (place/time/LATE) became the standard badge of rigor

The dialectic of "methodology"

- Towards "more perfect papers"
- Fundamentally *inward-facing*, leads to periods of hyper-refining one aspect of brick-making
- Rational choice, public choice, Rochester school
- 1990s-2000s: cross-country regressions, statistical modelling
- Credibility revolution: identification, internal validity, careful design
- Local answers (place/time/LATE) became the standard badge of rigor
- Next question: what do these *local* answers buy us outside context?

The dialectic of "methodology"

- Towards "more perfect papers"
- Fundamentally *inward-facing*, leads to periods of hyper-refining one aspect of brick-making
- Rational choice, public choice, Rochester school
- 1990s-2000s: cross-country regressions, statistical modelling
- Credibility revolution: identification, internal validity, careful design
- Local answers (place/time/LATE) became the standard badge of rigor
- Next question: what do these *local* answers buy us outside context?
- External validity the most pressing problem

- We live and make decisions in the *present*

- We live and make decisions in the *present*
- Our evidence is unavoidably from the *past*

My Contribution: Time

- We live and make decisions in the *present*
- Our evidence is unavoidably from the *past*
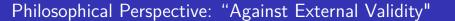- Policy and practice push us toward the *near future*

My Contribution: Time

- We live and make decisions in the *present*
- Our evidence is unavoidably from the *past*
- Policy and practice push us toward the *near future*
- Core question: how should past results inform todays world?

My Contribution: Time

- We live and make decisions in the *present*
- Our evidence is unavoidably from the *past*
- Policy and practice push us toward the *near future*
- Core question: how should past results inform todays world?
- In "internet time," this problem is paramount

My Contribution: Time

- We live and make decisions in the *present*
- Our evidence is unavoidably from the *past*
- Policy and practice push us toward the *near future*
- Core question: how should past results inform todays world?
- In "internet time," this problem is paramount
- Social science often underweights time compared to place/population

My Contribution: Time

- We live and make decisions in the *present*
- Our evidence is unavoidably from the *past*
- Policy and practice push us toward the *near future*
- Core question: how should past results inform todays world?
- In "internet time," this problem is paramount
- Social science often underweights time compared to place/population
- Why care about austere causal identification in the past if we'll accept hand-waving evaluations of "temporal validity"?

My Contribution: Time

- We live and make decisions in the *present*
- Our evidence is unavoidably from the *past*
- Policy and practice push us toward the *near future*
- Core question: how should past results inform todays world?
- In "internet time," this problem is paramount
- Social science often underweights time compared to place/population
- Why care about austere causal identification in the past if we'll accept hand-waving evaluations of "temporal validity"?
- We lack causal time series to test this; new paper with Drew Dimmery on "Generalization to the Present"

Philosophical Perspective: "Against External Validity"

Julian Reiss (*Synthese*, 2019): "External validity is normally juxtaposed with internal validity, and the former defined in terms of the latter...I maintain that an epistemic community that puts certainty before relevance will get neither (for the reasons discussed)"

"Methodological foundationalism is a flawed epistemology. I should perhaps say explicitly here that I do not accuse any philosopher of science of subscribing to a flawed epistemology. Methodological foundationalism is, however, widely held across the social sciences, and by trying to rationalise certain inferential strategies employed by scientists in my view some philosophers make themselves complicit in poor reasoning."

