

Investigating the Social World

Week 3: In Search for Rigour



Welcome back!

200
1823-2023 

A learning agent that acquires social norms from public sanctions in decentralized multi-agent settings

Collective Intelligence
Volume 2:2 1–14
© The Author(s) 2023
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/26339137231162025
journals.sagepub.com/home/col
 Sage

Eugene Vinitsky 
UC Berkeley, Berkeley, CA, USA

Raphael Köster, John P Agapiou, Edgar A Duéñez-Guzmán, Alexander S Vezhnevets and Joel Z Leibo 
Deepmind, London, UK

Abstract

Society is characterized by the presence of a variety of social norms: collective patterns of sanctioning that can prevent miscoordination and free-riding. Inspired by this, we aim to construct learning dynamics where potentially beneficial social norms can emerge. Since social norms are underpinned by sanctioning, we introduce a training regime where agents can access all sanctioning events but learning is otherwise decentralized. This setting is technologically interesting because sanctioning events may be the only available public signal in decentralized multi-agent systems where reward or policy-sharing is infeasible or undesirable. To achieve collective action in this setting, we construct an agent architecture containing a classifier module that categorizes observed behaviors as approved or disapproved, and a motivation to punish in accord with the group. We show that social norms emerge in multi-agent systems containing this agent and investigate the conditions under which this helps them achieve socially beneficial outcomes.

Keywords

Multi-agent systems, social norms, reinforcement learning



Wrapping up the loose ends

- ▶ Varieties of "theory" in social research: They contrast each other but (logically and intellectually) related to each other
 - Theory as paradigm and approach (Geddes 2003)
 - Theory for interpretation, explanation and emancipation (Cornelissen et al 2021)
- ▶ Key takeaway messages and tasks
 - Explain your analytical decisions as a researcher (but no need to be apologetic)
 - Recognize and appreciate the contribution of others
 - Search for the common ground for fruitful collaborations; avoid dogmatic feud and squid game

Critical Race Theory: A Brief History

How a complicated and expansive academic theory developed during the 1980s has become a hot-button political issue 40 years later.

 Share full article



 130



Thousands gathered in Washington on Aug. 28, 2020, to commemorate the historic March on Washington and the Rev. Dr. Martin Luther King Jr.'s "I Have a Dream" speech on that date in 1963. Jason Andrew for The New York Times

CRT as an interpretative theory



CRT is “ ‘is a way of seeing, attending to, accounting for, tracing and analyzing **the ways** that race is produced,’ she said, ‘**the ways** that racial inequality is facilitated, and **the ways** that our history has created these inequalities that now can be almost effortlessly reproduced unless we attend to the existence of these inequalities.’”

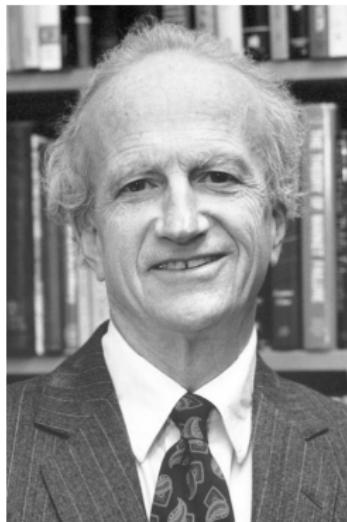
CRT as an emancipatory theory



“ ‘The founders of CRT critiqued liberal ideologies, and that they called on research scholars to seek out and **understand the roots of why racial disparities are so persistent, and to systematically dismantle racism.’ ”**

Explanatory theory of inter-group inequality

Explanatory theory of inter-group inequality



- ▶ **Statistical discrimination** is based on stereotypes that occur when rational, information-seeking decision makers use aggregate group characteristics to evaluate individuals
- ▶ **Taste-based discrimination** is based on prejudice or racial and gender bias

Breaking the ice: Where we are now

- ▶ Recap: To **produce valid knowledge claims** (cool, but how?)

Breaking the ice: Where we are now

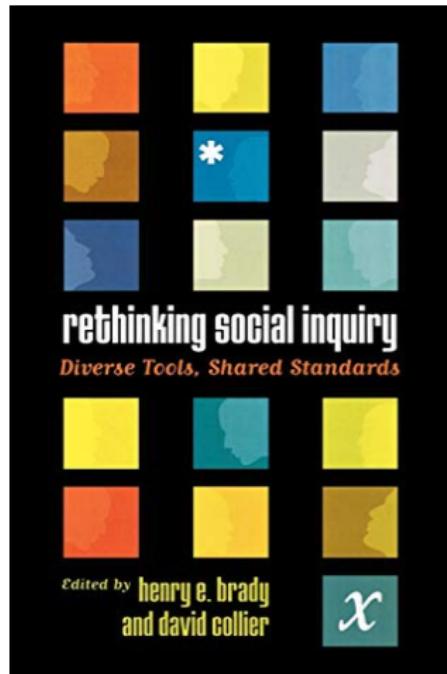
- ▶ Recap: To **produce valid knowledge claims** (cool, but how?)
- ▶ Coming weeks: "**Empirics comes to rescue**" (?)
 - How do social researchers use theory and data in their work?
 - Why should we use numbers to understand the social world?
 - How can we put our research in a broader theoretical and social contexts?

Gary King / Robert O. Keohane / Sidney Verba

Designing Social Inquiry



Scientific Inference in Qualitative Research



Part II



Building rigor in social research by aligning theory and empirics

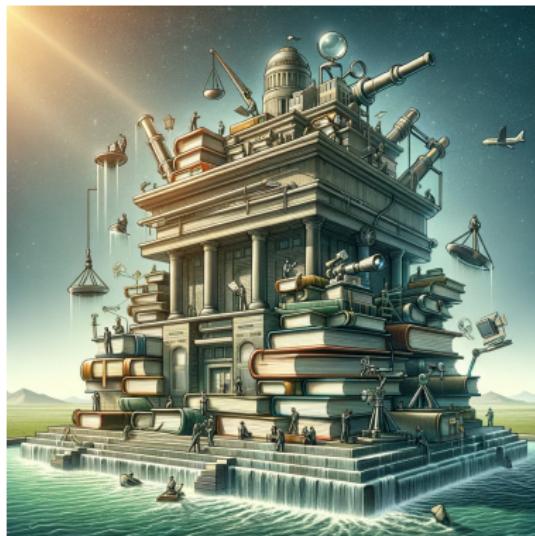
200
1823-2023 

What is rigor according to AI



- ▶ The image showcases a large, fortress-like structure made of books and **scientific instruments**, symbolizing the strength and solid foundation of rigorous research.
- ▶ This structure is encircled by a moat of flowing **data and statistics**, highlighting the **depth and thoroughness** required in research.

What is rigor according to AI



- ▶ Atop the fortress, researchers are depicted using telescopes, microscopes, and computers, representing their **meticulous and detailed** work.
- ▶ The clear and bright sky above the structure suggests the **clarity and enlightenment** achieved through rigorous study.

- ▶ Why we should pay attention to empirics
- ▶ How theory and empirics can be aligned
- ▶ Epistemology and theory-empirics alignment

Why empirics: Tradeoff of theory-building

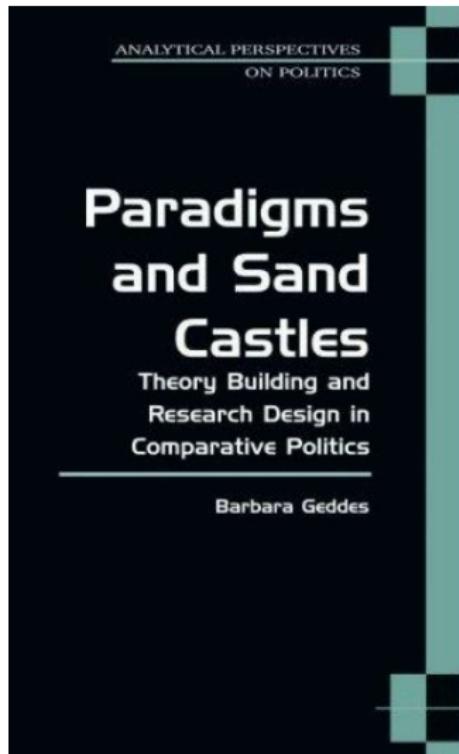
► Theory as interpretation

- Pros: Generate ideas and concepts to reveal nuances of the social world
- Cons: "Everyone is right" – a collection of opinions and perspectives