Slough & Tyson 2024: "a tension between design-based arguments for within-study identification of causal effects and the reliance on structural or model-based extrapolation approaches for causal generalization"

The Meta-Scientific Turn in Methodology

- Shift the unit of analysis

The Meta-Scientific Turn in Methodology

- Shift the unit of analysis
- From single studies → **literatures** that accumulate

The Meta-Scientific Turn in Methodology

- Shift the unit of analysis
- From single studies $\rightarrow$ **literatures** that accumulate
- Not perfecting the papers but improving the usefulness of the literature

The Meta-Scientific Turn in Methodology

- Shift the unit of analysis
- From single studies → **literatures** that accumulate
- Not perfecting the papers but improving the usefulness of the literature
- From statistical tools → **institutions** as methods

The Meta-Scientific Turn in Methodology

- Shift the unit of analysis
- From single studies → **literatures** that accumulate
- Not perfecting the papers but improving the usefulness of the literature
- From statistical tools → **institutions** as methods
- Peer review is a decision rule we can theorize, study, and improve

The Meta-Scientific Turn in Methodology

- Shift the unit of analysis
- From single studies → **literatures** that accumulate
- Not perfecting the papers but improving the usefulness of the literature
- From statistical tools → **institutions** as methods
- Peer review is a decision rule we can theorize, study, and improve
- The *Journal of Quantitative Description: Digital Media* is a **methodological** contribution

The Meta-Scientific Turn in Methodology

- Shift the unit of analysis
- From single studies → **literatures** that accumulate
- Not perfecting the papers but improving the usefulness of the literature
- From statistical tools → **institutions** as methods
- Peer review is a decision rule we can theorize, study, and improve
- The *Journal of Quantitative Description: Digital Media* is a **methodological** contribution
- 2025 and 2024 APSA ITP Section Award for Best Published Paper in the JQD (previously had been AJPS, JOP)

Chain of Knowledge Only As Strong As Its Weakest Link

- Metaphor from Nancy Cartwright

Chain of Knowledge Only As Strong As Its Weakest Link

- Metaphor from Nancy Cartwright
- Reality of the dialectic of methodology is that we only have so much "rigor" to evaluate given papers

Chain of Knowledge Only As Strong As Its Weakest Link

- Metaphor from Nancy Cartwright
- Reality of the dialectic of methodology is that we only have so much "rigor" to evaluate given papers
- Metaphor is unfortunately binary; better: stacked window panes, where light dispersion is multiplicative in the flaws in each layer

Chain of Knowledge Only As Strong As Its Weakest Link

- Metaphor from Nancy Cartwright
- Reality of the dialectic of methodology is that we only have so much "rigor" to evaluate given papers
- Metaphor is unfortunately binary; better: stacked window panes, where light dispersion is multiplicative in the flaws in each layer
- If we propagate error from each stage, metascientific imperative is to do methodological work on the *weakest* links

Chain of Knowledge Only As Strong As Its Weakest Link

- Metaphor from Nancy Cartwright
- Reality of the dialectic of methodology is that we only have so much "rigor" to evaluate given papers
- Metaphor is unfortunately binary; better: stacked window panes, where light dispersion is multiplicative in the flaws in each layer
- If we propagate error from each stage, metascientific imperative is to do methodological work on the *weakest* links
- "Theory" is too far from the world; we accept things that make sense to us

Chain of Knowledge Only As Strong As Its Weakest Link

- Metaphor from Nancy Cartwright
- Reality of the dialectic of methodology is that we only have so much "rigor" to evaluate given papers
- Metaphor is unfortunately binary; better: stacked window panes, where light dispersion is multiplicative in the flaws in each layer
- If we propagate error from each stage, metascientific imperative is to do methodological work on the *weakest* links
- "Theory" is too far from the world; we accept things that make sense to us
- Theoretical refinement and methodology are both in the middle – can't *see* them if the exterior panes are opaque

Chain of Knowledge Only As Strong As Its Weakest Link

- Metaphor from Nancy Cartwright
- Reality of the dialectic of methodology is that we only have so much "rigor" to evaluate given papers
- Metaphor is unfortunately binary; better: stacked window panes, where light dispersion is multiplicative in the flaws in each layer
- If we propagate error from each stage, metascientific imperative is to do methodological work on the *weakest* links
- "Theory" is too far from the world; we accept things that make sense to us
- Theoretical refinement and methodology are both in the middle – can't *see* them if the exterior panes are opaque
- First and last links: agenda-setting (with description) and prediction – both point directly to *the world*

Start at the Beginning

Start at the Beginning

- Decide what questions we want to answer

Start at the Beginning

- Decide what questions we want to answer
- Think seriously about the degree of precision with which we can hope to answer them

Start at the Beginning

- Decide what questions we want to answer
- Think seriously about the degree of precision with which we can hope to answer them
- My intuition: for most questions we want to answer, ceiling on precision is very low (Hofman, Watts et al 2017).