► Theory as explanation

- Pros: Lay out a specific and parsimonious process and mechanisms through which the social world "works"
- Cons: "Everyone loses" – domination and loss of complexities

Why empirics: The “sandcastle” pitfall



“ ‘My name is Ozymandias, king of kings: Look on my works, ye Mighty, and despair!’ Nothing beside remains. Round the decay of that colossal wreck, boundless and bare The lone and level sands stretch far away.”

How theory aligns with empirics

► Induction

- Producing broad claims based on our observations showing systematic patterns
- Prone to any unfair or unwarranted rejections with only an (or a few) "exception" or "outlier"
- Better for showing the existence a "theoretical" possibility awaiting further research

► Deduction

► Abduction

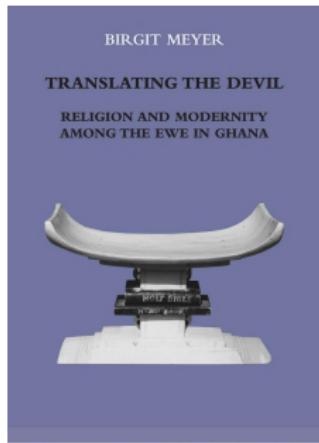
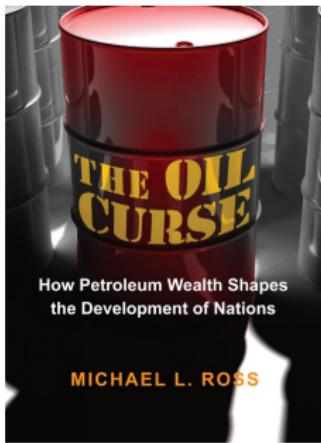
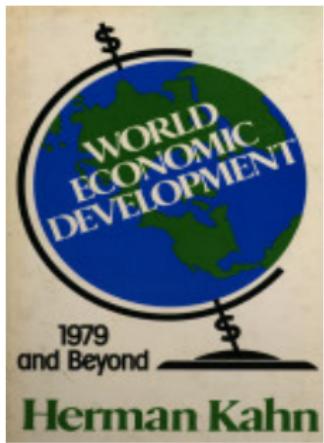
How theory aligns with empirics

- ▶ Induction
- ▶ Deduction
 - Evaluating theoretical claims by using data to "test" the hypotheses based from these claims
 - Ideal if attempting to adjudicate competing theoretical claims; often in line with positivism
 - Rejecting hypotheses does not necessarily rule out the theory (but it does reduce our faith in the theory, however)
- ▶ Abduction

How theory aligns with empirics

- ▶ Induction
- ▶ Deduction
- ▶ **Abduction**
 - Starting by (surprisingly) finding that a hypothesis does NOT work well and may trigger the need for new data and interpretations/explanations
 - Can be seen as a sequential integration of deductive-inductive processes
 - Common for the task of theoretical refinement (or improvement)

Examples from “culture” and development



- ▶ Research question?
- ▶ Theoretical claims or arguments?
- ▶ Empirical strategy (and epistemology)?
- ▶ Contribution and significance?

Induction: “World Economic Development: 1979 and Beyond” (Khan 1979)

- ▶ **Research question:** Why did NICs in East Asia achieve high economic growth rates between the 1950s and 1990s?
- ▶ **Theoretical claims or arguments:** These countries were economically successful because of their "Asian values"
- ▶ **Empirical strategy:** (Descriptive) case studies mostly
- ▶ **Significance and contribution**
 - No longer popular (or accepted) due to the lack of empirical support and conceptual setbacks (circular and essentialist)
 - Critically reflecting on the relationship between culture and development for explanatory purpose

Deduction: “The Oil Curse” (Ross 2012)