Start at the Beginning

- Decide what questions we want to answer
- Think seriously about the degree of precision with which we can hope to answer them
- My intuition: for most questions we want to answer, ceiling on precision is very low (Hofman, Watts et al 2017).
- The biggest "lift" is in description, qualitative or quantitative: revealing things that no one knew

Start at the Beginning

- Decide what questions we want to answer
- Think seriously about the degree of precision with which we can hope to answer them
- My intuition: for most questions we want to answer, ceiling on precision is very low (Hofman, Watts et al 2017).
- The biggest "lift" is in description, qualitative or quantitative: revealing things that no one knew
- Quantitative description broadly useful, to scientists, policymakers, fellow citizens

Start at the Beginning

- Decide what questions we want to answer
- Think seriously about the degree of precision with which we can hope to answer them
- My intuition: for most questions we want to answer, ceiling on precision is very low (Hofman, Watts et al 2017).
- The biggest "lift" is in description, qualitative or quantitative: revealing things that no one knew
- Quantitative description broadly useful, to scientists, policymakers, fellow citizens
- Progress in computational tools continues to make quantitative description more tractable and scalable

Start at the End

Start at the End

- Define decision/prediction targets *first*

Start at the End

- Define decision/prediction targets *first*
- For my interests, prediction *under intervention*

Start at the End

- Define decision/prediction targets *first*
- For my interests, prediction *under intervention*
- Register and score predictions *about the future*

Start at the End

- Define decision/prediction targets *first*
- For my interests, prediction *under intervention*
- Register and score predictions *about the future*
- Agnostic about what goes into the prediction

Start at the End

- Define decision/prediction targets *first*
- For my interests, prediction *under intervention*
- Register and score predictions *about the future*
- Agnostic about what goes into the prediction
- "The Bitter Lesson" from ML: "building in how we think we think does not work in the long run" (Sutton 2019)

Start at the End

- Define decision/prediction targets *first*
- For my interests, prediction *under intervention*
- Register and score predictions *about the future*
- Agnostic about what goes into the prediction
- "The Bitter Lesson" from ML: "building in how we think we think does not work in the long run" (Sutton 2019)
- An empirical revolution has taken place over the past 10, 5, 2 years–should consider the implications for our epistemology

Start at the End

- Define decision/prediction targets *first*
- For my interests, prediction *under intervention*
- Register and score predictions *about the future*
- Agnostic about what goes into the prediction
- "The Bitter Lesson" from ML: "building in how we think we think does not work in the long run" (Sutton 2019)
- An empirical revolution has taken place over the past 10, 5, 2 years–should consider the implications for our epistemology
- Limitations the human brain as the ultimate synthesizer of knowledge

Start at the End

- Define decision/prediction targets *first*
- For my interests, prediction *under intervention*
- Register and score predictions *about the future*
- Agnostic about what goes into the prediction
- "The Bitter Lesson" from ML: "building in how we think we think does not work in the long run" (Sutton 2019)
- An empirical revolution has taken place over the past 10, 5, 2 years–should consider the implications for our epistemology
- Limitations the human brain as the ultimate synthesizer of knowledge
- Measure usefulness by out-of-sample performance, not elegance or interestingness

On the Uniqueness of Field Experiments

- "Gold standard" claim is epistemological: confident in the exclusion restriction

On the Uniqueness of Field Experiments

- "Gold standard" claim is epistemological: confident in the exclusion restriction
- This damns them with faint praise

On the Uniqueness of Field Experiments

- "Gold standard" claim is epistemological: confident in the exclusion restriction
- This damns them with faint praise
- Ontologically unique: create counterfactual worlds

- "Gold standard" claim is epistemological: confident in the exclusion restriction
- This damns them with faint praise
- Ontologically unique: create counterfactual worlds
- When we take actions, we can learn unexpected things

On the Uniqueness of Field Experiments

- "Gold standard" claim is epistemological: confident in the exclusion restriction
- This damns them with faint praise
- Ontologically unique: create counterfactual worlds
- When we take actions, we can learn unexpected things
- American pragmatism, truth is defined as useful for taking actions that align outcomes with intentions

Field Experiments, Ethics and Democracy

- Ethics of experiments are primarily about the ethics of the actions; adding a control group does not meaningfully change the ethical stakes

Field Experiments, Ethics and Democracy

- Ethics of experiments are primarily about the ethics of the actions; adding a control group does not meaningfully change the ethical stakes
- Power imbalances exist in the world

Field Experiments, Ethics and Democracy

- Ethics of experiments are primarily about the ethics of the actions; adding a control group does not meaningfully change the ethical stakes
- Power imbalances exist in the world
- Policy implementers have an ethical *obligation* to experiment

Field Experiments, Ethics and Democracy

- Ethics of experiments are primarily about the ethics of the actions; adding a control group does not meaningfully change the ethical stakes
- Power imbalances exist in the world
- Policy implementers have an ethical *obligation* to experiment
- Don Campbell: "Social Scientist as Methodological Servant"

Field Experiments, Ethics and Democracy

- Ethics of experiments are primarily about the ethics of the actions; adding a control group does not meaningfully change the ethical stakes
- Power imbalances exist in the world
- Policy implementers have an ethical *obligation* to experiment
- Don Campbell: "Social Scientist as Methodological Servant"
- Community-driven experimentation: especially with the internet (CS, HCI)

Field Experiments, Ethics and Democracy

- Ethics of experiments are primarily about the ethics of the actions; adding a control group does not meaningfully change the ethical stakes
- Power imbalances exist in the world
- Policy implementers have an ethical *obligation* to experiment
- Don Campbell: "Social Scientist as Methodological Servant"
- Community-driven experimentation: especially with the internet (CS, HCI)
- Address the problem of generalizability by generating *and applying* local knowledge, locally

Field Experiments, Ethics and Democracy

- Ethics of experiments are primarily about the ethics of the actions; adding a control group does not meaningfully change the ethical stakes
- Power imbalances exist in the world
- Policy implementers have an ethical *obligation* to experiment
- Don Campbell: "Social Scientist as Methodological Servant"
- Community-driven experimentation: especially with the internet (CS, HCI)
- Address the problem of generalizability by generating *and applying* local knowledge, locally
- "The Experimenting Society" and "Reforms as Experiments" as Deweyan, cybernetic democracy

Beyond "More Perfect Papers"

Beyond "More Perfect Papers"

- Even out the "rigor" applied to every stage of the research process

Beyond "More Perfect Papers"

- Even out the "rigor" applied to every stage of the research process
- Appreciate our intuitions and institutions for meta-scientific reforms–but also their "scope conditions"

Beyond "More Perfect Papers"

- Even out the "rigor" applied to every stage of the research process
- Appreciate our intuitions and institutions for meta-scientific reforms–but also their "scope conditions"
- "Methodology" fails at these tasks because it's too inward-facing

Beyond "More Perfect Papers"

- Even out the "rigor" applied to every stage of the research process
- Appreciate our intuitions and institutions for meta-scientific reforms–but also their "scope conditions"
- "Methodology" fails at these tasks because it's too inward-facing
- More apertures to "the world": description (at the beginning), prediction (at the end), *action* directly in the world

Thank you!