- ▶ **Research question:** Why are women in the Middle East politically marginalized?
- ▶ **Theoretical claims or arguments:** Petroleum perpetuates patriarchy because oil creates a specific industrial structure that reduces women's participation in the labor market
- ▶ **Empirical strategy:** (Inferential) multiple regression to test the following hypotheses:
 - A rise in the value of oil production will reduce female participation in the labor force
 - A rise in the value of oil production will reduce female political influence
- ▶ **Significance and contribution:** Pushing us to think away from only using Islam as the explanation

Abduction: “Translating the Devil” (Meyer 1999)

- ▶ **Research question:** How have Christian beliefs been accepted by local people in former European colonies in Africa?
- ▶ **Theoretical claims or arguments:** Locals "appropriated" Christian beliefs and transformed Peki Ewe through the (re)imagination of the notion of devil
- ▶ **Empirical strategy:** Using archival research (in Germany and the UK) and ethnography (in Ghana) were employed to document various lived experiences as an outsider from Europe
- ▶ **Significant and contribution:** Enrich Weber's theoretical framework by considering a more complicated, nuanced process of religious conversion by focusing on the agency of local communities

Concluding remarks

- ▶ Alignment between theory-empirics provide the source of rigor in social research
- ▶ Bolster the (credibility) of theory through the establishment theory-empirics alignment and *high-quality* evidence
 - Why is it a good idea to use "numbers" to study the social world (even if you do not know quants)?
 - Why is causality important? Is causality always a must?
 - How can we generalize our insights beyond a case?
- ▶ Bring epistemology into the discussion: **How rigid are these epistemology-methodology links?**
 - **Positivism:** Theory as explanation; quantitative
 - **Constructivism** Theory as interpretation; qualitative

	<i>A positivist approach</i>	<i>A constructivist approach</i>
What is ‘reality’?	A definable ‘reality’ or ‘truth’ exists and is observable	There is no ‘reality’ or ‘truth’ beyond our experiences
What is the goal of academic enquiry?	Acquisition of the ‘truth’	A more informed construction of the world
How are the researcher and the ‘researched’ related?	The researcher is independent of the ‘researched’	The researcher is not independent of the ‘researched’
What should be the role for values?	None - objectivity sought	Part of ‘reality’ - subjectivity celebrated
What kind of approach?	Predominantly based on observability or measurability and with the aim of seeking ‘evidence’	Predominantly based on discourse and meaning with the aim of seeking a more informed understanding of the world
What kind of data is preferred?	Predominantly quantitative	Traditionally associated with a predominantly qualitative approach
Examples of such studies in Development Studies	Dollar and Kraay (2002) <i>Growth is Good for the Poor</i>	Narayan <i>et al.</i> , (2002) <i>Voices of the Poor</i>

Discussion: Two worlds of development studies (really?)

- ▶ Do you agree with the binary distinction between positivism and constructivism?
- ▶ Does your research fall into one of these epistemological stances? How do you define your epistemological stance?
- ▶ Do you agree with the alignment between epistemological positions and the choice of methods?
- ▶ Is such alignment always the case? Can we be somewhere in between? How do we know when a research work is rigorous?

	Economics	Politics	Sociology	Social Anthropology
What is reality?	One reality exists; reality is <i>independent</i> of our thoughts; what is observable is real	One reality exists. reality is <i>independent</i> of our thoughts	One reality exists; reality is <i>independent</i> of our thoughts, but much of reality is unobservable	There are <i>different</i> realities associated with different standpoints and cultures
Goal of enquiry?	Acquisition of a <i>single</i> truth -- a universal, general law	We can establish truths or <i>generalizations</i> about human beings	Truth needs to be understood in terms of <i>practical adequacy</i>	<i>Interpretation</i> of local meanings; there is <i>no universal truth</i>
How are the researcher (you) and the "researched" related?	The researcher is <i>objective</i> and is <i>independent</i> of the "researched"	The researcher is <i>subjective</i> and is <i>not independent</i> of the "researched"	The researcher is <i>subjective</i> and is <i>not independent</i> of the "researched"	The researcher is <i>subjective</i> and is <i>not independent</i> of the "researched"



Thank you!



c.cheng@bbk.ac.uk

200
1823-2023 